



การออกแบบและพัฒนาระบบมองเห็นสำหรับระบบทำสวนยางพาราอัตโนมัติ

Design and Development of Machine Vision for
Automatic Rubber Latex Farming System

รัฐชัย วงศ์ธนวิจิต

Rattachai Wongtanawijit

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา
ปรัชญาดุษฎีบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์
มหาวิทยาลัยสงขลานครินทร์

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Computer Engineering
Prince of Songkla University

2563

ลิขสิทธิ์ของมหาวิทยาลัยสงขลานครินทร์



การออกแบบและพัฒนาระบบมองเห็นสำหรับระบบทำสวนยางพาราอัตโนมัติ

Design and Development of Machine Vision for
Automatic Rubber Latex Farming System

รัฐชัย วงศ์ธนวิจิต

Rattachai Wongtanawijit

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา
ปรัชญาดุษฎีบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์
มหาวิทยาลัยสงขลานครินทร์

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Computer Engineering
Prince of Songkla University

2563

ลิขสิทธิ์ของมหาวิทยาลัยสงขลานครินทร์

ชื่อวิทยานิพนธ์ การออกแบบและพัฒนาระบบมองเห็นสำหรับระบบทำสวนยางพาราอัตโนมัติ
 ผู้เขียน นายรัฐชัย วงศ์ธนวิจิต
 สาขาวิชา วิศวกรรมคอมพิวเตอร์

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

คณะกรรมการสอบ

.....
 (ผู้ช่วยศาสตราจารย์ ดร.ธเนศ เคารพพวงค์)

..... ประธานกรรมการ
 (รองศาสตราจารย์ ดร.วัฒน์พงศ์ เกิดทองมี)

..... กรรมการ
 (ผู้ช่วยศาสตราจารย์ ดร.ธเนศ เคารพพวงค์)

..... กรรมการ
 (ผู้ช่วยศาสตราจารย์ ดร.นิคม สุวรรณวร)

..... กรรมการ
 (รองศาสตราจารย์ ดร.มนตรี กาญจนเดชะ)

..... กรรมการ
 (ดร.อนันท์ ชกสุริวงศ์)

บัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์ อนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่ง
 ของการศึกษา ตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์

.....
 (ศาสตราจารย์ ดร.ดำรงศักดิ์ ฟ้ารุ่งแสง)
 คณบดีบัณฑิตวิทยาลัย

ขอรับรองว่า ผลงานวิจัยนี้มาจากการศึกษาวิจัยของนักศึกษาเอง และได้แสดงความขอบคุณบุคคลที่มีส่วนช่วยเหลือแล้ว

ลงชื่อ.....

(ผู้ช่วยศาสตราจารย์ ดร.ธเนศ เคารพพงศ์)

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ลงชื่อ.....

(นายรัฐชัย วงศ์ธนวิจิต)

นักศึกษา

ข้าพเจ้าขอรับรองว่า ผลงานวิจัยนี้ไม่เคยเป็นส่วนหนึ่งในการอนุมัติปริญญาในระดับใดมาก่อน และไม่ได้ถูกใช้
ในการยื่นขออนุมัติปริญญาในขณะนี้

ลงชื่อ.....

(นายรัฐชัย วงศ์ธนวิจิต)

นักศึกษา

| | |
|-----------------|---|
| ชื่อวิทยานิพนธ์ | การออกแบบและพัฒนาระบบมองเห็นสำหรับระบบทำสวนยางพาราอัตโนมัติ |
| ผู้เขียน | นายรัฐชัย วงศ์ธนวิจิตร |
| สาขาวิชา | วิศวกรรมคอมพิวเตอร์ |
| ปีการศึกษา | 2563 |

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้นำเสนอการพัฒนาระบบมองเห็นของระบบทำสวนยางพาราอัตโนมัติ ซึ่งทำหน้าที่ตรวจหาตำแหน่งของรอยกรีดบนหน้ายาง และตำแหน่งถ้วยรองน้ำยางที่แขวนติดกับแต่ละต้น สำหรับแขนหุ่นยนต์ในการกรีดและเก็บน้ำยาง โดยระบบมองเห็นใช้งานกล้องถ่ายภาพสีและภาพความลึก พร้อมทั้งไฟส่องสว่างในการถ่ายภาพในเวลาากลางคืน ระบบมองเห็นประกอบด้วยกล้องถ่ายภาพสองตำแหน่ง ได้แก่ กล้องระยะไกลที่ออกแบบให้สามารถติดตั้งบริเวณฐานหุ่นยนต์ เพื่อมุมรับภาพที่ครอบคลุมตำแหน่งรอยกรีดบนหน้ายาง และกล้องระยะใกล้เพื่อถ่ายภาพในรายละเอียดของแนวรอยกรีด

ผู้จัดทำได้รวบรวมชุดข้อมูลภาพถ่ายจากกล้องระยะไกล สำหรับใช้พัฒนาวิธีตรวจหารอยกรีดและถ้วยรองน้ำยาง โดยนำเสนอวิธีกำหนดตำแหน่งอ้างอิงของรอยกรีดและถ้วยรองน้ำยางในภาพด้วยการห่อหุ้มสี่เหลี่ยมล้อมรอบ (Bounding Box) พร้อมทั้งนำเสนอวิธีตรวจหารอยกรีดและถ้วยรองน้ำยางในภาพจากกล้องมุมกว้าง 2 วิธี ได้แก่ 1.วิธีตรวจหารอยกรีดโดยอาศัยสีโทนแดงบริเวณพื้นที่รอยกรีดเก่าและความเอียงของรอยกรีด ผลการตรวจหามีความแม่นยำเฉลี่ยสูงสุดที่ 35.4% และ 2.วิธีตรวจหารอยกรีดและถ้วยรองน้ำยางด้วยโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN ทำงานร่วมกับ MobileNetV2 มีความแม่นยำเฉลี่ยที่ 89.9% เมื่อวัดผลด้วยอัตราส่วนการซ้อนทับสำหรับวัดผล (Intersection-Over-Union) ที่ 0.5

การรวบรวมชุดภาพถ่ายระยะใกล้สำหรับใช้พัฒนาวิธีตรวจหาแนวรอยกรีด วิทยานิพนธ์ได้นำเสนอการออกแบบอุปกรณ์ถ่ายภาพสำหรับถ่ายภาพแนวรอยกรีดระยะใกล้ในที่มืด, วิธีกำหนดตำแหน่งแนวรอยกรีดอ้างอิงในภาพถ่ายระยะใกล้ด้วยเส้นโค้งหลายมุม (Polygonal Curve) พร้อมทั้งวิธีการปรับตำแหน่งแนวรอยกรีดอ้างอิงโดยอาศัยข้อมูลจากภาพถ่ายความลึก, และนำเสนอวิธีตรวจหาแนวรอยกรีดในภาพจากกล้องระยะใกล้ โดยใช้เทคนิคการหาความต่างระหว่างภาพ (Image Differencing) เพื่อค้นหาเงาตามแนวรอยกรีดและประมาณขอบของเงาเป็นแนวรอยกรีดเป้าหมาย ซึ่งจากการทดลองวิธีตรวจหาแนวรอยกรีดระยะใกล้สามารถตรวจหาแนวรอยกรีดได้ด้วยความแม่นยำเฉลี่ย 80.3% เมื่อวัดผลด้วยกรอบสี่เหลี่ยมที่อัตราส่วนการซ้อนทับที่ 0.5 โดยความคลาดเคลื่อนของแนวรอยกรีดที่ได้ เมื่อกำหนดด้วยระยะทางแบบเฮาส์ดอร์ฟมีค่าเฉลี่ยประมาณ 13 พิกเซลในภาพขนาด 1280x720 พิกเซล หรือเท่ากับ 9.0 มิลลิเมตร ในระนาบความลึกที่ขนานกับระนาบรับภาพของกล้อง

| | |
|----------------------|--|
| Thesis Title | Design and Development of Machine Vision for Automatic Rubber Latex Farming System |
| Author | Mr. Rattachai Wongtanawijit |
| Major Program | Computer Engineering |
| Academic Year | 2020 |

ABSTRACT

This thesis presents a machine vision for an automatic rubber latex-farming robot. The design adopts two RGB-Depth cameras (Intel RealSense D400 Series) with the additional lights that aim under the low-light rubber tree orchards. Two RGB-Depth cameras face a rubber tree at two different depth distances wherefore the tapping-panel coverages and the details of the tapping path.

A far-range image dataset is build up with bounding box annotations. The thesis includes two different detection algorithms that handle the RGB and the depth image of the far-range image dataset. A color-based with a sliding window algorithm can detect the tapping position bounding box up to 35.4% average precision at 0.5 IoU. A CNN-based detection algorithm, Faster-RCNN with pre-trained MobileNetV2, achieves 80.3% average precision of tapping position and cup detections at 0.5 IoU.

For a near-range image dataset, the thesis manifests the annotation of the tapping line using a bounding box and a polygonal curve with a refinement algorithm, and a tapping line detection algorithm that benefits the tapping path shadow for extracting the line. The evaluation of the near-range algorithm is in two steps, which are the bounding box to obtain the detection precision and then measures the distance error of the detected results. The algorithm produces 89.9% average precision at 0.5 IoU for the bounding box and the average distance error at 13 pixels using Hausdorff Distances within the high-resolution 1280 by 720-pixel images, which equals to 9.0 millimeters respected to camera geometry.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความช่วยเหลือจากคุณพ่อและคุณแม่ ที่ให้การสนับสนุนตลอดช่วงการดำเนินงานวิจัย โดยเฉพาะในการเตรียมสถานที่และความช่วยเหลือขณะถ่ายภาพเก็บข้อมูลในสวนยางพารา ขอขอบคุณเกษตรกรคุณจันและคุณรีสำหรับต้นยางพาราตัวอย่าง ขอขอบคุณอาจารย์ธเนศ อาจารย์ที่ปรึกษาที่ให้คำแนะนำสำหรับปัญหาหลายๆ อย่างที่เกิดขึ้น ขอขอบคุณรุ่นพี่และเจ้าหน้าที่ทุกๆ คนในภาควิชาวิศวกรรมคอมพิวเตอร์ สำหรับคำแนะนำและการช่วยเหลือปัญหาการดำเนินงานทางเอกสาร สุดท้ายนี้ขอขอบคุณเพื่อนๆ รวมถึงรุ่นน้องทุกๆ คนในห้องวิจัยที่มีส่วนเกี่ยวข้องกับงานวิจัยนี้ทั้งหมด

รัฐชัย วงศ์ธนวิจิต

สารบัญ

| เรื่อง | หน้า |
|---|-----------|
| บทคัดย่อ | (5) |
| ABSTRACT | (6) |
| กิตติกรรมประกาศ | (7) |
| สารบัญ | (8) |
| รายการตาราง..... | (13) |
| รายการภาพประกอบ..... | (14) |
| สัญลักษณ์คำย่อและตัวย่อ | (23) |
| บทที่ 1 บทนำ | 1 |
| 1.1 ความสำคัญและที่มา..... | 1 |
| 1.2 การตรวจเอกสาร | 3 |
| 1.2.1 ระบบหุ่นยนต์อัตโนมัติในงานเกษตรกับการปรับใช้สำหรับสวนยางพารา | 3 |
| 1.2.2 สวนยางพาราและต้นยางพารา..... | 4 |
| 1.2.3 หน้ายางและรอยกรีดในระบบกรีตลงแบบเวียนขวา | 4 |
| 1.2.4 ถ้วยรองน้ำยาง | 7 |
| 1.2.5 เทคโนโลยีการทำสวนยางพาราจนถึงปัจจุบัน | 7 |
| 1.2.6 ระบบมองเห็นของระบบหุ่นยนต์อัตโนมัติในงานเกษตร..... | 10 |
| 1.2.7 การประมวลผลภาพเพื่อการตรวจหาวัตถุ | 11 |
| 1.3 วัตถุประสงค์..... | 14 |
| 1.4 ขอบเขตของการวิจัย | 14 |
| 1.5 ประโยชน์ที่คาดว่าจะได้รับ..... | 14 |
| บทที่ 2 วิธีการวิจัย..... | 15 |
| 2.1 ระบบหุ่นยนต์ทำสวนยางพาราอัตโนมัติ..... | 15 |
| 2.1.1 ส่วนประกอบของหุ่นยนต์ | 15 |
| 2.1.2 ระบบมองเห็นของหุ่นยนต์ | 15 |
| 2.2 การออกแบบระบบมองเห็นสำหรับหุ่นยนต์ทำสวนยางพาราอัตโนมัติ | 17 |
| 2.2.1 ขั้นตอนการทำงาน..... | 17 |
| 2.2.2 กล้องถ่ายภาพ | 18 |
| 2.2.3 กล้องถ่ายภาพมุมกว้าง (กล้อง 1)..... | 20 |
| 2.2.4 กล้องระยะใกล้ (กล้อง 2) | 23 |
| 2.3 การพัฒนาวิธีตรวจหาตำแหน่งรอยกรีดและถ้วยรองน้ำยางของกล้องถ่ายภาพมุมกว้าง (กล้อง 1)..... | 25 |
| 2.3.1 การถ่ายภาพต้นยางพาราด้วยกล้องถ่ายภาพมุมกว้าง | 25 |
| 2.3.1.1 อุปกรณ์ถ่ายภาพ | 25 |
| 2.3.1.2 ขั้นตอนการถ่ายภาพ | 27 |

สารบัญ (ต่อ)

| เรื่อง | หน้า |
|---|------|
| 2.3.1.3 ชุดข้อมูลภาพถ่าย..... | 28 |
| 2.3.2 การกำหนดตำแหน่งอ้างอิงของรอยกรีดและถ้วยรองน้ำยางในภาพ (Ground Truth Annotation)..... | 31 |
| 2.3.3 วิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี..... | 33 |
| 2.3.3.1 การหาพื้นที่ลำต้นในภาพ (Trunk Segmentation) | 34 |
| 2.3.3.2 การสร้างกรอบสี่เหลี่ยมเพื่อค้นหาตำแหน่งรอยกรีด (Candidate Bounding Box Generation)..... | 36 |
| 2.3.3.3 การจัดกลุ่มสีภายในพื้นที่ลำต้นเพื่อสร้างพีเจอร์แมป (Color Feature Map)..... | 39 |
| 2.3.3.4 การเลือกกรอบสี่เหลี่ยมล้อมรอบรอยกรีด (Bounding Box Detection)..... | 43 |
| 2.3.4 วิธีตรวจหารอยกรีดและถ้วยรองน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยวิธีตรวจหาวัตถุชนิดโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN (Faster Region with Deep-Convolutional Neural Network)..... | 46 |
| 2.3.4.1 การพิจารณาเลือกใช้เน็ตเวิร์กที่ผ่านการสอนมาก่อน | 46 |
| 2.3.4.2 การพิจารณาโครงสร้างของเน็ตเวิร์กตรวจหาวัตถุแบบโครงข่ายประสาทเชิงลึกแบบคอนโวลูชัน สำหรับตรวจหารอยกรีดและถ้วยรองน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้าง..... | 47 |
| 2.3.4.3 การกำหนด Faster R-CNN ร่วมกับ MobileNetV2..... | 49 |
| 2.3.4.4 การสอนเน็ตเวิร์กแบบปรับละเอียด (Fine Tuning)..... | 55 |
| 2.3.4.5 การเตรียมข้อมูลภาพสำหรับสอนและทดสอบเน็ตเวิร์ก..... | 57 |
| 2.3.5 ตำแหน่งในสามมิติของรอยและกรีดถ้วยรองน้ำยางอ้างอิงกับกล้องถ่ายภาพมุมกว้าง..... | 62 |
| 2.3.5.1 การแปลงตำแหน่งในพิกัดภาพเป็นตำแหน่งในพิกัดฉากสามมิติของกล้องถ่ายภาพมุมกว้าง..... | 62 |
| 2.3.5.2 ตำแหน่งของรอยกรีดในสามมิติอ้างอิงกับตำแหน่งกล้องถ่ายภาพมุมกว้าง..... | 65 |
| 2.3.5.3 ตำแหน่งของถ้วยรองน้ำยางในสามมิติอ้างอิงกับตำแหน่งกล้องถ่ายภาพมุมกว้าง..... | 66 |
| 2.4 การตรวจหาแนวรอยกรีดด้วยกล้องระยะใกล้ (กล้อง 2)..... | 67 |
| 2.4.1 การถ่ายภาพแนวรอยกรีดด้วยกล้องระยะใกล้ | 67 |
| 2.4.1.1 การหาตำแหน่งและมุมของกล้องระยะใกล้สำหรับถ่ายภาพแนวรอยกรีด..... | 68 |

สารบัญ (ต่อ)

| เรื่อง | หน้า |
|--|------------|
| 2.4.1.2 ระบบไฟส่องสว่างของกล้องระยะใกล้สำหรับถ่ายภาพรอยกรีด | 73 |
| 2.4.1.3 ภาพถ่ายแนวรอยกรีดระยะใกล้..... | 79 |
| 2.4.2 การกำหนดแนวรอยกรีดอ้างอิงภายในภาพถ่ายระยะใกล้ (Reference Tapping-Line Annotation) | 81 |
| 2.4.2.1 การกำหนดแนวรอยกรีดอ้างอิงในภาพถ่ายด้วยมือ (Manual Reference Tapping-Line Annotation)..... | 81 |
| 2.4.2.2 วิธีปรับตำแหน่งแนว รอยกรีดอ้างอิงที่กำหนดด้วยมือ (Reference Tapping-Line Annotation Refinement Algorithm) | 84 |
| 2.4.3 วิธีตรวจหาแนวรอยกรีดในภาพถ่ายระยะใกล้..... | 95 |
| 2.4.4 ตำแหน่งแนวรอยกรีดในพิกัดสามมิติอ้างอิงกล้องระยะใกล้ (กล้อง 2)..... | 108 |
| 2.5 การวัดผลการตรวจหารอยกรีด, ถ้วยร่องน้ำยาง, และแนวรอยกรีดในภาพ | 111 |
| 2.5.1 การตรวจหารอยกรีดและถ้วยร่องน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้าง | 111 |
| 2.5.1.1 อัตราส่วนการซ้อนทับระหว่างกรอบสี่เหลี่ยม (Intersection-Over-Union, IoU) | 111 |
| 2.5.1.2 ความแม่นยำเฉลี่ย | 111 |
| 2.5.2 การวัดตำแหน่งของแนวรอยกรีดในภาพถ่ายจากกล้องระยะใกล้ | 112 |
| 2.5.2.1 การรู้จำแนวรอยกรีด | 112 |
| 2.5.2.2 ระยะคลาดเคลื่อนของแนวรอยกรีด..... | 112 |
| บทที่ 3 ผล | 114 |
| 3.1 การตรวจหารอยกรีดและถ้วยร่องน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้าง | 114 |
| 3.1.1 ผลการทดลอง: วิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี่ | 114 |
| 3.1.1.1 กรณีจัดกลุ่มสี่ด้วยภาพโมเดลสี HSV..... | 115 |
| 3.1.1.2 กรณีจัดกลุ่มสี่ด้วยภาพโมเดลสี $L^*a^*b^*$ | 119 |
| 3.1.1.3 สรุปผลการทดลอง: วิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี่ | 123 |
| 3.1.2 ผลการทดลอง: วิธีตรวจหารอยกรีดและถ้วยร่องน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้างด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN..... | 125 |
| 3.1.2.1 ตัวแปรต้นของการทดลอง..... | 125 |
| 3.1.2.2 ผลการสอนและค่าลอสมฟังก์ชันของการสอน (Training Loss)..... | 126 |
| 3.1.2.3 ความแม่นยำของการตรวจหารอยกรีด | 130 |
| 3.1.2.4 ความแม่นยำของการตรวจหาถ้วยร่องน้ำยาง | 134 |

สารบัญ (ต่อ)

| เรื่อง | หน้า |
|--|------------|
| 3.1.2.5 เวลาที่ใช้ประมวลผลต่อภาพ..... | 138 |
| 3.1.2.6 ผลของจำนวนพื้นที่ย่อยจาก RPN หากกับความแม่นยำในการตรวจหา รอยกรีดและถ้วยรองน้ำยาง..... | 139 |
| 3.1.2.7 สรุปผลการทดลอง: วิธีตรวจหารอยกรีดและถ้วยรองน้ำยางในภาพ จากกล้องถ่ายภาพมุมกว้างด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชัน ชนิด Faster-RCNN..... | 142 |
| 3.1.3 เปรียบเทียบผลการตรวจหารอยกรีดในภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง ของทั้งสองวิธี..... | 143 |
| 3.1.4 กรณีเปรียบเทียบ: ผลการตรวจหารอยกรีดและถ้วยรองน้ำยาง ระหว่างโครง ข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันชนิด YOLO กับชนิด Faster-RCNN เมื่อ ทำงานบน MobileNetV2 | 145 |
| 3.2 การทดลองภาพถ่ายจากกล้องระยะใกล้ (กล้อง 2) | 147 |
| 3.2.1 ผลการทดลอง: วิธีปรับตำแหน่งแนวรอยกรีดอ้างอิงจากแนวรอยกรีดอ้างอิงที่ กำหนดด้วยมือ | 147 |
| 3.2.1.1 ตัวแปรต้นของการทดลอง..... | 147 |
| 3.2.1.2 ผลการปรับตำแหน่งแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือ | 147 |
| 3.2.1.3 สรุปผลการทดลอง: วิธีปรับตำแหน่งแนวรอยกรีดอ้างอิงจากแนวรอย กรีดอ้างอิงที่กำหนดด้วยมือ | 151 |
| 3.2.2 ผลการทดลอง: วิธีตรวจหาแนวรอยกรีดในภาพถ่ายระยะใกล้ | 152 |
| 3.2.2.1 ตัวแปรต้นของการทดลอง..... | 152 |
| 3.2.2.2 ผลการตรวจหาแนวรอยกรีด..... | 152 |
| 3.2.2.3 สรุปผลการทดลอง..... | 152 |
| บทที่ 4 บทวิจารณ์..... | 155 |
| 4.1 การตรวจหาแนวรอยกรีดในภาพถ่ายมุมกว้าง | 155 |
| 4.1.1 การตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อน กรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี่..... | 155 |
| 4.1.2 การตรวจหารอยกรีดและถ้วยรองน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN | 157 |
| 4.1.2.1 จุดเด่นของสถาปัตยกรรมเน็ตเวิร์กคอนโวลูชันสำหรับตรวจหาวัตถุ แบบ Faster-RCNN | 157 |
| 4.1.2.2 ตำแหน่งของ RPN ภายใน MobileNetV2 | 158 |
| 4.1.2.3 การแบ่งกลุ่มข้อมูลทดลองตามช่องสัญญาณภาพอินพุต | 158 |
| 4.1.2.4 การแบ่งข้อมูลเพื่อสอนและทดสอบเน็ตเวิร์กของแต่ละกลุ่มทดลอง | 159 |
| 4.1.2.5 การเพิ่มข้อมูล (Data Augmentation)..... | 159 |

สารบัญ (ต่อ)

| เรื่อง | หน้า |
|--|------------|
| 4.1.2.6 ขนาดภาพอินพุต..... | 160 |
| 4.2 การตรวจหาแนวรอยกรีดในภาพถ่ายระยะใกล้..... | 161 |
| 4.2.1 การกำหนดแนวรอยกรีดอ้างอิงในภาพ..... | 161 |
| 4.2.1.1 การกำหนดแนวรอยกรีดอ้างอิงด้วยมือ..... | 161 |
| 4.2.1.2 การปรับตำแหน่งแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือ..... | 162 |
| 4.2.2 การตรวจหาตำแหน่งแนวรอยกรีดในภาพถ่ายระยะใกล้..... | 163 |
| 4.2.2.1 วิธีตรวจหาแนวรอยกรีด..... | 163 |
| 4.2.2.2 ผลการตรวจหาแนวรอยกรีดในภาพ..... | 164 |
| บทที่ 5 บทสรุปและข้อเสนอแนะ..... | 168 |
| บรรณานุกรม..... | 170 |
| ภาคผนวก ก ผลงานตีพิมพ์เผยแพร่จากวิทยานิพนธ์..... | 178 |
| ก.1 บทความวิจัยเสนอในที่ประชุมวิชาการและมีการพิมพ์รวมเล่ม..... | 178 |
| ก.2 บทความวิชาการ..... | 178 |
| ประวัติผู้เขียน..... | 179 |

รายการตาราง

| ตาราง | หน้า |
|--|------|
| ตารางที่ 2.1 ตารางแสดงคุณสมบัติของโมดูลกล้องถ่ายภาพสีและภาพความลึก ภายในกล้องอินเทลเรียลเซนส์โมเดล D415 และ D435 [10] | 18 |
| ตารางที่ 3.1 ค่าตัวแปรของแต่ละกลุ่มทดลอง วิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี | 114 |
| ตารางที่ 3.2 ความแม่นยำเฉลี่ย (Average Precision) ที่อัตราส่วนการซ้อนทับสำหรับวัดผลระดับต่างๆ ของวิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV | 115 |
| ตารางที่ 3.3 เวลาที่ใช้ประมวลผลด้วยโปรเซสเซอร์หลัก (CPU) เฉลี่ยต่อภาพในหน่วยวินาที ของวิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV | 116 |
| ตารางที่ 3.4 ความแม่นยำเฉลี่ย (Average Precision) ที่อัตราส่วนการซ้อนทับสำหรับวัดผลระดับต่างๆ ของวิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี ในกรณีเมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ | 119 |
| ตารางที่ 3.5 เวลาที่ใช้ประมวลผลด้วยโปรเซสเซอร์หลัก (CPU) เฉลี่ยต่อภาพในหน่วยวินาที ของวิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี ในกรณีเมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ | 120 |
| ตารางที่ 3.6 ความแม่นยำเฉลี่ย (Average Precision) ของการตรวจหารอยกรีดที่อัตราส่วนการซ้อนทับสำหรับวัดผลระดับต่างๆ ด้วย Faster-RCNN ที่มี RPN ในตำแหน่งต่างๆ ของแต่ละกลุ่มทดลอง | 130 |
| ตารางที่ 3.7 ความแม่นยำเฉลี่ย (Average Precision) ของการตรวจหาถ้วยรองน้ำยางที่อัตราส่วนการซ้อนทับสำหรับวัดผลระดับต่างๆ ด้วย Faster-RCNN ที่มี RPN ในตำแหน่งต่างๆ ของแต่ละกลุ่มทดลอง | 134 |
| ตารางที่ 3.8 เวลาเฉลี่ยที่ใช้ตรวจหาวัตถุต่อภาพของ Faster-RCNN ที่สร้างบน MobileNetV2 ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 3 และบล็อก 10 ในกรณีประมวลผลด้วย CPU และเมื่อประมวลผลด้วย GPU | 138 |
| ตารางที่ 3.9 ผลการปรับตำแหน่งแนวรอยกรีด ในกรณี sr แบบค่าเดียว | 148 |
| ตารางที่ 3.10 ผลการปรับตำแหน่งแนวรอยกรีด ในกรณี sr เป็นช่วง $[1, sr_{max}]$ | 149 |

รายการภาพประกอบ

| รูป | หน้า |
|---|------|
| รูปที่ 1.1 หน้ายาง (หมายเลข 1), ลื่นร่องน้ำยาง (หมายเลข 2), และถ้วยรองน้ำยาง (หมายเลข 3)..... | 6 |
| รูปที่ 1.2 หน้ายางเมื่อผ่านการกรีตเป็นระยะเวลาหนึ่ง (รูปซ้าย) และหน้ายางที่เป็ดกรีตใหม่ (รูปขวา) | 6 |
| รูปที่ 1.3 มีดกรีตยางแบบกำหนดความลึกได้ [74] | 8 |
| รูปที่ 1.4 มีดกรีตยางมอเตอร์ไฟฟ้า [72] | 8 |
| รูปที่ 1.5 เครื่องกรีตยางรางเลื่อนระบบไฟฟ้าแบบเคลื่อนที่ได้ตามแนวแถว [81] | 9 |
| รูปที่ 1.6 หุ่นยนต์สำรวจสวนยางพาราอัตโนมัติ (UGV) [82] | 9 |
| รูปที่ 2.1 ภาพวาดแนวคิดหุ่นยนต์ทำสวนยางพาราอัตโนมัติ..... | 16 |
| รูปที่ 2.2 ภาพวาดแสดงหุ่นยนต์เมื่อเข้าประจำด้านยางพารา..... | 16 |
| รูปที่ 2.3 แผนผังแสดงขั้นตอนการทำงานของระบบมองเห็นของหุ่นยนต์ทำสวนยางพาราอัตโนมัติ | 17 |
| รูปที่ 2.4 ภาพโมเดลกล้อง D415 และมุมรับภาพตามแนวแกนนอนตามปกติ (FOV-H) และมุมรับภาพตามแนวแกนตั้งตามปกติ (FOV-V) | 19 |
| รูปที่ 2.5 ภาพโมเดลกล้อง D435 และมุมรับภาพตามแนวแกนนอนตามปกติ (FOV-H) และมุมรับภาพตามแนวแกนตั้งตามปกติ (FOV-V) | 19 |
| รูปที่ 2.6 การวางกล้องถ่ายภาพมุมกว้าง (กล้อง 1) ในกรณีวางกล้องตามปกติ (ภาพซ้าย) ซึ่งมุมรับภาพตามแนวแกนตั้งปกติ (FOV-V) มีค่าน้อยกว่ามุมรับภาพตามแนวแกนนอน (FOV-H) ทำให้ต้องวางกล้องห่างจากหน้ายางมากกว่า ($a > b$) เพื่อให้ทำให้มุมรับภาพครอบคลุมพื้นที่หน้ายาง | 20 |
| รูปที่ 2.7 กรอบอ้างอิงกล้องถ่ายภาพมุมกว้าง (กล้อง 1) | 21 |
| รูปที่ 2.8 แผนภาพแสดงมุมมองด้านบน (ระนาบ YZ) และมุมมองด้านข้าง (ระนาบ XZ) ซึ่งกล้องถ่ายภาพมุมกว้าง (กล้อง 1) อยู่ห่างจากหน้ายางเป็นระยะตั้งฉากเท่ากับ z_0 โดยหน้ายางของต้นยางแต่ละต้นไม่ได้เรียง กันเป็นแนวเส้นตรงโดยสมบูรณ์ (มีระยะเหลื่อมดังแสดงในมุมมองด้านบน) | 22 |
| รูปที่ 2.9 กรอบอ้างอิงกล้องระยะใกล้ (กล้อง 2)..... | 23 |
| รูปที่ 2.10 มุมมองด้านข้าง (ระนาบ YZ ตามกรอบอ้างอิงกล้องระยะใกล้) แสดงตำแหน่งมุมรับภาพของกล้องระยะใกล้ (กล้อง 2) กับพื้นที่รอยกรีต | 24 |
| รูปที่ 2.11 อุปกรณ์ถ่ายภาพกล้องถ่ายภาพมุมกว้าง..... | 25 |
| รูปที่ 2.12 (มุมมองจากด้านบน) แผนภาพแสดงตำแหน่งของอุปกรณ์ถ่ายภาพเมื่อวางหน้าต้นยางพารา แสดงตำแหน่งไฟส่องสว่างทั้งสอง (ไฟซ้าย, ไฟขวา) กับตำแหน่งกล้องถ่ายภาพมุมกว้าง (กล้อง 1)..... | 27 |
| รูปที่ 2.13 (มุมมองจากด้านบน) แผนภาพแสดงการวางอุปกรณ์ถ่ายภาพเมื่อย้ายตำแหน่งไปจากเดิมประมาณ 50-100 เซนติเมตร ซึ่งยังทำให้การถ่ายภาพรอยกรีตเป็นไปตามเงื่อนไข โดยหันกล้องเข้าหาลำต้นและยังคงระยะห่างระหว่างกล้องกับหน้ายางในช่วง 90-120 เซนติเมตร..... | 28 |
| รูปที่ 2.14 ตัวอย่างถ้วยรองน้ำยางชนิดพลาสติกสีดำขนาด 1000 และ 700 มิลลิลิตร | 29 |
| รูปที่ 2.15 ตัวอย่างภาพถ่ายจากกล้องถ่ายภาพมุมกว้างจำนวน 1 ชุดภาพ..... | 29 |
| รูปที่ 2.16 ภาพถ่ายตัวอย่างสวนยางพารา..... | 30 |

รายการภาพประกอบ (ต่อ)

| รูป | หน้า |
|---|------|
| รูปที่ 2.17 ตัวอย่างการวาดกรอบสี่เหลี่ยมล้อมรอบรอยกริดและถ้วยรองน้ำยาง..... | 32 |
| รูปที่ 2.18 แผนผังวิธีการตรวจหารอยกริดบนภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคเลื่อนกรอบสี่เหลี่ยมภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี..... | 33 |
| รูปที่ 2.19 ภาพถ่ายความลึกโทนาเทา (I_{D1})..... | 35 |
| รูปที่ 2.20 ภาพขาวดำของพื้นที่ลำต้น (I_M)..... | 35 |
| รูปที่ 2.21 ภาพถ่ายสีจากกล้องถ่ายภาพมุมกว้าง..... | 35 |
| รูปที่ 2.22 ภาพถ่ายสีเมื่อผ่านการหาพื้นที่ลำต้น..... | 35 |
| รูปที่ 2.23 แสดงกรอบสี่เหลี่ยมล้อมรอบรอยกริดภายในพื้นที่ลำต้น (A_i) ที่ขนาดความกว้างมากที่สุด ($w_b = w_i$) และที่ขนาดความกว้างน้อยที่สุด ($w_b = \frac{1}{2}w_i$) โดยแนวรอยกริดจริง (เส้นโค้งประทแยงมุม) ถูกประมาณด้วยเส้นทแยงมุมของกรอบสี่เหลี่ยม | 36 |
| รูปที่ 2.24 ภาพแสดงกรอบสี่เหลี่ยมที่เป็นไปได้ภายในพื้นที่ลำต้นที่มีอัตราส่วนความกว้างต่อความสูง (Aspect Ratio) และขนาดต่างๆ..... | 38 |
| รูปที่ 2.25 ภาพวาดหน้ายางและลำต้นยางพาราแสดงพื้นที่รอยกริดเก่าและตำแหน่งของแนวรอยกริดโดยแนวรอยกริดคือแนวแบ่งระหว่าง เปลือกยางที่ยังไม่ผ่านการกริดกับพื้นที่รอยกริดเก่า..... | 39 |
| รูปที่ 2.26 ทรงกรวยอธิบายโมเดลสี HSV [35] และการจัดกลุ่มค่าสีออกเป็น 4 กลุ่ม | 40 |
| รูปที่ 2.27 ตัวอย่างการจัดกลุ่มค่าสีออกเป็น 4 กลุ่ม ของภาพในโมเดลสี HSV เมื่อถูกดัดแปลงค่าด้วยสมการที่ 2.8 ตามรูปที่ 2.26, โดยรูปทางซ้ายแสดงภาพสี RGB ต้นฉบับที่ถูกจำกัดพื้นที่ลำต้นจากรูปที่ 2.22 และกลุ่มภาพทางขวาแสดงภาพขาวดำของกลุ่มสีที่มีค่ากลางใกล้เคียงกับค่ากลางเริ่มต้น ซึ่งได้แก่ 0.1, 0.2, 0.3 และ 0.4 เรียงตามลำดับ | 41 |
| รูปที่ 2.28 ตัวอย่างการจัดกลุ่มสีออกเป็น 3 กลุ่มในช่องสัญญาณ a^* ของโมเดลสี CIELAB โดยรูปทางซ้ายมือแสดงภาพสีตัวอย่างต้นฉบับ RGB (รูปที่ 2.22) และภาพทางขวามือแสดงกลุ่มสีที่มีโทนสีแดงมาก, ปานกลาง และ โทนสีเขียวมาก ตามลำดับ ซึ่งใช้วิธีจัดกลุ่มแบบค่ากลาง K กลุ่ม ในการจัดกลุ่มเช่นเดียวกันกับการจัดกลุ่มสีในช่องสัญญาณ Hue ของโมเดลสี HSV..... | 42 |
| รูปที่ 2.29 ตัวอย่างกรอบสี่เหลี่ยมที่ตรงตำแหน่งรอยกริด | 43 |
| รูปที่ 2.30 พื้นที่สามเหลี่ยมบนและล่างจากการแบ่งพื้นที่ภายในกรอบสี่เหลี่ยมด้วยเส้นทแยงมุม | 43 |
| รูปที่ 2.31 ภาพตัวอย่างแสดงกรอบสี่เหลี่ยมผลลัพธ์ (Output Box) ที่เกิดจากการรวมกรอบสี่เหลี่ยมที่มีค่าคะแนนสูงสุดในแต่ละขนาดความกว้าง..... | 45 |
| รูปที่ 2.32 ข้อมูลจาก [6] แผนภาพโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันที่ผ่านการสอนมาก่อน โดยในกรอบสี่เหลี่ยมเส้นประเป็นกลุ่มที่ของเน็ตเวิร์กที่เหมาะสมสำหรับระบบหุ่นยนต์ทำสวนยางพารา | 48 |
| รูปที่ 2.33 แผนภาพแสดงขนาดของข้อมูล (รูปภาพที่อยู่ในรูปแบบเมทริกซ์สามมิติ) หรือเรียกว่าเทนเซอร์ (Tensors) ภายในเน็ตเวิร์ก MobileNetV2 [69] โดยขนาดของกล่องสี่เหลี่ยมแสดงถึงขนาดเทนเซอร์ (กว้าง × ยาว × ลึก) ซึ่งภายในและระหว่างกล่องมีชั้นประมวลผลคอนโวลูชันแบบ Inverted Residuals และ Linear Bottlenecks หรือคอนโวลูชันเลเยอร์แบบปกติ (Conv2d) | 49 |

รายการภาพประกอบ (ต่อ)

| รูป | หน้า |
|--|------|
| รูปที่ 2.34 โครงสร้างเลเยอร์แบบ Inverted Residuals และ Linear Bottlenecks ของ MobileNetV2 จำนวน 1 บล็อก แสดง ณ ตำแหน่ง “block_10” ซึ่งรับอินพุตจาก “block_9” และแสดงเลเยอร์ส่วนปลายของ RPN ได้แก่ เลเยอร์คอนโวลูชัน (Conv2d), ReLU, เลเยอร์ทำนายประเภทวัตถุ (Softmax), เลเยอร์ทำนายตำแหน่งกรอบสี่เหลี่ยม (Regressor) และเลเยอร์พูลิงพื้นที่ที่สนใจ (ROI Pooling) สำหรับคำนวณหาเทนเซอร์เฉพาะในพื้นที่กรอบสี่เหลี่ยมที่สนใจ เพื่อป้อนให้ชั้นของเน็ตเวิร์กในลำดับสูงกว่าใน “block_11” | 52 |
| รูปที่ 2.35 ภาพรวมเน็ตเวิร์กตรวจหาตำแหน่งวัตถุแบบ Faster-RCNN ที่สร้างบน MobileNetV2 สำหรับตรวจหารอยกริดและถ้วยร่อนน้ำยาง โดย RPN ทำงานบนพีเจอร์แมปขนาดกว้าง 28x28..... | 53 |
| รูปที่ 2.36 ภาพรวมเน็ตเวิร์กตรวจหาตำแหน่งวัตถุแบบ Faster-RCNN ที่สร้างบน MobileNetV2 สำหรับตรวจหารอยกริดและถ้วยร่อนน้ำยาง โดย RPN ทำงานบนพีเจอร์แมปขนาดกว้าง 14x14..... | 53 |
| รูปที่ 2.37 ภาพรวมเน็ตเวิร์กตรวจหาตำแหน่งวัตถุแบบ Faster-RCNN ที่สร้างบน MobileNetV2 สำหรับตรวจหารอยกริดและถ้วยร่อนน้ำยาง โดย RPN ทำงานบนพีเจอร์แมปขนาดกว้าง 7x7 | 54 |
| รูปที่ 2.38 รูปวาดเปรียบเทียบขนาดของภาพถ่ายสีและภาพความลึก (RGB-Depth) และขนาดของอินพุตเลเยอร์ของ MobileNetV2 | 57 |
| รูปที่ 2.39 กลุ่มทดลองจัดสรรภาพ RGB-Depth ให้ตรงกับขนาดอินพุตเลเยอร์ของ MobileNetV2..... | 58 |
| รูปที่ 2.40 การเตรียมข้อมูลภาพสำหรับสอนและทดสอบเน็ตเวิร์ก | 59 |
| รูปที่ 2.41 ตัวอย่างการแปลงข้อมูลภาพในกลุ่มทดลองที่ 1 จากภาพตัวอย่างภาพถ่ายสี RGB 1 ภาพ เมื่อถูกแปลงขนาดให้เท่ากับ 224x224 พิกเซล และเมื่อซูมแปลงภาพด้วยการหมุนและขดเคยความสว่างใน 6 รูปแบบที่ไม่ซ้ำกัน พร้อมทั้งแสดงกรอบสี่เหลี่ยมล้อมรอบอ้างอิงของรอยกริดและถ้วยร่อนน้ำยางในภาพ | 60 |
| รูปที่ 2.42 ตัวอย่างการแปลงข้อมูลภาพในกลุ่มทดลองที่ 2 จากภาพตัวอย่างภาพถ่ายความลึก 1 ภาพ เมื่อถูกแปลงขนาดให้เท่ากับ 224x224 พิกเซล และเมื่อซูมแปลงภาพด้วยการหมุนและขดเคยความสว่างใน 6 รูปแบบที่ไม่ซ้ำกัน พร้อมทั้งแสดงกรอบสี่เหลี่ยมล้อมรอบอ้างอิงของรอยกริดและถ้วยร่อนน้ำยางในภาพ | 61 |
| รูปที่ 2.43 ตัวอย่างการแปลงข้อมูลภาพในกลุ่มทดลองที่ 3 จากภาพตัวอย่างภาพ 1 ชุดภาพ เมื่อถูกแปลงขนาดให้เท่ากับ 224x224 พิกเซล และเมื่อซูมแปลงภาพด้วยการหมุนและขดเคยความสว่างใน 6 รูปแบบที่ไม่ซ้ำกัน พร้อมทั้งแสดงกรอบสี่เหลี่ยมล้อมรอบอ้างอิงของรอยกริดและถ้วยร่อนน้ำยางในภาพ..... | 61 |
| รูปที่ 2.44 ภาพสามมิติโมเดลกล้อง D415 [10, 11] แสดงตำแหน่งกล้องถ่ายภาพสี (Color Camera) และกล้องถ่ายภาพความลึก (Stereo Depth Camera) ที่อยู่ภายใน และกรอบอ้างอิงพิกัดฉากที่มีจุดกำเนิด ณ ตำแหน่งกลางกล้องถ่ายภาพสี โดยกรอบอ้างอิงมีทิศทางของแกนตามเดิมจากผู้ผลิตกล้อง และระนาบความลึกศูนย์ (Z=0) อยู่ถัดจากกระจกหน้าเลนส์ (Cover Glass) เข้าไปในตัวกล้องเป็นระยะ 1.1 มิลลิเมตร | 63 |

รายการภาพประกอบ (ต่อ)

| รูป | หน้า |
|--|------|
| รูปที่ 2.45 ภาพโมเดลกล้อง D415 [11] เมื่อกล้องวางตัวในแนวตั้ง (Portrait Orientation) โดยแสดงกรอบอ้างอิง มีจุดกำเนิด ณ ตำแหน่งกลางกล้องถ่ายภาพสี (Color Camera) โดยแกน Y มีทิศทางพุ่งขึ้นจากระนาบพื้น, แกน X ขนานกับระดับพื้น และแกนความลึก Z ชี้เข้าหาต้นยางพารา..... | 64 |
| รูปที่ 2.46 แสดงตำแหน่งภายในกรอบสี่เหลี่ยมล้อมรอบรอยกริด และ (u_{b0}, v_{b0}) คือจุดกึ่งกลางของกรอบสี่เหลี่ยมในพิกัดภาพ | 65 |
| รูปที่ 2.47 แสดงตำแหน่งกรอบล้อมรอบถ้วยรองน้ำยางในภาพ และ (u_{b0}, v_{b0}) คือจุดกึ่งกลางของกรอบสี่เหลี่ยมล้อมรอบในพิกัดภาพ..... | 66 |
| รูปที่ 2.48 แสดงกรอบอ้างอิงของกล้องระยะใกล้ในตำแหน่งปลายแขนหุ่นยนต์ | 67 |
| รูปที่ 2.49 ตำแหน่งและการหันกล้องระยะใกล้เพื่อถ่ายภาพรอยกริด, (x_c, y_c, z_c) , โดยแสดงความสัมพันธ์ระหว่างตำแหน่งในระนาบ XZ ของกรอบอ้างอิงของกล้องถ่ายภาพมุมกว้าง กับตำแหน่งในกรอบล้อมรอบรอยกริดตามแกนนอนของภาพ (U) ภายในพิกัดภาพจากกล้องถ่ายภาพมุมกว้าง..... | 69 |
| รูปที่ 2.50 แสดงมุมรับภาพของกล้องระยะใกล้ในระนาบ YZ เมื่อกำหนดให้แกนความลึกตั้งฉากกับผิวหน้ายาง ซึ่งทำให้ความลึกบริเวณลำต้นในภาพส่วนบน, ส่วนล่าง, และบริเวณกลางภาพ มีค่าประมาณเท่ากัน $z_{2top} \approx z_{2bot} \approx z_2$ | 71 |
| รูปที่ 2.51 ตำแหน่งกล้องระยะใกล้ในระนาบ XZ ในกรณีที่มีมุมรับภาพครอบคลุมแนวรอยกริดได้ทั้งหมด ซึ่งทำให้กล้องสามารถถ่ายภาพแนวรอยกริดได้สมบูรณ์ โดยแสดงความลึกที่จุดเริ่มต้นแนวรอยกริด (z_{start}) และความลึกจุดสิ้นสุดแนวรอยกริด (z_{end}), และความลึกตำแหน่งกลางภาพ (z_2)..... | 72 |
| รูปที่ 2.52 ตำแหน่งกล้องระยะใกล้ในระนาบ XZ เมื่อเคลื่อนจากจุด O (ตามรูปที่ 2.51) ไปยังตำแหน่ง O' ซึ่งเป็นตำแหน่งที่มีมุมรับภาพไม่ครอบคลุมแนวรอยกริดทั้งหมด ทำให้ไม่สามารถอ่านค่าความลึกของจุดสิ้นสุดแนวรอยกริดได้..... | 73 |
| รูปที่ 2.53 การให้แสงในระนาบ YZ ของไฟส่องสว่างที่ 1, ไฟส่องสว่างที่ 2, และมุมรับภาพของกล้องระยะใกล้..... | 75 |
| รูปที่ 2.54 การให้แสงของไฟส่องสว่างที่ 2 กับหน้ายางในระนาบ YZ เพื่อสร้างเงาตามแนวรอยกริด..... | 75 |
| รูปที่ 2.55 ภาพวาดอุปกรณ์ระบบไฟส่องสว่างสำหรับถ่ายภาพแนวรอยกริดระยะใกล้ | 76 |
| รูปที่ 2.56 ภาพวาดในระนาบ YZ ของอุปกรณ์ให้แสงเพื่อการถ่ายภาพแนวรอยกริดระยะใกล้ โดยแสดงระยะ d_1, d_2 , และ d_3 | 77 |
| รูปที่ 2.57 อุปกรณ์ถ่ายภาพแนวรอยกริดระยะใกล้ | 78 |
| รูปที่ 2.58 อุปกรณ์ถ่ายภาพแนวรอยกริดระยะใกล้ขณะถ่ายภาพต้นยางพาราในสภาพแสงน้อย | 78 |
| รูปที่ 2.59 ตัวอย่างภาพถ่ายแนวรอยกริดระยะใกล้, ภาพถ่ายสีที่ 1 | 80 |
| รูปที่ 2.60 ตัวอย่างภาพถ่ายแนวรอยกริดระยะใกล้, ภาพถ่ายสีที่ 2 | 80 |
| รูปที่ 2.61 ตัวอย่างภาพถ่ายความลึกแนวรอยกริดระยะใกล้, (แสดงในรูปแบบ 8 บิต, I_{D2_8})..... | 80 |

รายการภาพประกอบ (ต่อ)

| รูป | หน้า |
|---|------|
| รูปที่ 2.62 ภาพวาดแสดงกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริดจริง (แสดงด้วยเส้นประ) ในภาพถ่ายสีที่ 1 ซึ่งกำหนดให้มุมซ้ายบนของกรอบ (u_b, v_b) อยู่ที่ตำแหน่งแนวรอยกริดซ้ายสุดที่ปรากฏในภาพ และมุมขวาล่างของกรอบ ($u_b + w_b, v_b + h_b$) อยู่ที่ตำแหน่งขวาสุดของแนวรอยกริดที่ปรากฏในภาพ..... | 82 |
| รูปที่ 2.63 รูปหลายเหลี่ยมภายในกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริด ซึ่งผู้กำหนดได้กำหนดจุดมุมเพิ่มเติม (Control Points) ให้กับรูปหลายเหลี่ยมตามแนวรอยกริดในระหว่างพิกเซลมุมซ้ายบนและมุมขวาล่างของกรอบสี่เหลี่ยม | 83 |
| รูปที่ 2.64 “เส้นขอบล่าง” (“Lower boundary”) ของรูปหลายเหลี่ยมในรูปที่ 2.63..... | 83 |
| รูปที่ 2.65 แผนผังขั้นตอนวิธีการปรับตำแหน่งรอยกริดที่กำหนดด้วยมือและการเลือกแนวรอยกริดอ้างอิงสำหรับชุดข้อมูล..... | 85 |
| รูปที่ 2.66 ภาพแสดงพื้นที่กรอบค้นหา R_{sr} ในระนาบพิกัด C_2 ซึ่งเกิดจากการขยายแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ (T_{manual}) ไปตามแกนตั้งเป็นระยะ sr ในทั้งสองทิศทาง..... | 88 |
| รูปที่ 2.67 ตัวอย่างภาพโทเนทาแสดงค่าเกรเดียนต์ที่เป็นบวกของภาพถ่ายความลึก ($I_{D2_8_hp}$) เมื่อผ่านการทำคอนโวลูชันกับฟิลเตอร์เกาส์ที่มีค่าการกระจายของน้ำหนักต่างกัน โดยแสดงขนาดขยาย 100x100 พิกเซล (Cropped) และตำแหน่งที่เป็นจุดยอดตามแนวตั้ง (Vertical Local Maxima) | 89 |
| รูปที่ 2.68 ภาพตัวอย่างเปรียบเทียบระหว่างแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ (T_{manual}) กับแนวรอยกริดอ้างอิงใหม่จากวิธีปรับตำแหน่ง (T_{refined})..... | 94 |
| รูปที่ 2.69 แผนผังขั้นตอนวิธีตรวจหารอยกริดในภาพถ่ายจากกล้องระยะใกล้ | 95 |
| รูปที่ 2.70 ตัวอย่างภาพโทเนทาที่ผ่านการปรับค่าความสว่างให้มีค่าความสว่างเฉลี่ยภายในพื้นที่ลำต้นเท่ากัน ระหว่างภาพโทเนทาของภาพถ่ายสีที่ 1 (I_{21_n}) ในรูป (ก) และภาพโทเนทาของภาพถ่ายสีที่ 2 (I_{22_n}) ในรูป (ข) | 98 |
| รูปที่ 2.71 ตัวอย่างภาพผลลัพธ์การลบระหว่างภาพโทเนทาที่ผ่านการปรับความสว่างของภาพโทเนทาของภาพถ่ายสีที่ 1 และภาพโทเนทาของภาพถ่ายสีที่ 2 แสดงในรูป (ก) และภาพผลลัพธ์ของการเลือกค่าตามข้อ 3 ในอัลกอริทึม 2.4.5 (I_{2_diff}) ในรูป (ข)..... | 99 |
| รูปที่ 2.72 ตัวอย่างการประเมินพื้นที่ย่อย (Regions), จากภาพ เส้นขอบล่าง 3 ของพื้นที่ย่อย R_3 ถูกกีดขวางจากขอบล่างของภาพอย่างสมบูรณ์โดยพื้นที่ย่อยอื่นในภาพ ในขณะที่เส้นขอบล่าง 2 ของพื้นที่ย่อย R_2 ถูกกีดขวางจากขอบล่างของภาพเพียงบางส่วนด้วยพื้นที่ย่อย R_1 | 100 |
| รูปที่ 2.73 ตัวอย่างภาพวาดแสดงลักษณะพื้นที่ย่อยที่เป็นเงาแนวรอยกริดและพื้นที่บริเวณข้างเคียง | 102 |
| รูปที่ 2.74 ตัวอย่างภาพวาดผลลัพธ์จากอัลกอริทึม 2.4.5 ในรูป (ก) และผลจากอัลกอริทึม 2.4.6 ในรูป (ข)..... | 103 |
| รูปที่ 2.75 ตัวอย่างการประมาณเส้นตรงในช่วงว่างระหว่างพื้นที่ย่อย..... | 105 |

รายการภาพประกอบ (ต่อ)

| รูป | หน้า |
|---|------|
| รูปที่ 2.76 กราฟแสดงเส้นขอบล่างของรูปที่ 2.75 และอนุพันธ์อันดับหนึ่งของ $g(u)$ เทียบกับค่า u แทนด้วย $g'(u)$ และแสดงการเลือกค่า T_{up}, T_{down} ที่สามารถแยกเส้นขอบล่างที่เป็นแนวรอยกรีดออกจากเส้นขอบล่างทั้งหมดได้ถูกต้อง | 107 |
| รูปที่ 2.77 ตัวอย่างผลการตรวจหาแนวรอยกรีดในภาพถ่ายระยะใกล้ เมื่อแสดงแนวรอยกรีด (แนวจุดพิกเซลสีขาว) แบบซ้อนทับกับภาพขยายของภาพถ่ายสีที่ 1 | 108 |
| รูปที่ 2.78 กรอบอ้างอิงของกล้องโมเดล D435 (กล้องระยะใกล้) มีจุดกำเนิดที่ตำแหน่งกึ่งกลางของกล้องถ่ายภาพสี (Color Camera), โมเดลสามมิติของกล้องจาก [11]..... | 109 |
| รูปที่ 2.79 แผนภาพแสดงตำแหน่งแนวรอยกรีดในพิกัดสามมิติของกล้องระยะใกล้ โดยค่า XYZ ในหน่วยเมตร..... | 110 |
| รูปที่ 2.80 การซ้อนทับของกรอบสี่เหลี่ยมอ้างอิงกับกรอบสี่เหลี่ยมจากการตรวจหาวัตถุ | 111 |
| รูปที่ 2.81 ตัวอย่างผลการตรวจหากรอบสี่เหลี่ยมแนวรอยกรีดในภาพถ่ายจากกล้องระยะใกล้ (กล้อง 2) ซึ่งทั้งสองกรณีมีค่าอัตราส่วนการซ้อนทับเท่ากัน..... | 112 |
| รูปที่ 2.82 แสดงตัวอย่างเส้นโค้งหลายมุม (Polygonal Curve) ของแนวรอยกรีดอ้างอิง (T_{GT}) และเส้นโค้งหลายมุมของแนวรอยกรีดที่ตรวจพบ ($T_{Detected}$) ซึ่งมีจำนวนจุด (พิกเซล) ไม่เท่ากัน และแสดงระยะทางแบบแฮสค็อพฟ์ จาก $T_{Detected}$ ไปยัง T_{GT} , $h(T_{detected}, T_{GT})$ | 113 |
| รูปที่ 3.1 ความแม่นยำเฉลี่ยในการตรวจหารอยกรีด ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV ของกลุ่มทดลองที่ 1, 2 และ 3 ซึ่งใช้ค่าระยะเลื่อนเท่ากับ 5 พิกเซล ($n_s = 5$)..... | 116 |
| รูปที่ 3.2 ความแม่นยำเฉลี่ยในการตรวจหารอยกรีด ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV ของกลุ่มทดลองที่ 4, 5 และ 6 ซึ่งใช้ค่าระยะเลื่อนเท่ากับ 15 พิกเซล ($n_s = 15$)..... | 117 |
| รูปที่ 3.3 ความแม่นยำเฉลี่ยในการตรวจหารอยกรีด ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV ของกลุ่มทดลอง 1 กับกลุ่มทดลอง 4 ซึ่งใช้ค่าระยะเลื่อน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน | 117 |
| รูปที่ 3.4 ความแม่นยำเฉลี่ยในการตรวจหารอยกรีด ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV ของกลุ่มทดลอง 2 กับกลุ่มทดลอง 5 ซึ่งใช้ค่าระยะเลื่อน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน | 118 |
| รูปที่ 3.5 ความแม่นยำเฉลี่ยในการตรวจหารอยกรีด ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV เปรียบเทียบกลุ่มทดลอง 3 กับกลุ่ม 6 ซึ่งใช้ค่าระยะเลื่อน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน | 118 |
| รูปที่ 3.6 ความแม่นยำเฉลี่ยในการตรวจหารอยกรีด ในกรณีเมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ ของกลุ่มทดลอง 1, 2 และ 3 ซึ่งใช้ค่าระยะเลื่อนเท่ากับ 5 พิกเซล ($n_s = 5$)..... | 120 |
| รูปที่ 3.7 ความแม่นยำเฉลี่ยในการตรวจหารอยกรีด ในกรณีเมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ ของกลุ่มทดลอง 4, 5 และ 6 ซึ่งใช้ค่าระยะเลื่อนเท่ากับ 15 พิกเซล ($n_s = 15$)..... | 121 |
| รูปที่ 3.8 ความแม่นยำเฉลี่ยในการตรวจหารอยกรีด ในกรณีเมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ ของกลุ่มทดลอง 1 กับ 4 ซึ่งใช้ค่าระยะเลื่อน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน | 121 |
| รูปที่ 3.9 ความแม่นยำเฉลี่ยในการตรวจหารอยกรีด ในกรณีเมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ ของกลุ่มทดลอง 2 กับ 5 ซึ่งใช้ค่าระยะเลื่อน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน | 122 |

รายการภาพประกอบ (ต่อ)

| รูป | หน้า |
|---|------|
| รูปที่ 3.10 ความแม่นยำเฉลี่ยในการตรวจหารอยกริด ในกรณีเมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ ของกลุ่มทดลองที่ 3 กับ 6 ซึ่งใช้ค่าระยะเลือน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน..... | 122 |
| รูปที่ 3.11 ความแม่นยำเฉลี่ยในการตรวจหารอยกริดด้วยวิธีตรวจหารอยกริดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี เปรียบเทียบระหว่างในกรณีเมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ และโมเดลสี HSV..... | 123 |
| รูปที่ 3.12 ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 3..... | 127 |
| รูปที่ 3.13 ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 10..... | 127 |
| รูปที่ 3.14 ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 3..... | 128 |
| รูปที่ 3.15 ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 10..... | 128 |
| รูปที่ 3.16 ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 3..... | 129 |
| รูปที่ 3.17 ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 10..... | 129 |
| รูปที่ 3.18 ความแม่นยำเฉลี่ยในการตรวจหารอยกริดด้วย Faster-RCNN กลุ่มทดลองที่ 1 ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2..... | 131 |
| รูปที่ 3.19 ความแม่นยำเฉลี่ยในการตรวจหารอยกริดด้วย Faster-RCNN กลุ่มทดลองที่ 2 ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2..... | 131 |
| รูปที่ 3.20 ความแม่นยำเฉลี่ยในการตรวจหารอยกริดด้วย Faster-RCNN กลุ่มทดลองที่ 3 ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2..... | 132 |
| รูปที่ 3.21 ความแม่นยำเฉลี่ยในการตรวจหารอยกริดด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 ของ MobileNetV2 แสดงกลุ่มทดลองที่ 1, 2, และ 3..... | 133 |
| รูปที่ 3.22 ความแม่นยำเฉลี่ยในการตรวจหารอยกริดด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 10 ของ MobileNetV2 แสดงกลุ่มทดลองที่ 1, 2, และ 3..... | 133 |
| รูปที่ 3.23 ความแม่นยำเฉลี่ยในการตรวจหาถ้วยรองน้ำยางด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2 ในกลุ่มทดลองที่ 1..... | 135 |
| รูปที่ 3.24 ความแม่นยำเฉลี่ยในการตรวจหาถ้วยรองน้ำยางด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2 ในกลุ่มทดลองที่ 2..... | 135 |
| รูปที่ 3.25 ความแม่นยำเฉลี่ยในการตรวจหาถ้วยรองน้ำยางด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2 ในกลุ่มทดลองที่ 3..... | 136 |
| รูปที่ 3.26 ความแม่นยำเฉลี่ยในการตรวจหาถ้วยรองน้ำยางด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 ของ MobileNetV2 แสดงกลุ่มทดลองที่ 1, 2, และ 3..... | 137 |
| รูปที่ 3.27 ความแม่นยำเฉลี่ยในการตรวจหาถ้วยรองน้ำยางด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 10 ของ MobileNetV2 แสดงกลุ่มทดลองที่ 1, 2, และ 3..... | 137 |

รายการภาพประกอบ (ต่อ)

| รูป | หน้า |
|--|------|
| รูปที่ 3.28 ระยะเวลาประมวลผลเฉลี่ยต่อภาพของการตรวจหารอยกริดและถ้วยรองน้ำยางด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2 เมื่อประมวลผลด้วยหน่วยประมวลผลหลัก (CPU) หรือกราฟิกการ์ด (GPU) | 138 |
| รูปที่ 3.29 ความแม่นยำเฉลี่ยของการตรวจหารอยกริดเมื่อเปลี่ยนจำนวนพื้นที่ย่อยจาก RPN ในขณะตรวจหาวัตถุด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 3 ของ MobileNetV2 ... | 139 |
| รูปที่ 3.30 ความแม่นยำเฉลี่ยของการตรวจหารอยกริดเมื่อเปลี่ยนจำนวนพื้นที่ย่อยจาก RPN ในขณะตรวจหาวัตถุด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 10 ของ MobileNetV2 | 140 |
| รูปที่ 3.31 ความแม่นยำเฉลี่ยของการตรวจหาถ้วยรองน้ำยางเมื่อเปลี่ยนจำนวนพื้นที่ย่อยจาก RPN ในขณะตรวจหาวัตถุด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 3 ของ MobileNetV2 | 140 |
| รูปที่ 3.32 ความแม่นยำเฉลี่ยของการตรวจหาถ้วยรองน้ำยางเมื่อเปลี่ยนจำนวนพื้นที่ย่อยจาก RPN ในขณะตรวจหาวัตถุด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 10 ของ MobileNetV2 | 141 |
| รูปที่ 3.33 ความแม่นยำเฉลี่ยในการตรวจหารอยกริดของแต่ละวิธีในกลุ่มทดลองที่ดีที่สุด | 144 |
| รูปที่ 3.34 ความแม่นยำเฉลี่ยในการตรวจหารอยกริดและถ้วยรองน้ำยาง (mAP) ระหว่าง Faster-RCNN กับ YOLOv2 ที่ใช้งานเลเยอร์ทั้งหมดของ MobileNetV2, เมื่อใช้ภาพสี RGB เป็นภาพอินพุต | 145 |
| รูปที่ 3.35 ความแม่นยำเฉลี่ยของการตรวจหารอยกริดและถ้วยรองน้ำยาง (mAP) ด้วย YOLOv2 ที่ใช้งานเลเยอร์เฉพาะเลเยอร์บล็อก 1-3, เลเยอร์บล็อก 1-10 และเลเยอร์ทั้งหมดของ MobileNetV2 เมื่อภาพอินพุตคือภาพสี RGB | 146 |
| รูปที่ 3.36 เปอร์เซ็นต์ของจำนวนแนวรอยกริดที่ปรับตำแหน่งได้สำเร็จ ในกรณีเมื่อกำหนดค่า sr แบบค่าเดียว และผลการปรับตำแหน่งสำเร็จกรณีเมื่อกำหนดค่า sr แบบช่วง | 149 |
| รูปที่ 3.37 ค่าคะแนนคุณภาพแนวรอยกริดสัมพัทธ์เฉลี่ย (RQS) ของแนวรอยกริดที่ปรับตำแหน่งสำเร็จ ในกรณีการกำหนดค่า sr แบบค่าเดียว และผลการปรับตำแหน่งสำเร็จกรณีกำหนดค่า sr แบบช่วง | 150 |
| รูปที่ 3.38 คะแนนความตรงตำแหน่งแนวรอยกริดสัมพัทธ์เฉลี่ย (RLS) ของแนวรอยกริดที่ปรับตำแหน่งได้สำเร็จ ในกรณีกำหนดค่า sr แบบค่าเดียว และผลการปรับตำแหน่งสำเร็จกรณีกำหนดค่า sr แบบช่วง | 150 |
| รูปที่ 3.39 คะแนนความตรงของแนวรอยกริดสัมพัทธ์เฉลี่ย (RSS) ของแนวรอยกริดที่ปรับตำแหน่งสำเร็จ ในกรณีกำหนดค่า sr แบบค่าเดียว และผลการปรับตำแหน่งสำเร็จกรณีใช้ค่า sr แบบช่วง | 151 |
| รูปที่ 3.40 ความแม่นยำเฉลี่ย (Average Precision) ของการตรวจหาตำแหน่งแนวรอยกริดด้วยกรอบสี่เหลี่ยมที่อัตราส่วนการซ้อนทับสำหรับวัตถุผลเท่ากับ 0.5 | 153 |

รายการภาพประกอบ (ต่อ)

| รูป | หน้า |
|--|------|
| รูปที่ 3.41 ระยะทางแบบเฮาส์ดอร์ฟฟ์ (Hausdorff Distance) โดยเฉลี่ย ในหน่วยพิกเซลระหว่างแนวรอยกรีดที่ตรวจพบกับแนวรอยกรีดอ้างอิง..... | 154 |
| รูปที่ 4.1 ภาพขยายบริเวณพื้นที่รอยกรีดเก่า, รอยกรีดล่าสุด, และพื้นที่ที่ยังไม่ถูกกรีด ของภาพถ่ายมุมกว้าง โดยในรูป (ก) แสดงกรอบสี่เหลี่ยมอ้างอิงในรูปแบบเส้นประ, (ข) แสดงกรอบสี่เหลี่ยมผลลัพธ์จากการตรวจหาด้วยภาพโมเดลสี HSV กับภาพโมเดลสี $L^*a^*b^*$ พร้อมทั้งกรอบสี่เหลี่ยมอ้างอิง, และ (ค) แสดงบริเวณที่มีสีแตกต่างกัน โดยพื้นที่ในกรอบหมายเลข 1, 2, 3, และ 4 คือพื้นที่ที่ยังไม่ถูกกรีด, พื้นที่รอยกรีดเก่าที่ติดกับรอยกรีดล่าสุด, พื้นที่รอยกรีดเก่า, และพื้นที่รอยกรีดที่เก่ากว่าตามลำดับ..... | 156 |
| รูปที่ 4.2 ภาพวาดแสดงลักษณะหน้ายางที่พื้นที่รอยกรีดเก่ามีการขยายขนาดเมื่อผ่านการกรีดมากขึ้น.... | 157 |
| รูปที่ 4.3 ภาพขยายของตัวอย่างผลการตรวจหาแนวรอยกรีดในภาพถ่ายระยะใกล้เมื่อใช้ค่า T_{up} และ T_{down} ต่างกัน..... | 165 |
| รูปที่ 4.4 ระยะห่างเฉลี่ยระหว่างแนวรอยกรีดที่ตรวจพบกับแนวรอยกรีดอ้างอิง โดยแสดงช่วงของค่า T_{up}, T_{down} ที่ทำให้ระยะห่างมีค่าน้อยกว่าตำแหน่งอื่นโดยรอบ (ในกรอบเส้นประ)..... | 166 |
| รูปที่ 4.5 ตำแหน่งในภาพถ่ายเมื่อถูกแปลงไปเป็นตำแหน่งในพิกัดสามมิติ ในระนาบความลึกที่ 25 เซนติเมตร กับ 40 เซนติเมตร จากกล้องระยะใกล้..... | 167 |

สัญลักษณ์คำย่อและตัวย่อ

| | |
|---------------|---|
| CNN | Convolutional Neural Network |
| ROI | Region of Interest |
| RPN | Region Proposal Network |
| FOV-V | มุมมองภาพตามแนวตั้งของภาพ (Vertical Field of View) |
| FOV-H | มุมมองภาพตามแนวนอนของภาพ (Horizontal Field of View) |
| FOV-D | มุมมองภาพตามแนวทแยงมุมภาพ (Diagonal Field of View) |
| θ_{yz} | (ครึ่งหนึ่งของ) ขนาดมุมมองภาพในระนาบ YZ |
| θ_{xz} | (ครึ่งหนึ่งของ) ขนาดมุมมองภาพในระนาบ XZ |
| β_{yz} | (ครึ่งหนึ่งของ) ขนาดมุมฉายแสงของไฟส่องสว่างในระนาบ YZ |
| β_{xz} | (ครึ่งหนึ่งของ) ขนาดมุมฉายแสงของไฟส่องสว่างในระนาบ XZ |
| w_b | ขนาดความกว้างของกรอบสี่เหลี่ยม |
| h_b | ขนาดความสูงของกรอบสี่เหลี่ยม |
| θ_b | มุมที่เส้นทแยงมุมหลักทำกับด้านกว้างของกรอบสี่เหลี่ยม |
| w_1 | ความกว้างของภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง |
| h_1 | ความสูงของภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง |
| w_2 | ความกว้างของภาพถ่ายจากกล้องระยะใกล้ |
| h_2 | ความสูงของภาพถ่ายจากกล้องระยะใกล้ |
| w_t | ความกว้างเฉลี่ยของลำต้นที่ปรากฏในภาพ |
| I | ภาพถ่ายใดๆ |
| $p(u, v)$ | ค่าของพิกเซล ณ พิกัดตำแหน่ง (u, v) ในภาพใดๆ |
| R | พื้นที่ย่อย (Region) ใดๆ ในภาพ |
| U | เซตของจำนวนเต็มบวกสำหรับชี้ตำแหน่งในแนวแกนของภาพถ่าย |
| V | เซตของจำนวนเต็มบวกสำหรับชี้ตำแหน่งในแนวแกนตั้งของภาพ |
| C | เซตของพิกัดตำแหน่งในภาพ (Pixel Coordinate) |
| z_0 | ระยะความลึกตั้งฉากจากกล้องถ่ายภาพมุมกว้างถึงผิวหน้ายาง |
| z_2 | ระยะความลึกตั้งฉากจากกล้องระยะใกล้ถึงผิวหน้ายาง |
| (u, v) | พิกัดตำแหน่งในภาพ ณ ตำแหน่งคอลัมน์ (แกนนอน) เท่ากับ u และตำแหน่งแถว (แกนตั้ง) เท่ากับ v |
| (x, y, z) | พิกัดตำแหน่งสามมิติ |
| T | เซตของตำแหน่งพิกเซลในภาพแทนแนวรอยกรีด |
| S/2 | ระบบกรีดลงแบบเวียนขวาที่มีความยาวรอยกรีดเท่ากับครึ่งเส้นรอบวงลำต้น |
| S/3 | ระบบกรีดลงแบบเวียนขวาที่มีความยาวรอยกรีดเท่ากับหนึ่งในสามของเส้นรอบวงลำต้น |
| S/4 | ระบบกรีดลงแบบเวียนขวาที่มีความยาวรอยกรีดเท่ากับหนึ่งในสี่ของเส้นรอบวงลำต้น |

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มา

การทำสวนยางพาราในเขตพื้นที่ภาคใต้ของไทย เกษตรกรนิยมเก็บเกี่ยวน้ำยางจากต้นยางพาราด้วยวิธีการกรีดยาง โดยอาศัยมีดกรีดยางแบบดั้งเดิม (“มีดเจ๊ะบง” [53]) ซึ่งมีรูปร่างที่เอื้อต่อการสร้างรอยกรีดบนลำต้นยางพารา การกรีดยางด้วยมีดอาศัยประสบการณ์และความชำนาญของเกษตรกรผู้กรีดยาง เพื่อให้สามารถไหลตามรอยกรีดได้และไม่ให้หน้ายางเกิดความเสียหาย อย่างไรก็ตามผู้กรีดยางที่มีประสบการณ์อาจสามารถทำให้หน้ายางเสียหายได้จากการกรีดยางได้เช่นกัน เนื่องจากลักษณะเปลือกหน้ายางพาราที่ต่างกันไปในแต่ละต้น ตัวอย่างความเสียหายของหน้ายางจากการกรีดยางที่เกิดขึ้นบ่อยจากผู้กรีดยางที่มีประสบการณ์ เช่น ในกรณีเปลือกยางมีความหนาไม่สม่ำเสมอ ทำให้รอยกรีดลึกเข้าไปในเนื้อไม้ไม่น้อยเกินไปจนน้ำยางไหลออกมาน้อยกว่าที่ควรจะเป็น หรือรอยกรีดลึกมากเกินไปจนถึงชั้นเนื้อเยื่อเจริญ (Cambium Layer) ทำให้ต้นยางพาราไม่สามารถฟื้นฟูสภาพหน้ายางสำหรับรอบกรีดถัดไปได้ รอยกรีดที่แนะนำมีความลึกก่อนถึงชั้นเนื้อเยื่อเจริญประมาณ 0.2 มิลลิเมตร [1] หรือรอยกรีดควรลึกประมาณ 3 ถึง 5 มิลลิเมตรจากผิวเปลือกชั้นนอกลำต้น ทั้งนี้ความลึกของรอยกรีดที่เหมาะสมจะต่างกันไปตามสายพันธุ์และอายุของต้นยางพารา ผู้กรีดยางที่มีประสบการณ์จะประมาณความลึกของรอยกรีดด้วยการคาดคะเนด้วยสายตาประกอบกับความลึกของแนวรอยกรีดล่าสุด, ความเอียงของแนวรอยกรีดที่เหมาะสมกับความชันของน้ำยางที่ทำให้หน้ายางสามารถไหลตามแนวรอยกรีดได้ดี และการประมาณน้ำหนักของมือกรีดส่งผลต่อความลึกของแนวรอยกรีด และยกตัวอย่างกรณีนี้ที่ผู้กรีดยางขาดประสบการณ์ เช่น การกรีดยางที่รอยกรีดมีความลึกไม่สม่ำเสมอซึ่งส่งผลให้น้ำยางไหลน้อยกว่าที่ควร หรือการกรีดยางที่ทำให้หน้ายางไหลออกจากรอยกรีด ไม่สามารถไหลเก็บสะสมในภาชนะรองรับน้ำยางได้ เป็นต้น

เกษตรกรจะเริ่มการกรีดยางในช่วงเวลา 3.00 นาฬิกา ถึง 6.00 นาฬิกา เนื่องจากต้นยางพาราสามารถผลิตน้ำยางได้สูงกว่าช่วงเวลาอื่นของวัน ผลจากปัจจัยของอุณหภูมิและความชื้นในพื้นที่ภาคใต้ [24, 63] แต่ในทางปฏิบัติการเลือกเวลากรีดยางของเกษตรกรยังมีปัจจัยเสริมอื่น เช่น เวลาหรือระยะทางในการเดินทางไปยังสวนยางพารา, สภาพพื้นที่ภูมิประเทศภายในสวนยาง, สภาพอากาศ, และจำนวนต้นยางพารา เป็นต้น จากปัจจัยต่างๆ ข้างต้น เกษตรกรจึงเลือกกรีดยางก่อนเวลา เป็นผลให้เกษตรกรต้องกรีดยางพาราในสภาพแวดล้อมที่มีแดด โดยจะอาศัยแสงสว่างจากไฟฉายขนาดพกพาแบบคาดศีรษะที่สามารถใช้งานได้นานพอ จนจะเข้าสู่ช่วงเช้าเพื่ออาศัยแสงสว่างจากดวงอาทิตย์ในการทำงานต่อไป การปฏิบัติงานในสภาพแสงน้อยส่งผลต่อคุณภาพของรอยกรีด เพราะเกษตรกรต้องอาศัยการทำงานร่วมกันของการมองเห็นและประสาทสัมผัสที่มือกรีด นอกจากนี้เกษตรกรยังต้องเสี่ยงภัยอันตรายจากสัตว์ภายในสวนยางพาราที่ออกหากินในเวลากลางคืน เช่น ยุง มด งู ตะขาบ หรือสัตว์มีพิษอื่น เป็นต้น

กระบวนการกรีดยางด้วยแรงงานเกษตรกรมีรายละเอียดดังนี้ เกษตรกรกรีดยางพารา 1 ต้น ใช้เวลาประมาณ 1 ถึง 2 นาที ขึ้นอยู่กับความชำนาญของผู้กรีดยาง เมื่อเสร็จสิ้นการกรีดยางในต้นหนึ่งๆ เกษตรกรจะเดินเท้าไปยังต้นยางพาราต้นถัดไปเพื่อกรีดยางครบทุกต้นทั่วทั้งสวนยางพารา โดยในเนื้อที่ 1 ไร่ ซึ่งมีจำนวนต้น

ยางพาราประมาณ 80-90 ต้น เกษตรกรจะใช้เวลากกรี๊ดโดยรวมประมาณ 1-2 ชั่วโมงต่อไร่ ซึ่งส่วนมากเกษตรกรจะต้องกรี๊ดสวนยางในขนาดที่มากกว่า 1 ไร่ ทำให้เกษตรกรประสบกับความเหนื่อยล้า เนื่องจากการเดินทั่วทั้งสวนยางพาราในเนื้อที่ขนาดหลายไร่, การใช้สายตาในสภาพแสงที่ไม่อำนวยนักในขณะกรี๊ด, และผลจากตำแหน่งความสูงของรอยกรี๊ดที่ไม่เท่ากันในหน้ายางแต่ละต้น ทำให้เกษตรกรอาจต้องก้มลงต่ำเพื่อกรี๊ดในบางต้น หรือต้องเอื้อมมือกรี๊ดขึ้นสูงในบางลำต้น ซึ่งล้วนส่งผลต่อสุขภาพในระยะยาวของผู้กรี๊ด

เกษตรกรจะเก็บน้ำยางหลังจากการกรี๊ดเสร็จสิ้น โดยจะทิ้งช่วงรอน้ำยางนานประมาณ 2 ถึง 4 ชั่วโมงหลังกรี๊ดเสร็จ เพื่อให้น้ำยางไหลเข้าสู่ระบบในถ้วยรองน้ำยางที่แขวนอยู่ติดกับแต่ละต้นยางพารา จากนั้นเกษตรกรจึงจะเก็บน้ำยางจากภาชนะสะสมน้ำยางจากต้นยางพาราแต่ละต้น ซึ่งกระบวนการทำงานของเกษตรกรยังคงอาศัยการเดิน เก็บน้ำยางจากภาชนะที่แขวนติดกับลำต้น เทสะสมเข้าสู่ภาชนะที่ใหญ่ที่สามารถถือได้ เกษตรกรจะถือภาชนะนี้เก็บสะสมน้ำยางจากต้นยางพาราทุกต้นทั่วทั้งสวนยางพารา ทำให้เกษตรกรยังคงประสบปัญหาความเหนื่อยล้าเช่นเดียวกันกับขั้นตอนการกรี๊ด ที่ภาชนะเก็บสะสมน้ำยางในแต่ละต้นอาจอยู่ในตำแหน่งสูงหรือต่ำ แตกต่างกันไปตามตำแหน่งของรอยกรี๊ด [71] ทั้งนี้ในขณะที่เก็บน้ำยางหากเกิดกรณีที่น้ำยางในถ้วยรองน้ำยางถูกเทออกไม่หมด หรือมีบางส่วนติดกับผิวภายในถ้วยรองน้ำยาง เกษตรกรจะใช้ไม้กวาดน้ำยางเข้าช่วย โดยใช้มือข้างหนึ่งจับถ้วยรองน้ำยาง ในขณะที่มืออีกข้างหนึ่งใช้ไม้กวาดน้ำยาง ซึ่งเพิ่มความยุ่งยากให้กับการทำงาน เพราะเกษตรกรต้องพกไม้กวาดน้ำยางกับภาชนะสะสมน้ำยางไปด้วยขณะเดิน ทำให้การเก็บน้ำยางล่าช้ากว่าเดิม และเสี่ยงต่อการทำน้ำยางหกเสียหายเนื่องจากสภาพพื้นที่ในสวนยางพาราเสี่ยงต่อการเกิดอุบัติเหตุทุกหล่ม

นอกจากนี้เกษตรกรประสบปัญหาความเสี่ยงต่อการเปลี่ยนแปลงของสภาพอากาศ ลม, ฝน หรือการทำงานที่ล่าช้าเกินไป ในช่วงการกรี๊ดที่จะทำให้ น้ำยางไม่สามารถไหลตามรอยกรี๊ดได้ หรือในช่วงรอน้ำยางไหลสะสมเข้าสู่ถ้วยรองน้ำยาง เช่น เกิดลมพัดแรงทำให้ถ้วยรองน้ำยางหกเสียหาย, ฝนตกขณะรอน้ำยางไหล ทำให้ผลผลิตน้ำยางถูกเจือปนด้วยน้ำฝน หรือเกิดการแห้งหรือจับตัวกันของน้ำยาง เป็นต้น

สามารถสรุปประเด็นปัญหาข้างต้น ได้แก่ ความชำนาญของผู้กรี๊ด, ความเหนื่อยล้าของเกษตรกร, สภาพแวดล้อมในสวนยางพาราในเวลากลางคืนที่ไม่เอื้อต่อการทำงาน ผู้เขียนจึงเล็งเห็นถึงความสามารถของเทคโนโลยีระบบหุ่นยนต์อัตโนมัติ ที่สามารถประยุกต์ใช้กับการทำสวนยางพาราแทนที่การใช้แรงงานคนได้ เนื่องจากเทคโนโลยีหุ่นยนต์มีข้อได้เปรียบกว่าแรงงานมนุษย์ในด้าน ความเร็ว, ความชำนาญ, ความเหนื่อยล้า ซึ่งถูกกำหนดด้วยโปรแกรมควบคุมการทำงานของหุ่นยนต์ ให้สามารถทำงานในปริมาณที่มากกว่าและละเอียดกว่ามนุษย์ โดยเฉพาะอย่างยิ่งการทนต่อสภาพแวดล้อมของระบบหุ่นยนต์ แต่เนื่องจากการพัฒนาระบบหุ่นยนต์ทั้งระบบมีขอบเขตการศึกษาที่กว้าง วิทยานิพนธ์นี้จึงจำกัดไว้เพียงการพัฒนาาระบบตรวจรู้ตำแหน่งของเป้าหมาย ซึ่งเป็นส่วนหนึ่งของระบบหุ่นยนต์ทำสวนยางพาราอัตโนมัติ ได้แก่ การตรวจหาตำแหน่งของแนวรอยกรี๊ดบนหน้ายางและตำแหน่งถ้วยรองน้ำยาง เพื่อประโยชน์ในการพัฒนาระบบทำสวนยางพาราแบบอัตโนมัติต่อไป

1.2 การตรวจเอกสาร

1.2.1 ระบบหุ่นยนต์อัตโนมัติในงานเกษตรกับการปรับใช้สำหรับสวนยางพารา

เทคโนโลยีเครื่องจักรกลและระบบอัตโนมัติในงานเกษตรที่มีใช้ในปัจจุบัน สามารถจำแนกตามลักษณะการทำงานออกเป็น 3 ประเภท ได้แก่

- 1 ระบบเครื่องจักรกล (Machine) ซึ่งใช้แรงงานคนปฏิบัติงานโดยมีปฏิสัมพันธ์กับเครื่องจักรเพื่อให้เครื่องจักรทำงาน (Manual System)
- 2 เครื่องจักรกลกึ่งระบบอัตโนมัติ (Semi-Automated System) เครื่องจักรสามารถดำเนินงานได้เองบ้างแต่ยังคงอาศัยแรงงานคนปฏิสัมพันธ์กับเครื่องจักร
- 3 ระบบอัตโนมัติแบบสมบูรณ์ หรือหุ่นยนต์อัตโนมัติ (Automation System) เป็นระบบที่ไม่มีมนุษย์เข้ามามีส่วนร่วมตลอดการทำงาน

การวิจัยวรรณกรรมในวิทยานิพนธ์ฉบับนี้จะเน้นในด้านระบบอัตโนมัติแบบสมบูรณ์ ซึ่งเป็นระบบที่ถูกออกแบบให้ทำงานโดยไม่พึ่งการดำเนินงานจากแรงงานมนุษย์ และถูกใช้ในงานเกษตรอย่างแพร่หลายโดยเฉพาะในงานเก็บเกี่ยวพืชผล ซึ่งสามารถทำงานในสภาพแวดล้อมทั้งในสถานที่ปลูกภายในเรือนเพาะ (Greenhouse) และในสภาพแวดล้อมเปิด (Outdoor) สอดคล้องกับการพัฒนาระบบทำสวนยางอัตโนมัติ ตัวอย่างเช่น ระบบสำหรับเก็บเกี่ยวผลองุ่น [57], ผลแอปเปิล [70], ผลแตงกวา [18] เป็นต้น ซึ่งระบบหุ่นยนต์เก็บเกี่ยวพืชผลอัตโนมัติเหล่านี้ ต่างมีส่วนประกอบและความสามารถที่คล้ายกัน ได้แก่

- ระบบขับเคลื่อน เหมือนการเดินทางของเกษตรกร ระบบหุ่นยนต์สามารถเคลื่อนที่ในแนวแปลงปลูก ซึ่งสอดคล้องกับการปลูกพืชแบบแถว (Row Planting) ของพืชสวน หรือพืชไร่ผลไม้ทั่วไป เช่น สวนแอปเปิล, สวนสตอเบอรี่, แปลงปลูกพริก เป็นต้น
- อุปกรณ์เก็บเกี่ยว หุ่นยนต์จะถูกติดตั้งเครื่องมือสำหรับเก็บเกี่ยวผลผลิต ซึ่งจะนิยมใช้งานแขนหุ่นยนต์ (Robot Arm) ร่วมกับมือจับ (Gripper) ที่ติดตั้งในส่วนปลายแขนกล โดยมือจับถูกออกแบบให้มีความสามารถเก็บเกี่ยวพืชผลแต่ละชนิดได้โดยเฉพาะ
- ระบบค้นหาตำแหน่งเป้าหมาย ทำหน้าที่ค้นหาตำแหน่งผลผลิต เพื่อให้อุปกรณ์เก็บเกี่ยวเข้าเก็บเกี่ยวผลผลิต โดยการตรวจหาตำแหน่งเป้าหมายจะทำผ่านเซนเซอร์หลากหลายชนิด ยกตัวอย่างเช่น เซนเซอร์วัดระยะ, เซนเซอร์รับภาพหรือกล้องถ่ายภาพ, เซนเซอร์อ่านค่าข้อมูลตำแหน่งชนิดอื่นๆ ที่ทำงานในย่านความถี่ที่แตกต่างออกไป การเลือกชนิดเซนเซอร์ขึ้นอยู่กับสภาพแวดล้อมที่หุ่นยนต์ปฏิบัติงาน, มุมมองหรือระนาบทำงานของหุ่นยนต์ ให้สอดคล้องกับวิธีตรวจหาเป้าหมายที่สามารถใช้งานและประมวลผลข้อมูลของเซนเซอร์ ได้อย่างมีประสิทธิภาพ

ดังนั้นการออกแบบระบบทำสวนยางพาราอัตโนมัติ โดยมีส่วนประกอบหลักอ้างอิงจากระบบเก็บเกี่ยวพืชผลอัตโนมัติข้างต้นนั้นมีความเป็นไปได้ เนื่องจากการทำสวนยางพารามีความคล้ายคลึงกับการเก็บเกี่ยวพืชผลจากแปลงปลูกในแนวแถว เช่น การกรีดยางและเก็บน้ำยางจากถ้วยรองน้ำยางของเกษตรกร จากแต่ละต้นหรือแถว, การกรีดยางโดยใช้มีดกรีดยางซึ่งสามารถพัฒนาระบบแขนหุ่นยนต์ให้มีอุปกรณ์ปลายแขนสำหรับกรีดยางหรือ

เก็บน้ำยางได้ กล่าวคือ ระบบทำสวนยางพาราอัตโนมัติ อาจประกอบไปด้วย ระบบการเคลื่อนที่จะทำให้หุ่นยนต์สามารถไปยังต้นยางพาราแต่ละต้นตามแนวแถวได้, แขนหุ่นยนต์ที่มีอุปกรณ์ปลายแขนสำหรับกรีตและเก็บน้ำยางจากถ้วยเก็บน้ำยาง และระบบตรวจหาตำแหน่งของรอยกรีตและถ้วยรองน้ำยาง

1.2.2 สวนยางพาราและต้นยางพารา

ตามคำแนะนำของกรมวิชาการเกษตรที่ให้เกษตรกรปลูกยางพาราสายพันธุ์สถาบันวิจัยยาง 251 (RRIT-251) หรือพันธุ์อาร์อาร์ไอเอ็ม 600 (RRIM-600) ของประเทศมาเลเซีย [61] ทำให้ทั้งสองพันธุ์เป็นที่นิยมปลูกในเขตจังหวัดภาคใต้ของไทย โดยการปลูกยางพาราจะปลูกเป็นแนวแถว ซึ่งในกรณีที่ไม้ปลูกพืชแซมระหว่างต้น ระยะห่างระหว่างต้นยางพาราในแถวจะห่างประมาณ 2.5 ถึง 3.0 เมตร และระยะห่างระหว่างแถวประมาณ 6 ถึง 8 เมตร ทำให้เนื้อที่ 1 ไร่ สามารถปลูกยางพาราได้ประมาณ 80-90 ต้น

ลำต้นยางพาราจะมีลักษณะตรงตั้งขึ้นในแนวตั้งฉากกับระนาบพื้นดิน หรือทำมุมเอียงเฉลี่ยไม่เกิน 15 องศาจากแนวตั้ง ในกรณีภูมิประเทศพื้นที่ปลูกเป็นพื้นราบ แต่ทั้งนี้ขึ้นกับสภาพลมและทิศทางของแสงแดดในแต่ละท้องที่ ส่งผลต่อความเอียงของลำต้นเมื่อเจริญเติบโต ลักษณะผิวเปลือกลำต้นจะมีสี ริวรอยที่ต่างกันออกไปตามสายพันธุ์ นอกจากนี้ความชื้นในอากาศจะส่งผลให้พืชกาฝากเจริญเติบโตในบริเวณลำต้นได้ เช่น คะไครน้ำ, มอส เป็นต้น ส่งผลต่อลักษณะพื้นผิวของเปลือกยาง

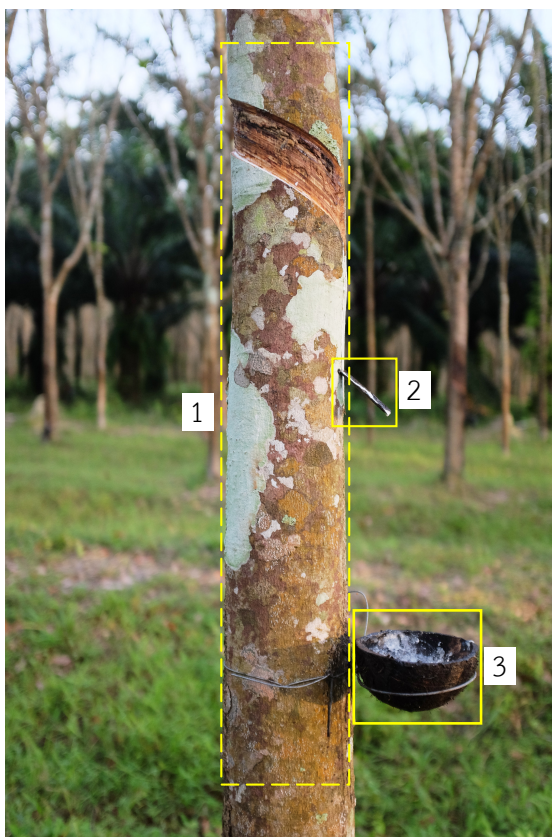
1.2.3 หน้ำยางและรอยกรีตในระบบกรีตลงแบบเวียนขวา

เกษตรกรจะเปิดกรีตยางพาราครั้งแรกเมื่อต้นยางพารามีอายุประมาณ 7 ปี หรือลำต้นมีเส้นรอบวงไม่น้อยกว่า 50 เซนติเมตร ซึ่งเกษตรกรในภาคใต้เลือกใช้ระบบกรีตลงแบบเวียนขวา (Spiral Downward Tapping System) โดยจะสร้างรอยกรีตแรก (“เปิดกรีต”) บนลำต้นที่ความสูง 150 เซนติเมตรจากพื้นดิน การกรีตแต่ละครั้งจะทำให้ความสูงของรอยกรีตลดลง และเมื่อรอยกรีตสูง 50 เซนติเมตรจากระดับพื้นดิน เกษตรกรจึงจะเปลี่ยนเปิดกรีตหน้ำยางใหม่ สาเหตุระบบกรีตลงแบบเวียนขวาได้รับความนิยมมากกว่าระบบกรีตอื่น เช่น การกรีตรูปตัววี (V-Cut) หรือการกรีตแบบสั้น (Mini-Cut) เนื่องจากการกรีตลงแบบเวียนขวานวรอยกรีตจะตัดผ่านท่อน้ำยางภายในเปลือกยางได้มาก ซึ่งส่งผลให้ได้น้ำยางในปริมาณที่มากกว่าในขณะที่สูญเสียเปลือกยางน้อยกว่าระบบกรีตอื่น

การกรีตแบบระบบกรีตลงแบบเวียนขวา [1, 8, 9, 23, 58, 75] เกษตรกรผู้กรีตจะสร้างรอยกรีตบนลำต้นยางพารา หรือ “หน้ำยาง” (“Tapping Panel”) แสดงในรูปที่ 1.1 แนวนรอยกรีตจะเริ่มต้นจากตำแหน่งซ้ายมือของผู้กรีตลาดลงต่ำไปตามส่วนโค้งของลำต้นจนสิ้นสุดในตำแหน่งทางขวามือของผู้กรีต โดยรอยกรีตจะตัดผ่านท่อน้ำยาง (Latex Vessels) ภายในเปลือกของลำต้นยางพาราก่อนถึงชั้นเนื้อเยื่อเจริญ (Cambium) เกษตรกรจะสร้างรอยกรีตให้มีลักษณะเป็นร่องที่มีความลาดเอียงเฉลี่ย 30 ถึง 45 องศากับแนวระดับ เพื่อให้หน้ำยางที่ซึมออกมาจากท่อน้ำยางสามารถไหลไปตามรอยกรีตจนถึงปลายรอยกรีตทางด้านขวา โดยในส่วนปลายรอยกรีตเกษตรกรจะสร้างร่องในแนวตั้งด้วยมีดกรีต ยาวประมาณ 5 ถึง 15 เซนติเมตร แล้วใช้ด้ามจับมีดกรีตยางตอกรางเหล็กขนาดเล็ก (ลิ้นรองน้ำยาง) เข้ากับลำต้น เพื่อให้หน้ำยางสามารถไหลไปสะสมที่ถ้วยรองน้ำยางที่แขวนติดกับลำต้นได้ลิ้นรองน้ำยางได้

ความกว้างรอยกรีดของระบบกรีดลงแบบเวียนขวา มีขนาดตั้งแต่หนึ่งในสี่ของรอบวงลำต้น (อักษรย่อ S/4) จนถึงครึ่งหนึ่งของรอบวงของลำต้น (อักษรย่อ S/2) ตามการจัดสรรหน้ายาง ในขณะที่ความลึกรอยกรีดสำหรับยางพาราอายุ 7 ปีขึ้นไปจะอยู่ระหว่าง 3.00 ถึง 5.00 มิลลิเมตรจากเปลือกชั้นนอกของลำต้น แต่จากงานศึกษา [1, 58] พบว่าความลึกของรอยกรีดควรอ้างอิงจากชั้นเนื้อเยื่อเจริญ โดยรอยกรีดควรเว้นไม่กรีดถึงชั้นเนื้อเยื่อเจริญ มีระยะความลึกก่อนถึงเนื้อเยื่อเจริญไว้อย่างน้อย 0.2-0.5 มิลลิเมตร แต่ในทางปฏิบัติของเกษตรกรผู้กรีดยาง ความลึกของรอยกรีดจะมาจากการคาดคะเน เช่นเดียวกับกับความเอียงของรอยกรีด หากผู้กรีดยางพบว่ารอยกรีดที่กรีดได้เมื่อรอไปสักระยะเวลาหนึ่ง หากน้ำยางยังไม่สามารถไหลตามแนวรอยกรีดได้ ผู้กรีดยางจะทำการกรีดใหม่เพื่อปรับระดับความเอียง หรือความลึกของแนวรอยกรีด ซึ่งเป็นการแก้ปัญหาเฉพาะหน้าโดยอาศัยประสบการณ์ของผู้กรีดยาง

การกรีดในแต่ละครั้งเกษตรกรผู้กรีดยางจะสร้างรอยกรีดใหม่ที่ทับรอยกรีดเดิม โดยกรีดเอาเปลือกของหน้ายางที่ติดกับรอยกรีดเดิมออก ซึ่งผู้กรีดยางจะประมาณความลึกของรอยกรีดใหม่อิงจากแนวรอยกรีดเดิมร่วมกับการคาดคะเนความลึกของเปลือกไม้ตามแนวรอยกรีดใหม่ เนื่องจากสภาพพื้นผิวลำต้นยางพาราที่ขรุขระแตกต่างกันไปในแต่ละต้น การกรีดในแต่ละครั้งจะสูญเสียเปลือกไม้เป็นระยะความสูงเฉลี่ย 3 ถึง 5 มิลลิเมตร หน้ายางเมื่อผ่านการกรีดไปสักระยะเวลาหนึ่ง พื้นที่หน้ายางจะแบ่งออกเป็น 2 บริเวณ ซึ่งสามารถสังเกตเห็นได้ชัด คือพื้นที่ที่ถูกกรีด (พื้นที่รอยกรีดเก่า) และพื้นที่ที่ยังไม่กรีด (เปลือกของลำต้นยางพาราตามปกติ) ดังแสดงในรูปที่ 1.2 การกรีดแบบทิศทางลงจะสร้างหน้ายางที่มีพื้นที่ที่ถูกกรีดขยายเพิ่มมากขึ้น การขยายตัวของพื้นที่ที่ถูกกรีดจะเริ่มตั้งแต่ความสูง 150 เซนติเมตรจากระดับพื้นดิน ขยายเพิ่มมากขึ้นตามจำนวนวันที่เปิดกรีด และจะสิ้นสุดเมื่อรอยกรีดมีความสูงประมาณ 50 เซนติเมตรจากระดับพื้นดิน จากนั้นเกษตรกรผู้กรีดยางจึงจะเปิดกรีดหน้ายางใหม่ เพื่อให้หน้ายางที่ถูกกรีดสร้างเปลือกชั้นนอกและท่อน้ำยางใหม่โดยอาศัยเนื้อเยื่อเจริญ ซึ่งหากเนื้อเยื่อเจริญเสียหายจากการกรีด เปลือกยางบริเวณนั้นจะไม่ถูกสร้างใหม่ ทำให้ไม่สามารถกรีดยางในหน้ายางดังกล่าวอีกได้ โดยหน้ายางหนึ่งๆ รองรับกรีดได้นาน 2-3 ปี



รูปที่ 1.1: หน้ายาง (หมายเลข 1), ลิ้นรองหน้ายาง (หมายเลข 2), และถ้วยรองหน้ายาง (หมายเลข 3)



รูปที่ 1.2: หน้ายางเมื่อผ่านการกรีดเป็นระยะเวลาหนึ่ง (รูปซ้าย) และหน้ายางที่เปิดกรีดใหม่ (รูปขวา)

1.2.4 ถ้วยรองน้ำยาง

ถ้วยรองน้ำยาง หรือ “จอกยาง” เป็นภาชนะรองรับน้ำยางจากรอยกรีด โดยจะถูวางบนห่วงโลหะ ที่ถูกมัดติดไว้กับลำต้นยางพาราตำแหน่งใต้ลำต้นรองน้ำยาง ในอดีตเกษตรกรจะใช้วัสดุที่ทำได้ง่ายหรือวัสดุเหลือใช้อื่นเป็นถ้วยรองน้ำยาง เช่น กะลามะพร้าว (ตามรูปที่ 1.1), ฤงพลาสติกหรือภาชนะรีไซเคิลอื่นๆ แต่ในปัจจุบันถ้วยรองน้ำยางถูกพัฒนาให้มีประสิทธิภาพมากขึ้น เช่น ถ้วยรองน้ำยางพลาสติก, ถ้วยรองน้ำยางเซรามิกที่มีพื้นผิวไม่เกาะน้ำยาง [3] มีข้อดีกว่าถ้วยรองน้ำยางจากกะลามะพร้าวซึ่งมีผิวด้านในไม่เรียบ ทำให้สูญเสียน้ำยางในระหว่างการเทเก็บ หรือเกษตรกรไม่สามารถกวาดน้ำยางออกหมดได้ด้วยไม้กวาดน้ำยาง

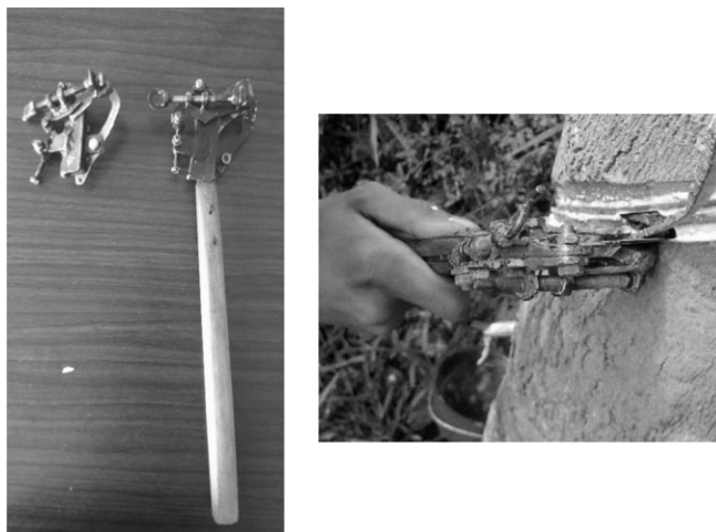
เกษตรกรจะเลือกชนิดถ้วยรองน้ำยางโดยพิจารณาจากข้อดีข้อเสีย เช่น ความคงทน, ราคา ซึ่งถ้วยเก็บน้ำยางชนิดพลาสติกได้รับความนิยมมากขึ้นในช่วงหลัง เพราะมีราคาถูกและมีประสิทธิภาพเหนือกว่ากะลามะพร้าว เกษตรกรมักเลือกใช้ถ้วยพลาสติกขนาด 0.7-1.0 ลิตร (ตัวอย่างตามรูปที่ 2.14) สอดคล้องกับปริมาณน้ำยางที่ได้ต่อต้น และสะดวกต่อการหยิบจับในขั้นตอนการเก็บน้ำยาง ถ้วยรองน้ำยางขนาด 0.7-1.0 ลิตรจะมีเส้นผ่านศูนย์กลางปากถ้วยระหว่าง 13-16 เซนติเมตร เส้นผ่านศูนย์กลางฐานถ้วยระหว่าง 5-7 เซนติเมตร และมีความสูงระหว่าง 10-13 เซนติเมตร

1.2.5 เทคโนโลยีการทำสวนยางพาราจนถึงปัจจุบัน

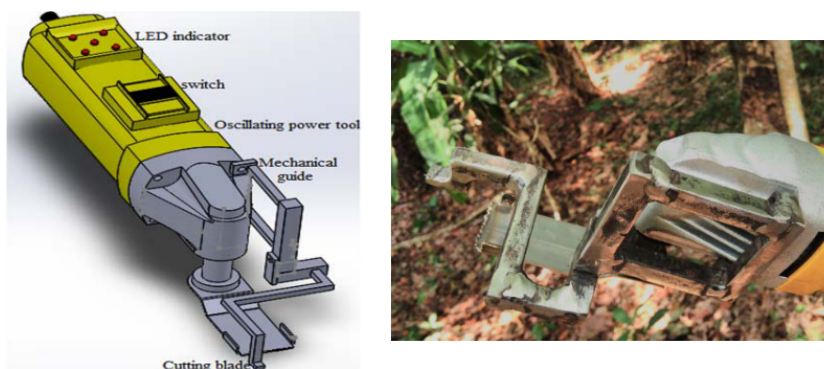
ระบบหุ่นยนต์อัตโนมัติยังไม่ถูกนำมาปรับใช้กับการทำสวนยางเมื่อเทียบกับหุ่นยนต์เก็บเกี่ยวผลผลิตทางการเกษตร โดยเทคโนโลยีในการทำสวนยางมีเพียงระบบที่ใกล้เคียงความเป็นระบบอัตโนมัติ ซึ่งยังอาศัยการทำงานร่วมกับแรงงานมนุษย์ ยกตัวอย่างเช่น การพัฒนาเม็ดกรีดแบบใหม่ที่เกษตรกรผู้กรีดสามารถกรีดโดยควบคุมความลึกรอยกรีดได้ง่ายขึ้น [74], มีดกรีดยางไฟฟ้าเพื่อทดแทนการขาดประสบการณ์และความชำนาญของผู้กรีด [72] ทำงานด้วยใบมีดที่ออกแบบมาทดแทนมีดแบบเดิม ทำงานด้วยมอเตอร์ไฟฟ้าและแบตเตอรี่ โดยผลการทดลองแสดงให้เห็นว่ามีดกรีดยางที่ถูกออกแบบมาใหม่นี้สามารถใช้ทดแทนมีดกรีดยางแบบเดิม (มีดเจ้ะบง) ได้ไม่สมบูรณ์นัก เพราะยังคงมีปัญหาประเด็นความลึกของรอยกรีดที่มีดทำได้ซึ่งส่งผลต่อความเสียหายของหน้ายาง อีกทั้งมีดดังกล่าวยังคงพึ่งพาผู้กรีด ในการกำหนดจุดกรีดและมุมมองของรอยกรีด

เครื่องกรีดยางรางเลื่อนระบบไฟฟ้าแบบติดตามลำต้น [54] เพื่อทดแทนแรงงานคนในขั้นตอนการกรีด แต่ยังมีจุดอ่อนที่สำคัญคือการนำมาใช้งานจริงโดยเฉพาะจำนวนเครื่องที่ส่งผลต่อการจัดการและการบำรุงรักษา เพราะเครื่องถูกออกแบบมาเพื่อติดตามลำต้นยางพารา ทำให้จำนวนเครื่องที่ต้องการในการทำสวนยางจริงจะเท่ากับจำนวนต้นยางพาราทั้งหมด เครื่องที่ต้นยางพาราแต่ละต้นต่างมีมีดกรีดยางแยกกันตามจำนวนเครื่อง ซึ่งระบบไฟฟ้าที่ต้องการสายเชื่อมต่อระหว่างเครื่องทุกเครื่อง หรือแม้ใช้พลังงานจากแบตเตอรี่ก็จะส่งผลถึงจำนวนแบตเตอรี่ ความยากในการบำรุงรักษาทุกๆ เครื่องในสวนยางจริง รวมถึงเครื่องไม่ได้ออกแบบให้เกษตรกรเก็บน้ำยางได้ง่าย

เมื่อไม่นานมานี้เกิดการพัฒนารูปแบบหนึ่งของระบบทำสวนยางพาราอัตโนมัติแบบสมบูรณ์มากขึ้น เช่น เครื่องกรีดรางเลื่อนแบบระบบไฟฟ้าของ [81] ปรับปรุงข้อด้อยของเครื่องกรีดยางอัตโนมัติติดตามลำต้น โดยออกแบบรางแขนที่ทำให้เครื่องกรีดยางเพียงเครื่องเดียว สามารถเคลื่อนที่ได้ตามแนวแถวต้นยางพารา แต่ลักษณะรอยกรีดที่ได้จากเครื่องยังคงไม่สมบูรณ์มากนัก ซึ่งทำให้หน้ายางยังคงได้รับความเสียหาย,



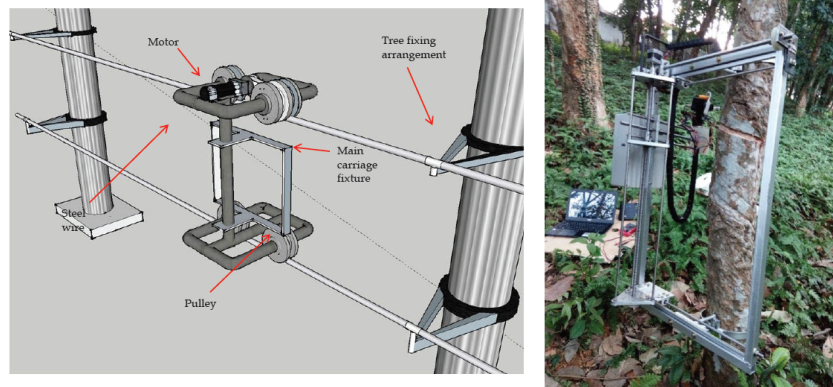
รูปที่ 1.3: มีดกรีดยางแบบกำหนดความลึกได้ [74]



รูปที่ 1.4: มีดกรีดยางมอเตอร์ไฟฟ้า [72]

การใช้งานระบบคอมพิวเตอร์สำหรับพัฒนาระบบกรีดยางเมื่องานวิจัยของ [44] ได้นำเสนอการใช้กล้องถ่ายภาพสี ภาพตรวจหาต้นยางพาราในภาพสำหรับติดตั้งให้กับหุ่นยนต์ โดยจะทำการตรวจหาวัตถุที่สร้างขึ้นเองที่ถูกลำดับไปติดไว้บนลำต้นของต้นยางพารา เพื่อให้สามารถคำนวณระยะห่างของต้นยางพารากับกล้องได้ผ่านโมเดลกล้องรูเข็ม (Pinhole Camera) แต่งานวิจัยดังกล่าวไม่ได้กล่าวถึงสภาพแวดล้อมที่มีต้นยางพาราขณะกรีด เพราะสภาพของแสงมีผลต่อกล้องถ่ายภาพสีที่ใช้หาวัตถุที่นำไปติดไว้บนลำต้นยางพารา

งานวิจัย [82] นำเสนอหุ่นยนต์ที่เคลื่อนที่ได้อย่างอัตโนมัติ (Unmanned Ground Vehicle, UGV) ภายในสวนยางพาราในสภาพแวดล้อมที่ไม่ต้องการแสงสว่าง ซึ่งทำได้โดยใช้เซนเซอร์ตรวจหาลำต้นยางพาราที่ทำงานในย่านความถี่อินฟราเรด (2D LiDAR) หุ่นยนต์สามารถเคลื่อนตัวด้วยล้อผ่านมอเตอร์ไฟฟ้าและแบตเตอรี่ เข้าประจำตรงลำต้นยางพาราเป้าหมายได้ที่ระยะห่างประมาณ 1.00 เมตร



รูปที่ 1.5: เครื่องกรีดยางรางเลื่อนระบบไฟฟ้าแบบเคลื่อนที่ได้ตามแนวแถว [81]



รูปที่ 1.6: หุ่นยนต์สำรวจสวนยางพาราอัตโนมัติ (UGV) [82]

1.2.6 ระบบมองเห็นของระบบหุ่นยนต์อัตโนมัติในงานเกษตร

ระบบหุ่นยนต์อัตโนมัติในงานเก็บเกี่ยวหนึ่งๆ อาจใช้งานเซนเซอร์หลากหลายชนิดตามความซับซ้อนของวิธีการทำงานของหุ่นยนต์ ยกตัวอย่างหลักการทำงานของระบบที่ใช้ทั้งเซนเซอร์วัดระยะแบบอินฟราเรด (Infrared Proximity Sensor) ร่วมกับการใช้กล้องถ่ายภาพในย่านความถี่สายตามนุษย์มองเห็น (Visible Light Camera) โดยระบบหุ่นยนต์จะค้นหาเป้าหมายด้วยการประมวลผลภาพจากกล้องถ่ายภาพ จากนั้นคำนวณตำแหน่งอย่างคร่าวๆของเป้าหมายด้วยการคำนวณความสัมพันธ์ของตำแหน่งกล้องกับมุมรับภาพ เช่นการใช้กล้องคู่ (Stereo Camera) จากนั้นให้อุปกรณ์สำหรับเก็บเกี่ยวเข้าเก็บเกี่ยวเป้าหมาย ซึ่งโดยส่วนมากจะทำผ่านแขนหุ่นยนต์ โดยในขณะที่เก็บเกี่ยวเซนเซอร์วัดระยะแบบอินฟราเรดจะทำหน้าที่วัดระยะห่างหรือตำแหน่งแบบละเอียด ระหว่างอุปกรณ์เก็บเกี่ยวกับผลผลิตเป้าหมายในระนาบสามมิติ ระบบในลักษณะดังกล่าวมีข้อดี โดยผู้ออกแบบระบบสามารถเลือกชนิดของเซนเซอร์ตามความต้องการได้, เพิ่มหรืออัปเดตเซนเซอร์สามารถทำได้ยืดหยุ่น แต่อาจประสบความยุ่งยากมากขึ้นในกรณีที่ใช้เซนเซอร์มาจากต่างผู้ผลิต ซึ่งอาจจะให้ข้อมูลทางเทคนิคที่ไม่เพียงพอ หรือความเข้ากันของเซนเซอร์หากนำมาใช้งานร่วมกัน ยกตัวอย่างระบบมองเห็นที่ใช้หลักการทำงานดังกล่าว เช่น ระบบมองเห็นของหุ่นยนต์เก็บเกี่ยวหน่อไม้ฝรั่ง (Asparagus) [36] ใช้งานกล้องถ่ายภาพสีเพื่อค้นหาหน่อไม้ และใช้งานเลเซอร์สแกนสามมิติ (Laser Scanner) ที่ติดตั้งบนรางเลื่อน (Linear Slider) เพื่อระบุตำแหน่งในสามมิติของเป้าหมายที่หาได้จากกล้องถ่ายภาพสี เพื่อให้แขนหุ่นยนต์สามารถเข้าเก็บเกี่ยวหน่อไม้ฝรั่งที่ตำแหน่งเป้าหมายได้

จนกระทั่งกล้องถ่ายภาพสีและภาพความลึก (RGB-D Camera) ได้รับความนิยมมากขึ้นในช่วงหลัง มีข้อดีที่การใช้งานที่ซับซ้อนลดลง, ผู้พัฒนาระบบสามารถเข้าถึงได้ง่ายขึ้น, และมีการสนับสนุนทางซอฟต์แวร์จากผู้ผลิตกล้อง ตัวอย่างเช่น กล้องคิเน็กซ์จากไมโครซอฟท์ (Microsoft Kinect, Kinect V2) ได้ถูกนำมาใช้งานในงานรู้จำและตรวจหาวัตถุในงานหุ่นยนต์เก็บผลแอปเปิล [40, 67] แต่จุดอ่อนสำคัญคือ ตัวกล้องถูกออกแบบมาให้ใช้งานภายในอาคาร ทำให้กล้องถูกใช้งานได้ไม่เต็มศักยภาพ หรือเกิดปัญหาเมื่อนำไปใช้งานในสภาพแวดล้อมภายนอก งานวิจัยที่ใช้กล้องดังกล่าวสำหรับงานภายนอก (Outdoor) จึงต้องใช้อุปกรณ์เสริมอื่น หรือจำเป็นต้องปรับแต่งอัลกอริทึมตรวจหาเป้าหมาย เพื่อให้สภาพแวดล้อมใกล้เคียงกับการใช้งานกล้องภายในอาคาร หลังจากนั้นก็มีผู้ผลิตรายอื่นๆ พัฒนาต่อยอดกล้องสามมิติให้สามารถใช้งานในสภาพแวดล้อมหลายหลายมากขึ้น รวมไปถึงระบบหุ่นยนต์อัตโนมัติสำหรับการเกษตร ดังใน [47] ที่เลือกใช้กล้องสามมิติจากอินเทล Intel RealSense Camera [41]

ทำให้ในปัจจุบันกล้องถ่ายภาพสีและภาพความลึก เป็นเซนเซอร์ที่นิยมใช้ในหุ่นยนต์ระบบอัตโนมัติ สำหรับการค้นหาและระบุตำแหน่งในสามมิติของวัตถุเป้าหมาย เนื่องจากเซนเซอร์สามารถแปลความข้อมูลได้ทั้งรูปภาพสีในสามช่องสัญญาณสี (RGB Channels), ภาพความลึก (Depth Channel), และจุดในสามมิติ (Point Cloud) ซึ่งการนำข้อมูลแต่ละอย่างมาใช้ประโยชน์นั้น ขึ้นอยู่กับ โจทย์ในด้านสภาพแวดล้อมหรือลักษณะที่ปรากฏของวัตถุเป้าหมาย ยกตัวอย่างเช่น การตรวจหาลำต้นปาล์มน้ำมัน [39], การตรวจหาลำต้นยางพารา [82], การตรวจหาหน่อไม้ฝรั่ง [46] ได้เลือกประมวลผลข้อมูลที่เป็นจุดในสามมิติ (PointCloud) เพียงอย่างเดียวก็เพียงพอต่อการตรวจหาวัตถุเป้าหมายได้ ด้วยอัลกอริทึมการแบ่งกลุ่มจุดในสามมิติ (Clustering) เพราะผู้ออกแบบระบบวางกล้องถ่ายภาพเพื่อทำให้วัตถุเป้าหมายยื่นตัวออกมาจากระนาบของพื้นหลังอย่างชัดเจน แต่ในงานค้นหาผลแอปเปิล [25] และมะเขือเทศ [79] วัตถุเป้าหมายที่เป็นผลไม้จะอยู่ปะปนกับพื้นหลัง

ในระนาบความลึกเดียวกัน การประมวลผลภาพเพื่อแยกบริเวณผลออกจากใบไม้สีเขียวซึ่งเป็นพื้นหลัง ซึ่งการประมวลผลโดยใช้ค่าสีทำได้ง่ายกว่าการใช้จุดในสามมิติ [28, 83] หรือออกแบบอัลกอริทึมที่มีความซับซ้อนมากกว่า เช่น งานตรวจจับผลแอปเปิลในที่มืด [49] โดยให้แสงจากไฟส่องสว่าง ผู้พัฒนาหลีกเลี่ยงการประมวลผลภาพสี เพราะถูกแสงรบกวน หรือสีของวัตถุในภาพมีความไม่แน่นอนในสภาพแสงที่ไม่สม่ำเสมอ แต่หันไปใช้การสะท้อนแสงของผลแอปเปิลแทน เพราะการสะท้อนแสงของผิวที่ผลแอปเปิลกับการสะท้อนแสงของใบไม้ต่างกัน นอกจากนี้ยังมีงานวิจัยอื่นๆ เช่น [25, 76, 79] นำเสนอการทำงานในสภาพแวดล้อมที่มีดี โดยออกแบบอุปกรณ์ให้แสงสว่างกับวัตถุในภาพ งานวิจัยเหล่านี้มีจุดเด่นที่การประมวลผลภาพ สามารถแยกวัตถุที่สนใจออกจากพื้นหลังได้ง่ายกว่าการทำงานในสภาพแสงตอนกลางวัน เพราะไม่มีแสงภายนอกตามธรรมชาติเข้ามารบกวนในภาพถ่าย โดยยังคงใช้งานลักษณะสีของวัตถุร่วมกับการหาขอบภาพของวัตถุเป้าหมาย เพื่อหาตำแหน่งที่แน่นอนสำหรับการเก็บเกี่ยวด้วยแขนหุ่นยนต์

การออกแบบระบบหุ่นยนต์ดังกล่าวจะต้องจัดการกับการจัดวางเซนเซอร์ เพื่อคำนวณความสัมพันธ์ระหว่างตำแหน่งเซนเซอร์กับตำแหน่งวัตถุที่อ่านค่าได้จากเซนเซอร์ โดยหากมีการใช้งานเซนเซอร์จำนวนมาก จะยิ่งเพิ่มความซับซ้อนในการจัดวาง และความซับซ้อนในสร้างอัลกอริทึมการหาตำแหน่งของวัตถุเป้าหมายผ่านการประมวลผลข้อมูลจากเซนเซอร์แต่ละตัว อีกทั้งการจัดวางเซนเซอร์จะต้องทำให้เซนเซอร์นั้นๆ ถูกใช้งานอย่างมีประสิทธิภาพ ซึ่งส่งผลต่อความซับซ้อนของอัลกอริทึมการตรวจหาวัตถุ โดยการออกแบบจะเน้นให้ระบบมองเห็นสามารถประมวลผลเพื่อหาตำแหน่งของวัตถุเป้าหมายได้ง่าย สังเกตได้จากหุ่นยนต์ทางการเกษตรที่ถูกใช้งานโดยทั่วไป จะใช้เทคนิคประมวลผลภาพเพื่อตรวจหาวัตถุเป้าหมายที่ไม่ซับซ้อนแต่ให้ผลที่ดีเพียงพอ โดยใช้ภาพถ่ายในมุมมองที่ทำให้ออกแบบการประมวลผลภาพทำได้ง่ายที่สุด

1.2.7 การประมวลผลภาพเพื่อการตรวจหาวัตถุ

อัลกอริทึมการประมวลผลภาพเพื่อตรวจหาวัตถุ (Image Processing for Object Detection) สามารถจำแนกออกเป็น 2 ประเภท ได้แก่

1. วิธีตรวจหาวัตถุในภาพรูปแบบเดิม (Traditional Object Detector)

ผู้พัฒนาอัลกอริทึมจะค้นหาลักษณะเด่นของวัตถุในภาพ เพื่อเป็นตัวบ่งชี้ความแตกต่างจากพื้นที่อื่นในภาพ ซึ่งเรียกว่า “ฟีเจอร์” (Features) ผ่านวิธีประมวลผลภาพ (Image Processing) และรู้จำวัตถุผ่านลักษณะเด่นดังกล่าวด้วยวิธีการรู้จำแบบต่างๆ (Machine Learning) เพื่อตรวจหาชนิดและตำแหน่งของวัตถุภายในภาพ

โดยทั่วไปผู้พัฒนาอัลกอริทึมจะออกแบบวิธีประมวลผลภาพขั้นต้น (Pre-processing) ก่อนการสกัดหาฟีเจอร์ (Feature Extraction) เพื่อลดข้อมูลที่ไม่ต้องการออกไปจากภาพ หรือการเน้นบางลักษณะในภาพที่ส่งผลต่อการหาฟีเจอร์ให้เด่นชัดขึ้น (Image Enhancement) โดยมีอัลกอริทึมพื้นฐานหรือเทคนิคที่เกี่ยวข้อง เช่น การกรองสัญญาณรบกวน (Noise Filtering), การลบพื้นที่พื้นหลังบางส่วนผ่านวิธีเทรชโฮลด์ (Foreground-Background Thresholding), การเน้นขอบภาพ (Edges Enhancement), การเบลอภาพ (Smoothing), การประสานระหว่างภาพ (Addition or Subtraction), การเลือกค่า (Thresholding), การแก้ไขค่าสีและความสว่าง (Color and Illumination Correction), การปรับเปลี่ยนรูปร่างของพื้นที่ใน

ภาพ (Morphological Filtering), การแปลงโมเดลค่าสี (Color Model Transformation) เป็นต้น ซึ่งต่างเป็นวิธีที่ใช้การคำนวณเชิงตัวเลข นอกเหนือจากนี้ยังมีวิธีประมวลผลภาพอื่นๆ เช่น การแปลงทางเรขาคณิต การหมุนภาพ (Rotation), การย่อหรือขยายภาพ (Scaling or Resizing) รวมถึงการประยุกต์ทฤษฎีกราฟสำหรับประมวลผลรูปภาพ เช่น Connected Component Labeling หรือ Connected Component Analysis [15, 26] เป็นต้น เมื่อผ่านการประมวลผลภาพขั้นต้นภาพจะถูกสกัดหาพีเจอร์ด้วยวิธีที่แตกต่างออกไปตามลักษณะปัญหา ซึ่งผู้ออกแบบอัลกอริทึมตรวจสอบหาจะเป็นผู้ออกแบบวิธีการสกัดหาพีเจอร์เอง เช่น การตรวจหาบุคคลภายในภาพ [13] ด้วยพีเจอร์ฮิสโทแกรมของเกรเดียนแบบหลายทิศทาง (Histogram of Oriented Gradients, HOG) ซึ่งมีพื้นฐานจากการหาขอบภาพ (Edge Detection) เพราะผู้พัฒนาเล็งเห็นว่ารูปร่างของบุคคลในภาพสามารถใช้ขอบภาพและทิศทางของค่าเกรเดียนต์ของขอบรอบโครงร่างคนมีความแตกต่างจากพื้นหลัง โดยการระบุตำแหน่งของบุคคลในภาพ จะค้นหากรอบสี่เหลี่ยมล้อมรอบบุคคล (Bounding Box) จากกรอบที่เป็นไปได้ทั้งหมดในภาพ (Sliding Window) โดยใช้วิธีการตัดสินใจด้วยการเรียนรู้

แต่ในงานเก็บเกี่ยวพืชผลทางการเกษตรนิยมใช้วิธีที่ซับซ้อนน้อยกว่าโดยเฉพาะการใช้สีและรูปร่างของผล ซึ่งมีความกำกวมระหว่าง พีเจอร์สำหรับรู้จำหรือพีเจอร์สำหรับการลดพื้นที่ที่รบกวน เช่น การแบ่งกลุ่มของค่าสี (Color-Clustering) กับการคำนวณความกลมของพื้นที่ (Contour Circularity) เช่น การตรวจหาผลมะเขือเทศ [79, 80] เพราะการใช้สีแดงของผลมะเขือเทศซึ่งมีความต่างอย่างชัดเจนจากใบไม้หลังสีเขียวในส่วนพื้นหลัง และการตรวจหาตำแหน่งของมะเขือเทศจะใช้รูปร่างมะเขือเทศในภาพที่มีลักษณะเป็นวงกลม ทั้งนี้สามารถพิจารณาได้ว่า การเลือกพื้นที่ที่เป็นผลโดยแยกความแตกต่างของสีวัตถุกับสีพื้นหลัง เป็นพีเจอร์หนึ่งสำหรับการรู้จำมะเขือเทศได้ หรือตัวอย่างการการสกัดพีเจอร์ทางความถี่ (Filter Banks) สำหรับแยกชนิดของพื้นผิว (Texture Classification) เช่น [38] เพื่อแยกความแตกต่างระหว่างพื้นผิวของวัตถุกับพื้นผิวของพื้นหลัง สามารถใช้เพื่อรู้จำและระบุตำแหน่งของวัตถุในภาพได้เช่นกัน

สังเกตได้ว่าวิธีตรวจรู้จำวัตถุแบบเดิม รูปแบบพีเจอร์จะขึ้นอยู่กับการวิเคราะห์ปัญหาของผู้พัฒนาอัลกอริทึมตรวจสอบหาวัตถุ ที่จะสามารถค้นหาวิธีดึงลักษณะเด่นของวัตถุที่ปรากฏในภาพ เช่นกันกับปัญหาการตรวจหารอยกรีดและถ้วยรองน้ำยางในภาพ การประมวลผลด้วยวิธีตรวจรู้จำวัตถุแบบเดิมอาจนำมาปรับใช้กับการตรวจหารอยกรีดยางพาราได้ เช่น การหาพีเจอร์ของพื้นผิวเปลือกยางบนหน้ายางที่ยังไม่ถูกกรีดกับหน้ายางที่ผ่านการกรีดมาแล้ว เพื่อหาแนวรอยต่อระหว่างสองพื้นผิวซึ่งคือรอยกรีดล่าสุด หรือ การใช้สีแยกบริเวณรอยต่อของหน้ายางที่ยังไม่ถูกกรีดกับหน้ายางที่ถูกกรีดไปแล้ว จากนั้นหาแนวรอยกรีดล่าสุดโดยการหาขอบภาพ เป็นต้น

2. วิธีตรวจหาวัตถุในภาพรูปแบบสมัยใหม่ (Deep-Learning Object Detector)

เป็นการใช้งานและพัฒนาโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชัน (Convolutional Neural Network) เพื่อรู้จำหรือตรวจหาวัตถุหลากหลายประเภท โดยงานวิจัย [43] นำเสนอระบบโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันสำหรับรู้จำชนิดวัตถุ เป็นโครงข่ายที่มีชั้นประสาทเชื่อมต่อกันแบบหลายชั้นลำดับ (Deep Network) ซึ่งในแต่ละชั้นมีการคำนวณแบบคอนโวลูชันด้วยเคอร์เนล (Kernel) ที่มีค่าภายในต่างกันไป โครงข่ายประสาทจะปรับเปลี่ยนค่าน้ำหนักของเคอร์เนลด้วยการเรียนรู้จากชุดข้อมูลรูปภาพ โดยผู้พัฒนาจะต้องกำหนดรูปภาพอินพุต และกำกับชนิดวัตถุให้กับภาพ การปรับเปลี่ยนค่าน้ำ

หนักตามข้อมูลภาพอินพุตเสมือนการสกัดพีเจอร์ในวิธีตรวจหาวัตถุในภาพรูปแบบเดิม ผลการทดลองจากในหลายงานวิจัยแสดงให้เห็นว่าโครงข่ายประสาทจำลองแบบคอนโวลูชัน มีความสามารถจำแนกประเภทของวัตถุในภาพได้หลายหลายประเภท และมีความแม่นยำมากกว่าวิธีจำแบบเดิม

การตรวจหาวัตถุในภาพด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชัน ซึ่งเป็นปัญหาที่ประกอบด้วย 2 ปัญหาย่อย ได้แก่ การรู้จำชนิดวัตถุ และการระบุตำแหน่งในภาพ ซึ่งโครงข่ายสำหรับรู้จำไม่ได้รองรับการระบุตำแหน่งวัตถุภายในภาพ งานวิจัยการตรวจหาวัตถุในภาพด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันแบบ R-CNN (Region with Convolutional Neural Network) [21] ได้นำเสนอวิธีตรวจหาวัตถุ โดยมีหลักการทำงานที่ผสมระหว่างวิธีประมวลผลภาพแบบเดิมกับการรู้จำของโครงข่าย โดยผู้พัฒนาอัลกอริทึมจะเป็นผู้ออกแบบส่วนของการค้นหาพื้นที่ที่เป็นไปได้ว่ามีวัตถุเป้าหมายอยู่ภายใน (Region Proposal) ในขณะที่โครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชัน จะทำหน้าที่ทำนายพื้นที่ที่เป็นไปได้ใดๆ ว่ามีวัตถุที่ต้องการหรือไม่ อาจกล่าวได้ว่างานวิจัยลักษณะนี้เป็นนิยามใหม่ของคำว่าวิธีตรวจหาวัตถุในภาพแบบสมัยใหม่ (Modern Object Detector หรือ Deep Learning Object Detector) เป็นผลให้เกิดงานวิจัยที่ปรับปรุงการทำงานของ R-CNN ตามมาเป็นจำนวนมาก เพื่อปรับปรุงข้อด้อยในด้านประสิทธิภาพการคำนวณ เช่น Fast R-CNN [20], Faster R-CNN [66] หรือการปรับปรุงในด้านโครงสร้างพร้อมทั้งความแม่นยำ เช่น R-FCN [12], SSD, YOLO, YOLO-9000 เป็นต้น ในขณะเดียวกันการพัฒนาโครงสร้างของโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันเพื่อสกัดพีเจอร์จากข้อมูล ถูกนำเสนอตามมาเป็นจำนวนมากเช่นกัน ยกตัวอย่างเช่น ResNet [30], MobileNet [32, 69] เป็นต้น

ข้อจำกัดที่สำคัญของวิธีตรวจหาวัตถุแบบสมัยใหม่นี้ คือ ข้อมูล เนื่องจากโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันต้องการข้อมูลในการเรียนรู้ ทั้งนี้ด้านปริมาณและความหลากหลาย มากพอที่จะทำให้สามารถพิจารณาวิธีตรวจหาวัตถุอื่นๆ ได้ว่า มีความสามารถหรือความแม่นยำในการอนุมานผลการตรวจหาด้วยชุดข้อมูลอื่นที่ไม่อยู่ในกลุ่มตัวอย่าง (Data Generalization) ทั้งนี้โครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันที่ถูกนำเสนอในงานวิจัยต่างๆ ข้างต้น ได้ถูกสอนและเรียนรู้กับชุดข้อมูลรูปภาพในปริมาณมาก (Big Data)

การนำวิธีการตรวจหาวัตถุรูปแบบสมัยใหม่ไปปรับใช้กับการตรวจหารอยกริดและถ้วยร่อนน้ำยาง สามารถทำได้เช่นกัน โดยใช้เทคนิคการเรียนรู้แบบส่งต่อ (Transfer Learning) ซึ่งสามารถทำให้โครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันที่ผ่านการเรียนรู้บนฐานข้อมูลขนาดใหญ่มาแล้ว เรียนรู้ชุดข้อมูลรูปภาพของรอยกริดต่อได้ (Fine Tuning) เพื่อให้โครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันทำงานที่ในขอบเขตปัญหาที่สนใจได้ คำถามที่ยังเป็นปัญหาสำหรับวิธีนี้คือ กลุ่มข้อมูลรอยกริดยางพาราตัวอย่างที่ใช้เรียนรู้และทดสอบนั้นเพียงพอต่อการอนุมานกับกลุ่มข้อมูลอื่นที่ยังไม่ได้เก็บรวบรวมหรือไม่ เพราะวิธีเรียนรู้ของวิธีตรวจหาวัตถุในรูปแบบสมัยใหม่ มีปัญหาด้านความน่าเชื่อถือและความแม่นยำ เนื่องจากอาจจะให้ค่าความแม่นยำของการตรวจหาวัตถุด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันสูงเกินไป หรือเจาะจงกับข้อมูลชุดนั้น (Overfitting)

งานวิจัยวิธีตรวจหาวัตถุทั้งรูปแบบเดิมและรูปแบบสมัยใหม่ต่างพัฒนาและทดสอบจากฐานข้อมูลรูปภาพที่ถูกเตรียมไว้ให้ผู้วิจัยได้เปรียบเทียบและทดลอง ตัวอย่างฐานข้อมูลเช่น ImageNet [14], MS-COCO Dataset [48], COCO-Stuff [7], Google Open Images Dataset [45], CIFAR-10/CIFAR-100

Datasets [42] แต่สำหรับปัญหาของหุ่นยนต์ทางการเกษตรที่มีเป้าหมายเพียงตรวจวัดเฉพาะอย่าง นั้นยังไม่มีการศึกษาหรือยังไม่มีเก็บข้อมูลรูปภาพที่มากเพียงพอ โดยเฉพาะหากต้องการศึกษาพัฒนาวิธีตรวจหารอยกริดและถ้วยรองน้ำยาง เป็นประเด็นศึกษาหนึ่งที่น่าสนใจ เพราะยังไม่มีเก็บรวบรวมข้อมูลภาพถ่ายรอยกริดบนต้นยางพาราและถ้วยรองน้ำยางในสภาพแวดล้อมจริง ทำให้การศึกษาสามารถทำได้อย่างยืดหยุ่น ตั้งแต่การออกแบบวิธีการรวบรวมข้อมูลภาพถ่ายที่อำนวยความสะดวกผลภาพ การออกแบบวิธีตรวจหาตำแหน่งรอยกริดและถ้วยรองน้ำยางในภาพ ทั้งนี้การกริดยางพาราของเกษตรกรอยู่ในช่วงเวลาที่ไม่มีแสงจากธรรมชาติหรือเป็นช่วงเวลาคาบเกี่ยวระหว่างช่วงที่มีหรือไม่มีแสงอาทิตย์ ดังนั้นในการพัฒนาระบบหากสามารถควบคุมการใช้แสงสว่างในขั้นตอนการรวบรวมภาพรอยกริด จะเป็นประโยชน์ต่อการพัฒนาวิธีตรวจหาแนวรอยกริดในภาพได้ ดังกรณีระบบมองเห็นของหุ่นยนต์เก็บเกี่ยวพืชผลในเวลากลางคืน เช่น ผลแอปเปิล [49, 51], ผลกีวี [52] หรือการพัฒนาวิธีตรวจหาที่หนาทนต่อการเปลี่ยนแปลงของสภาพแสงในสวนยางพารา

1.3 วัตถุประสงค์

- 1 นำเสนอระบบมองเห็นตามแนวคิดระบบหุ่นยนต์ทำสวนยางพาราอัตโนมัติ
- 2 พัฒนาเครื่องมือเก็บข้อมูลภาพถ่ายรอยกริดและถ้วยรองน้ำยาง
- 3 พัฒนาริธีตรวจหารอยกริดและถ้วยรองน้ำยางในภาพถ่าย
- 4 พัฒนาริธีตรวจหาตำแหน่งแนวรอยกริดในพิกัดสามมิติ

1.4 ขอบเขตของการวิจัย

- 1 ศึกษาและเก็บข้อมูลเฉพาะต้นยางพาราที่เปิดกริดเปลือกแรก, ใช้ระบบกริดลงแบบเวียนขวา, และเป็นยางพาราพันธุ์ RRIT-251 และ RRIM-600 ในช่วงเวลาที่มีสภาพแสงน้อย ในเขตพื้นที่ภาคใต้ของไทย
- 2 กำหนดถ้วยรองน้ำยางเป็นชนิดพลาสติกสีดำ ขนาด 0.7 และ 1.0 ลิตร

1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1 เทคนิคการใช้กล้องถ่ายภาพสีและภาพความลึกกับไฟส่องสว่างเพื่อถ่ายภาพรอยกริดบนลำต้นยางพาราในสวนยางพาราที่มีสภาพแสงน้อยหรือมืด
- 2 วิธีการตรวจหารอยกริดและถ้วยรองน้ำยางในภาพถ่าย ด้วยเทคนิคการประมวลผลภาพทั้งในรูปแบบเดิมและวิธีตรวจหาวัตถุในภาพรูปแบบใหม่ เพื่อเปรียบเทียบจุดเด่นหรือจุดด้อย
- 3 วิธีการตรวจหาตำแหน่งแนวรอยกริดอย่างละเอียดในภาพถ่าย และการคำนวณตำแหน่งในสามมิติของแนวรอยกริด เพื่อต่อยอดการพัฒนาขั้นต่อไป

บทที่ 2

วิธีการวิจัย

2.1 ระบบหุ่นยนต์ทำสวนยางพาราอัตโนมัติ

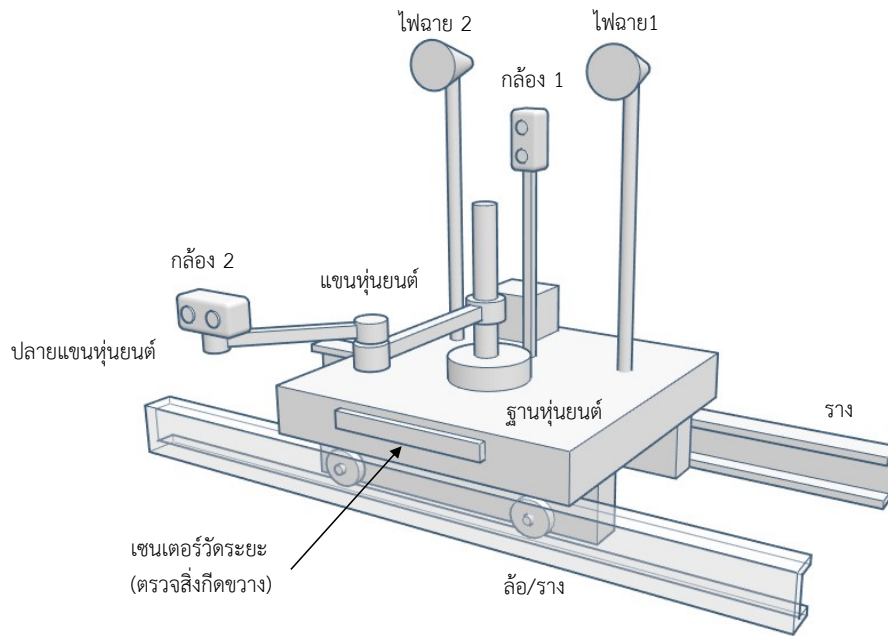
2.1.1 ส่วนประกอบของหุ่นยนต์

หุ่นยนต์ทำสวนยางพาราจะมีลักษณะเดียวกันกับหุ่นยนต์เก็บเกี่ยวพืชผลทางการเกษตรแบบอัตโนมัติ โดยหุ่นยนต์ทำสวนยางพาราจะมีความสามารถในการเคลื่อนที่ในแนวแถวของสวนยาง เพื่อการกรีดยางและเก็บน้ำยางจากต้นยางพาราแต่ละต้นได้ทั่วทั้งสวน ตามรูปที่ 2.1 แสดงส่วนประกอบหลักของหุ่นยนต์ ดังนี้

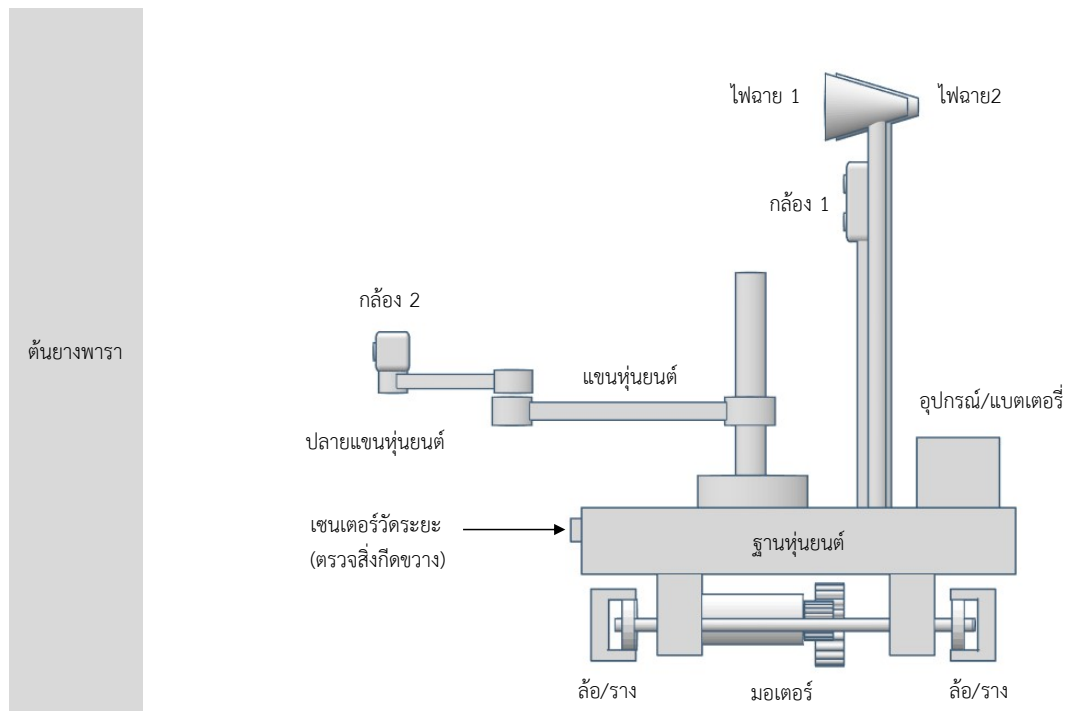
- 1 ฐานหุ่นยนต์ขับเคลื่อนด้วยมอเตอร์ไฟฟ้าและระบบล้อและราง สำหรับเคลื่อนที่ตามแนวแถวต้นยางพารา ซึ่งตามภาพรางจะถูกติดตั้งไว้บนพื้นประจำแถวต้นยางพารา
- 2 แขนหุ่นยนต์ซึ่งถูกติดตั้งอยู่บนฐาน สามารถเคลื่อนเข้าหาตำแหน่งรอยกรีดบนหน้ายางและถ้วยรองน้ำยางบนลำต้นได้อย่างครอบคลุม เพื่อกรีดหรือเก็บน้ำยางจากถ้วยรองน้ำยาง
- 3 ระบบมองเห็นทำหน้าที่ค้นหาตำแหน่งเป้าหมายในพื้นที่ทำงานของหุ่นยนต์ โดยอาศัยข้อมูลด้วยกล้องถ่ายภาพหรือเซนเซอร์วัดระยะรูปแบบอื่นๆ เพื่อตรวจหาตำแหน่งรอยกรีดและถ้วยรองน้ำยางอ้างอิงกับตำแหน่งหุ่นยนต์ เพื่อให้แขนหุ่นยนต์เข้ากรีดหรือเก็บน้ำยางจากถ้วยรองน้ำยางในบริเวณลำต้นยางพาราได้

2.1.2 ระบบมองเห็นของหุ่นยนต์

ระบบมองเห็นทำหน้าที่ตรวจหาตำแหน่งรอยกรีดและถ้วยรองน้ำยาง เพื่อให้แขนหุ่นยนต์เข้ากรีดหรือเก็บน้ำยางจากถ้วยรองน้ำยางได้ ระบบมองเห็นของหุ่นยนต์ที่ใช้กล้องถ่ายภาพเป็นเซนเซอร์รับข้อมูลจึงต้องการกล้องที่สามารถวัดตำแหน่งในภาพเป็นพิกัดสามมิติได้ และเนื่องจากการทำงานของหุ่นยนต์ต้องทำงานในเวลากลางคืน ระบบหุ่นยนต์จึงต้องการไฟส่องสว่างไปยังหน้ายางเพื่อช่วยในการทำงานของกล้องถ่ายภาพ ตามรูปที่ 2.1 และรูปที่ 2.2



รูปที่ 2.1: ภาพวาดแนวคิดหุ่นยนต์ทำสวนยางพาราอัตโนมัติ



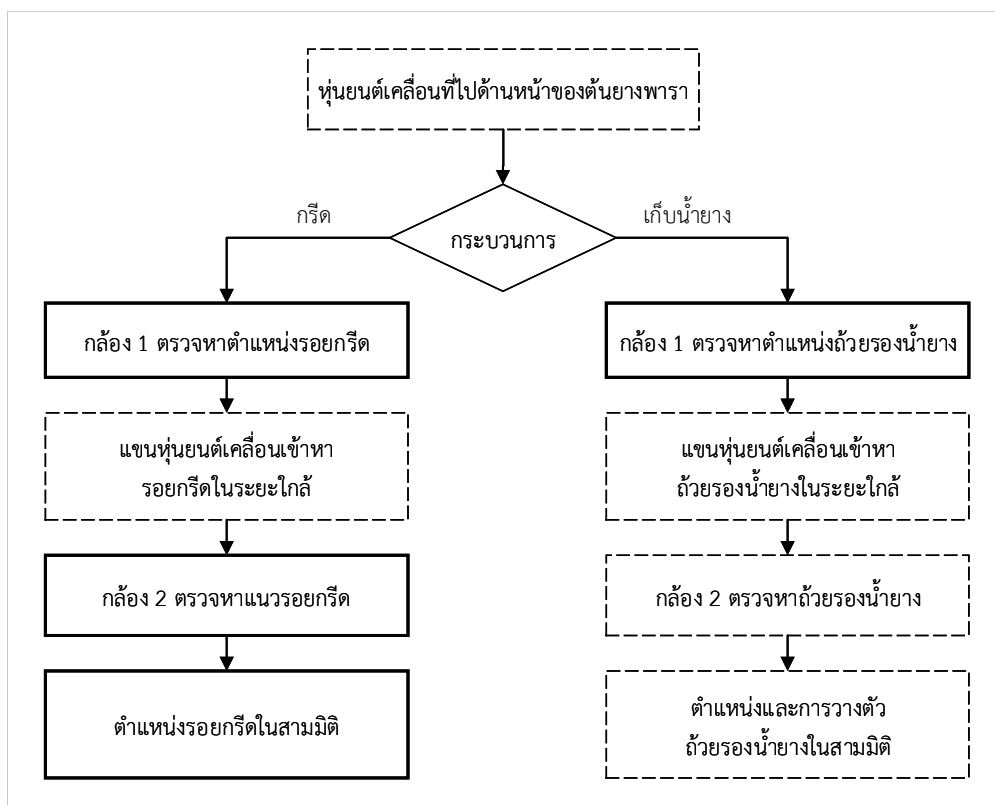
รูปที่ 2.2: ภาพวาดแสดงหุ่นยนต์เมื่อเข้าประจำต้นยางพารา

2.2 การออกแบบระบบมองเห็นสำหรับหุ่นยนต์ทำสวนยางพาราอัตโนมัติ

2.2.1 ขั้นตอนการทำงาน

ออกแบบให้ระบบมองเห็นใช้กล้องถ่ายภาพสีและภาพความลึก (RGB-D Camera) ซึ่งสามารถถ่ายภาพพร้อมทั้งให้ค่าพิกัดสามมิติของวัตถุที่ปรากฏในภาพได้ ตามรูปที่ 2.1 และรูปที่ 2.2 ประกอบไปด้วยกล้อง 2 ตำแหน่ง ได้แก่ กล้องถ่ายภาพมุมกว้าง (กล้อง 1) ติดตั้งบริเวณฐานหุ่นยนต์ ทำหน้าที่ตรวจหาตำแหน่งรอยกรีดหรือถ้วยรองน้ำยางในมุมมองภาพรวม และกล้องระยะใกล้ (กล้อง 2) ติดตั้งบริเวณปลายแขนหุ่นยนต์ ทำหน้าที่ตรวจหาตำแหน่งของเป้าหมายที่ได้จากกล้องถ่ายภาพมุมกว้างอย่างละเอียด ซึ่งได้แก่ แนวนรอยกรีดและตำแหน่งหน้าสัมผัสของมือจับ สำหรับเข้าจับถ้วยรองน้ำยางของมือจับในส่วนปลายแขนหุ่นยนต์

การทำงานหุ่นยนต์และระบบมองเห็นจะแบ่งออกเป็น 2 ขั้นตอน เริ่มต้นเมื่อหุ่นยนต์เคลื่อนตรงกลับลำต้นยางพารา โดยอาศัยเซนเซอร์ตรวจหาสิ่งกีดขวางในการขยับฐานหุ่นยนต์ให้ตรงกับลำต้น จากนั้นระบบมองเห็นจะใช้กล้องถ่ายภาพมุมกว้าง (กล้อง 1) ตรวจหาตำแหน่งสามมิติของรอยกรีด หรือถ้วยรองน้ำยางในมุมมองภาพรวมที่ครอบคลุมพื้นที่ทั้งหมดที่เป็นไปได้บนลำต้น จากนั้นแขนหุ่นยนต์ยื่นกล้องระยะใกล้ (กล้อง 2) ที่ติดตั้งอยู่ส่วนปลายแขนหุ่นยนต์ไปยังตำแหน่งที่ตรวจหาได้จากกล้องถ่ายภาพมุมกว้าง เพื่อให้กล้องระยะใกล้ตรวจหาตำแหน่งแนวนรอยกรีดหรือถ้วยรองน้ำยางอย่างละเอียดอีกครั้งหนึ่ง ตามแผนผังการทำงานรูปที่ 2.3



รูปที่ 2.3: แผนผังแสดงขั้นตอนการทำงานของระบบมองเห็นของหุ่นยนต์ทำสวนยางพาราอัตโนมัติ

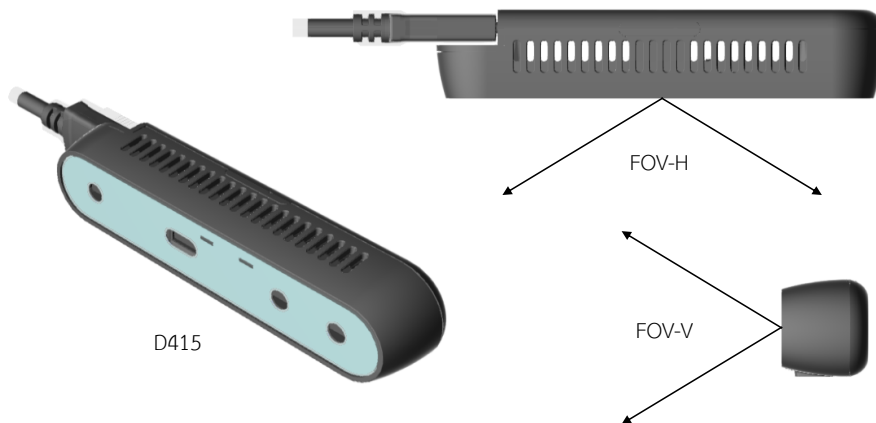
2.2.2 กล้องถ่ายภาพ

ระบบมองเห็นเลือกใช้กล้องถ่ายภาพสีและภาพความลึกรุ่นอินเทลเรียลเซนส์ D400s (Intel RealSense D400 Series) [41] ซึ่งเป็นกล้องที่มีทั้งโมดูลกล้องถ่ายภาพสีและโมดูลกล้องถ่ายภาพความลึกภายใน โดยกล้องถ่ายภาพความลึกจะทำงานแบบเอคทีฟสเตอริโอ (Active IR Stereo) ร่วมกับโปรเจคเตอร์ฉายแสงอินฟราเรดที่ติดตั้งอยู่ภายในตัวกล้อง

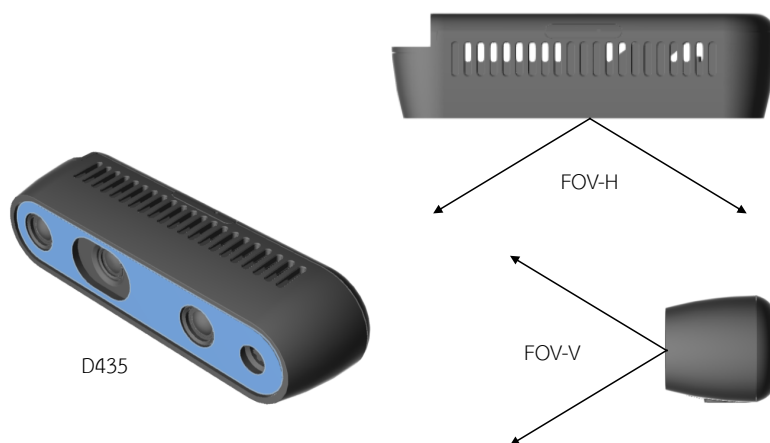
กล้องจากอินเทลมีจุดเด่นที่เอื้อต่อการนำไปใช้เก็บข้อมูลภาพถ่ายในสวนยางพารามากกว่ากล้องรุ่นอื่นๆ ที่มีจำหน่ายสำหรับผู้พัฒนา เช่นกล้องคิเน็กซ์ (Microsoft Kinect) ในด้านศักยภาพที่สามารถถ่ายภาพที่มีความละเอียดสูง ในขณะที่มีความคลาดเคลื่อนต่ำเหมาะสำหรับการตรวจหาแนวรอยกรีด, กล้องมีขนาดเล็กเหมาะต่อการใช้งานเป็นกล้องปลายแขนหุ่นยนต์, ระบบไฟเลี้ยงผ่านสาย USB เหมาะต่อการนำไปใช้งานหรือเก็บข้อมูลภาพถ่ายนอกสถานที่ รวมถึงการสนับสนุนด้านซอฟต์แวร์สำหรับนักพัฒนาจากผู้ผลิต โดยในตารางที่ 2.1 แสดงคุณสมบัติของโมดูลกล้องถ่ายภาพสีและภาพความลึกของแต่ละโมเดล ที่นำมาใช้พัฒนาระบบมองเห็นของหุ่นยนต์ทำสวนยางพาราในวิทยานิพนธ์ฉบับนี้

ตารางที่ 2.1: ตารางแสดงคุณสมบัติของโมดูลกล้องถ่ายภาพสีและภาพความลึก ภายในกล้องอินเทลเรียลเซนส์ โมเดล D415 และ D435 [10]

| คุณสมบัติ | โมเดล D415 (รูปที่ 2.4) | โมเดล D435 (รูปที่ 2.5) |
|--|-------------------------|-------------------------|
| มุมรับภาพกล้องถ่ายภาพสี (องศา) | | |
| ตามแนวนอนภาพ, แนวตั้งภาพ, แนวเส้นทแยงมุมภาพ (FOV-H°, FOV-V°, FOV-D°) | 69.4, 42.5, 77 | 69.4, 42.5, 77 |
| มุมรับภาพกล้องถ่ายภาพความลึก (องศา) | | |
| ตามแนวนอนภาพ, แนวตั้งภาพ, แนวเส้นทแยงมุมภาพ (FOV-H°, FOV-V°, FOV-D°) | 69.4, 42.5, 77 | 91.2, 65.5, 100.6 |
| ระยะความลึกที่ทำงานได้น้อยที่สุด (มิลลิเมตร) | | |
| ที่ภาพความละเอียด 1280x720 พิกเซล โดยไม่ปรับการตั้งค่าเพิ่มเติม | 450 | 280 |



รูปที่ 2.4: ภาพโมเดลกล้อง D415 และมุมรับภาพตามแนวแกนนอนตามปกติ (FOV-H) และมุมรับภาพตามแนวแกนตั้งตามปกติ (FOV-V)

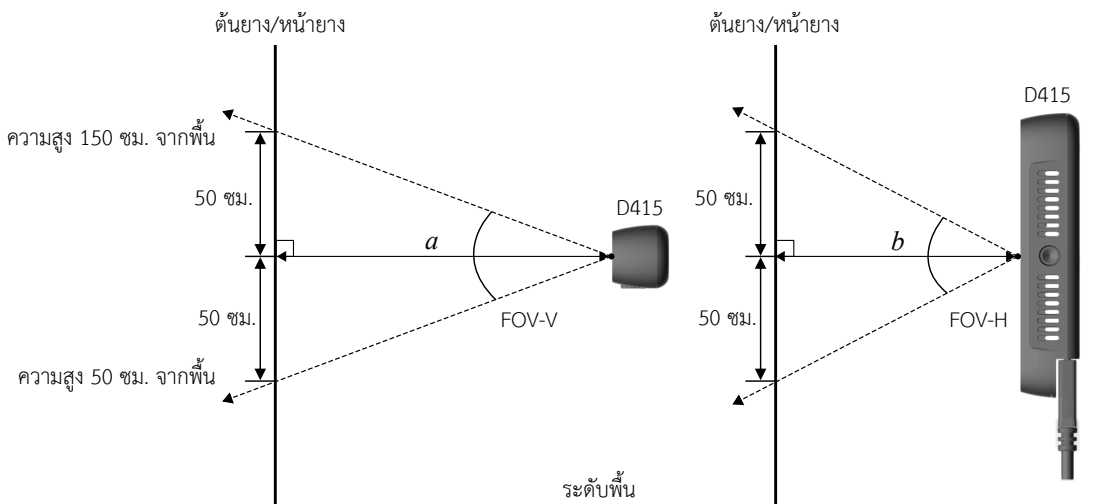


รูปที่ 2.5: ภาพโมเดลกล้อง D435 และมุมรับภาพตามแนวแกนนอนตามปกติ (FOV-H) และมุมรับภาพตามแนวแกนตั้งตามปกติ (FOV-V)

2.2.3 กล้องถ่ายภาพมุมกว้าง (กล้อง 1)

กำหนดให้กล้องมุมกว้าง (กล้อง 1) มีตำแหน่งและมุมรับภาพคงที่บนฐานหุ่นยนต์ ประโยชน์เพื่อความสะดวกต่อการคำนวณตำแหน่งรอยกรีดและถ้วยรองน้ำยาง ระหว่างพิกัดแกนหุ่นยนต์หรือฐานหุ่นยนต์กับพิกัดกล้อง การวางกล้องต้องทำให้มุมรับภาพครอบคลุมความสูงของหน้ายางและถ้วยรองน้ำยางบนลำต้น ซึ่งคือระดับความสูงที่ 50 ถึง 150 เซนติเมตรจากพื้น และต้องให้กล้องอยู่ใกล้ลำต้นยางพาราเพื่อลดระยะยื่นของแกนหุ่นยนต์

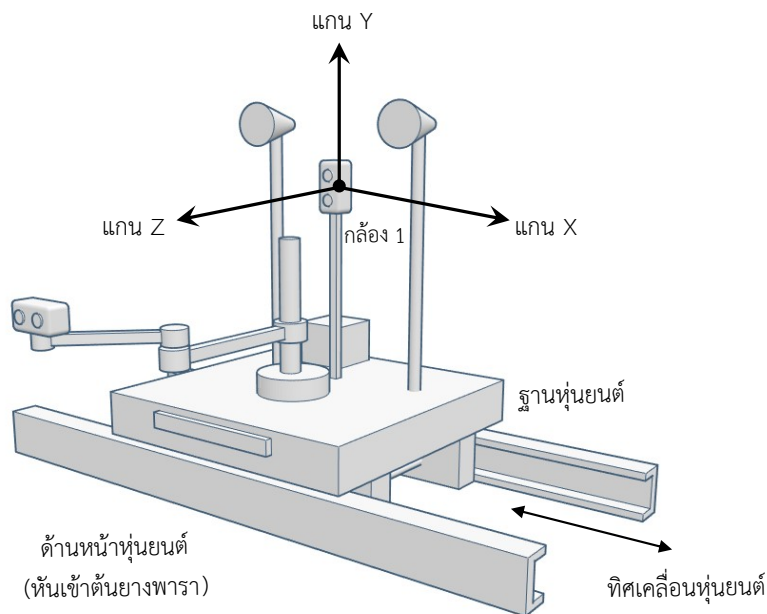
จากตารางที่ 2.1 มุมรับภาพของโมดูลกล้องถ่ายภาพสีทั้งสองโมเดลมีค่าเท่ากัน จึงเลือกให้กล้องโมเดล D415 ที่มีระยะความลึกที่น้อยที่สุดมากกว่าโมเดล D435 ทำหน้าที่เป็นกล้องถ่ายภาพมุมกว้าง โมดูลกล้องถ่ายภาพสีมีมุมรับภาพในแนวนอนตามปกติของกล้อง (FOV-H) กว้างกว่ามุมรับภาพในแนวตั้งตามปกติของกล้อง (FOV-V) จึงเลือกวางกล้องในแนวตั้ง ตามรูปที่ 2.6 เพื่อให้มุมรับภาพครอบคลุมหน้ายาง โดยที่สามารถวางกล้องเข้าใกล้หน้ายางได้มากกว่าการวางกล้องตามปกติ



รูปที่ 2.6: การวางกล้องถ่ายภาพมุมกว้าง (กล้อง 1) ในกรณีวางกล้องตามปกติ (ภาพซ้าย) ซึ่งมุมรับภาพตามแนวแกนตั้งปกติ (FOV-V) มีค่าน้อยกว่ามุมรับภาพตามแกนนอน (FOV-H) ทำให้ต้องวางกล้องห่างจากหน้ายางมากกว่า ($a > b$) เพื่อให้มุมรับภาพครอบคลุมพื้นที่หน้ายาง

กำหนดกรอบอ้างอิงพิกัดฉากสามมิติ (Cartesian Coordinate) ของกล้องถ่ายภาพมุมกว้างตามรูปที่ 2.7 มีจุดกำหนดอยู่ที่ตำแหน่งของกล้องถ่ายภาพมุมกว้าง แกน Z แทนแกนความลึกของกล้อง ซึ่งเป็นแนวเส้นสมมุติที่ลากผ่านศูนย์กลางเลนส์ มีทิศพุ่งออกจากหน้ากล้องเข้าหาต้นยางพาราขนานกับแนวระดับพื้นดิน, แกน Y มีทิศชี้ขึ้นตามแนวตั้ง, และแกน X มีทิศขนานกับแนวเคลื่อนของราง ทั้งนี้จะใช้กรอบอ้างอิงนี้เป็นกรอบอ้างอิงหลักของหุ่นยนต์

ระยะห่างระหว่างกล้องถ่ายภาพมุมกว้างกับหน้ายาง (z_0) ขึ้นอยู่กับมุมรับภาพของโมดูลกล้องแต่ละชนิดตามตาราง 2.1 จึงกำหนดให้ความสูงของกล้องมีค่าประมาณ 100 เซนติเมตรจากพื้น และให้แนวเส้นผ่านศูนย์กลางเลนส์ (Optical Centerline) หรือแกนความลึกตั้งฉากกับผิวลำต้น เพื่อมุมรับภาพครอบคลุมพื้นที่หน้ายางสำหรับกรีด โดยระยะความสูง (ตามแกน Y) ของครึ่งหน้ายางหรือ 50 เซนติเมตร จะ



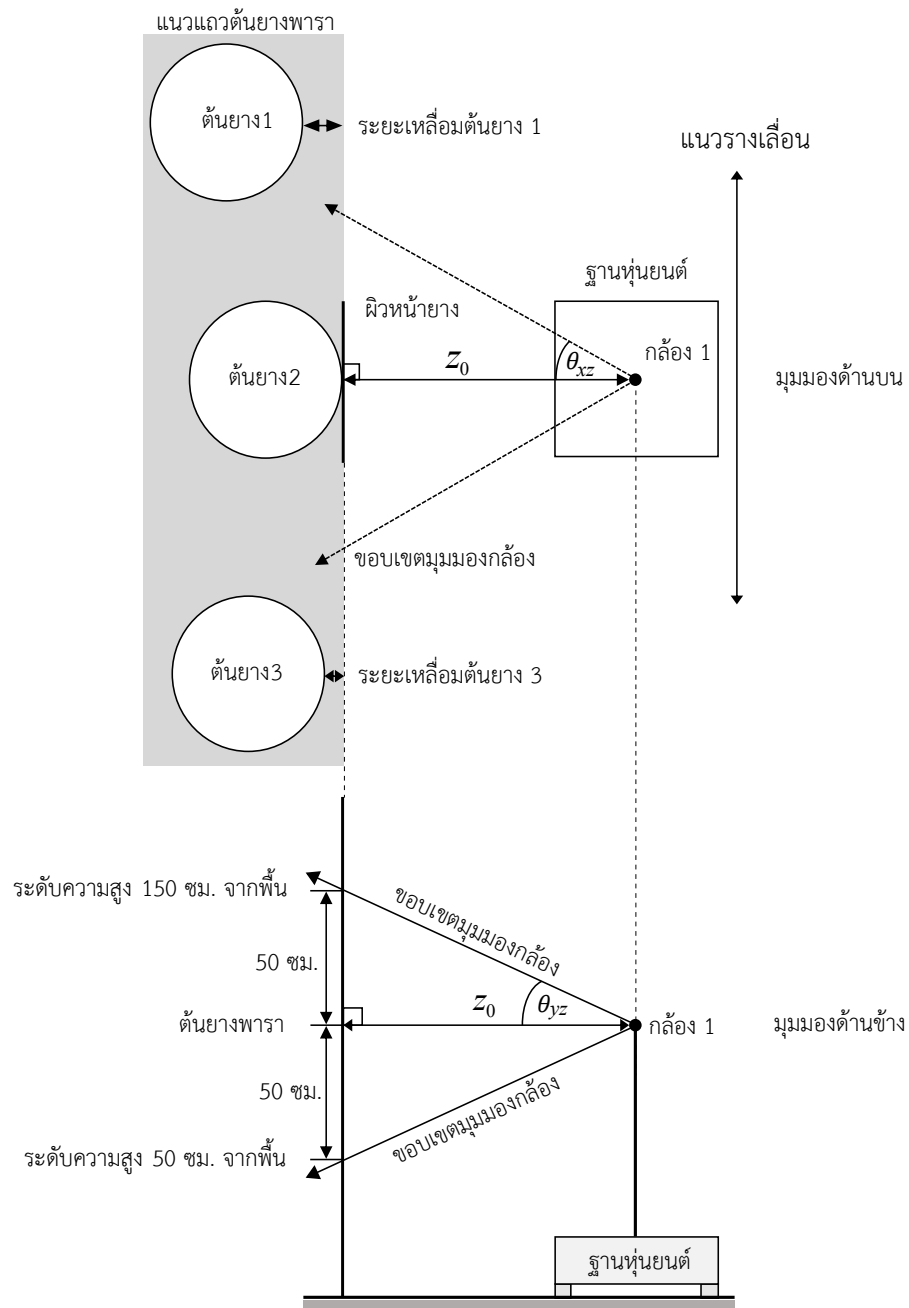
รูปที่ 2.7: กรอบอ้างอิงกล้องถ่ายภาพมุมกว้าง (กล้อง 1)

ถูกรวมด้วยมุมรับภาพขนาด θ_{yz} ซึ่งเท่ากับครึ่งหนึ่งของมุมรับภาพรวม ($\frac{1}{2}$ FOV-H) ดังแสดงในรูปที่ 2.8

$$\min(z_0) = 50 \tan \theta_{yz} \quad (2.1)$$

กำหนดให้ θ_{yz} เท่ากับครึ่งหนึ่งของมุมรับภาพในระนาบ YZ ระยะห่างที่น้อยที่สุด $\min(z_0)$ ในหน่วยเซนติเมตร สำหรับโมดูลกล้องแต่ละชนิดของกล้องโมเดล D415 และ D435 สามารถคำนวณได้จากสมการ 2.1 จากการคำนวณพบว่าระยะห่างน้อยที่สุดระหว่างกล้องถ่ายภาพกับหน้ายางของกล้องโมเดล D415 เท่ากับ 81.60 เซนติเมตร, กล้องโมเดล D435 เท่ากับ 76.40 เซนติเมตร เพื่อให้โมดูลเซนเซอร์รับภาพทั้งภายในกล้องทั้งหมดมีมุมรับภาพครอบคลุมหน้ายางที่ความสูงที่เป็นไปได้ของรอยกรีด

อย่างไรก็ตามกล้องถ่ายภาพมุมกว้างจะถูกติดตั้งบนฐานหุ่นยนต์ที่อยู่บนรางขนานกับแนวแถวของต้นยางพารา ซึ่งต้นยางพาราแต่ละต้นในแถวจะมีระยะเหลือมที่ทำให้ระยะห่างระหว่างรางกับลำต้นไม้เท่ากัน ส่งผลให้ระยะระหว่างกล้องกับต้นยางพาราในแต่ละต้นไม่เท่ากันเช่นกัน ตามรูปที่ 2.8 ดังนั้นจึงต้องเผื่อระยะวางกล้องให้ห่างจากหน้ายางให้มากขึ้น โดยระยะห่างระหว่างกล้องถ่ายภาพมุมกว้างกับต้นยางพาราจะเท่ากับ $\min(z_0)$ รวมกับระยะห่างจากต้นยางพาราถึงรางที่มากที่สุดแถว แต่จากขอบเขตของวิทยานิพนธ์ฉบับนี้จะเป็นการพัฒนาเพียงระบบมองเห็นเท่านั้น (เฉพาะกรอบเส้นประในรูปที่ 2.3) จึงกำหนดค่าอย่างง่ายที่เป็นไปได้ โดยให้ระยะห่างของกล้องกับต้นยางพาราแต่ละต้นมีค่าในช่วง 90 ถึง 120 เซนติเมตร ซึ่งเป็นช่วงที่เหมาะสมที่จะใช้กล้องโมเดล D415 เนื่องด้วยศักยภาพของกล้องทั้งสองโมเดลไม่แตกต่างกันสำหรับใช้เป็นกล้องถ่ายภาพมุมกว้าง

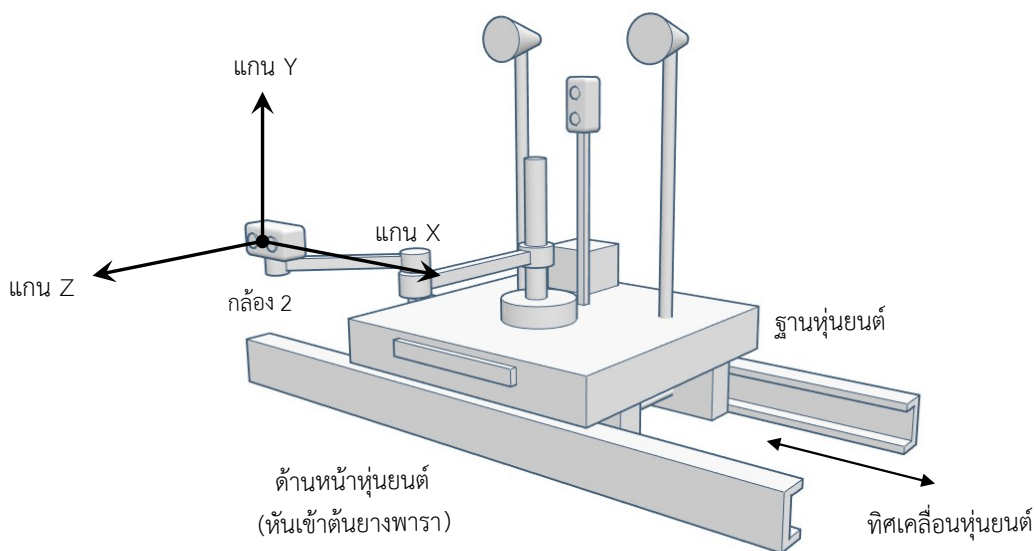


รูปที่ 2.8: แผนภาพแสดงมุมมองด้านบน (ระนาบ YZ) และมุมมองด้านข้าง (ระนาบ XZ) ซึ่งกล้องถ่ายภาพมุมกว้าง (กล้อง 1) อยู่ห่างจากหน้ายางเป็นระยะตั้งฉากเท่ากับ z_0 โดยหน้ายางของต้นยางแต่ละต้นไม่ได้เรียงกันเป็นแนวเส้นตรงโดยสมบูรณ์ (มีระยะเหลือมดังแสดงในมุมมองด้านบน)

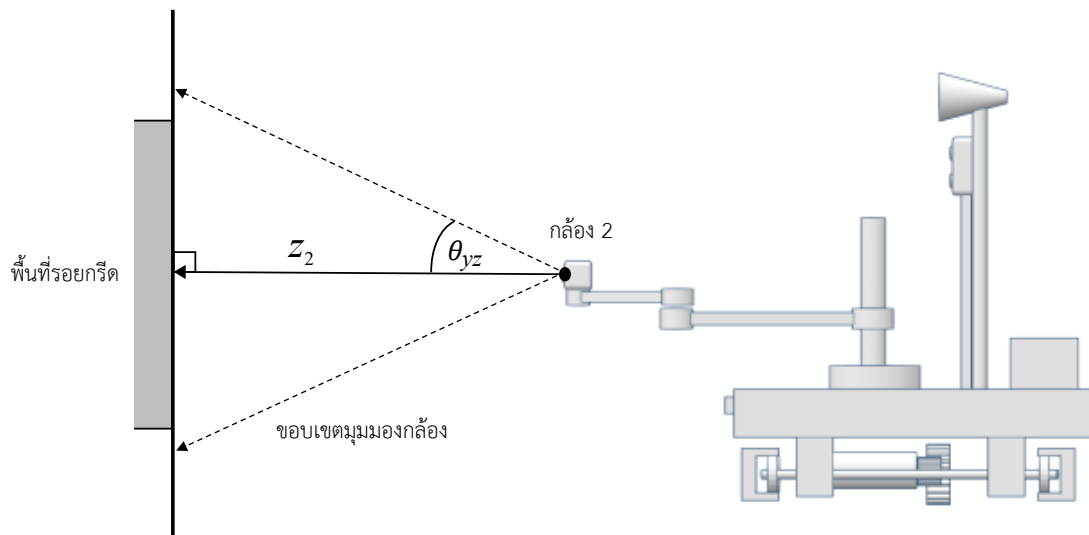
2.2.4 กล้องระยะใกล้ (กล้อง 2)

การเคลื่อนแขนหุ่นยนต์เข้าหาหน้ายางในระยะใกล้เพื่อกรีดหรือเก็บน้ำยาง อาจทำให้แขนหุ่นยนต์จะบังมุมรับภาพของกล้องถ่ายภาพมุมกว้าง หรือเกิดเงาของแขนหุ่นยนต์ในภาพจากกล้องถ่ายภาพมุมกว้าง เนื่องจากไฟส่องสว่างที่ตำแหน่งฐานอยู่ในระนาบความลึกเดียวกันกับกล้องถ่ายภาพมุมกว้าง นอกจากนี้กล้องถ่ายภาพมุมกว้าง (กล้อง 1) ซึ่งคือกล้องโมเดล D415 มีความคลาดเคลื่อนของความลึกของกล้องที่ระยะห่างจากหน้ายาง 90-120 เซนติเมตร [27] มากกว่าความลึกระหว่างเปลือกยางโดยเฉลี่ยตามแนวรอยกรีด ดังนั้นกล้องถ่ายภาพมุมกว้างมีความสามารถไม่เพียงพอสำหรับหุ่นยนต์ทำสวนยางพาราอัตโนมัติ จึงออกแบบให้ระบบมองเห็นใช้งานกล้องระยะใกล้สำหรับติดตั้งบริเวณปลายแขนกล จึงเลือกใช้กล้องโมเดล D435 เป็นกล้องระยะใกล้ โดยกำหนดให้ระยะทำงานของกล้องมีความลึกห่างจากผิวหน้ายางในช่วง 20.0-40.0 เซนติเมตร (z_2) เป็นช่วงที่เหมาะสมเนื่องจากคุณสมบัติกล้องโมเดล D435 มีระยะความลึกน้อยที่สุดที่กล้องสามารถทำงานได้ประมาณ 15 เซนติเมตร และมีค่าความคลาดเคลื่อนของความลึกค่อนข้างต่ำในระยะไม่เกิน 50 เซนติเมตร

กำหนดกรอบอ้างอิงระบบพิกัดฉากสามมิติของกล้องระยะใกล้ตามรูปที่ 2.9 ให้จุดกำเนิดอยู่ที่ตำแหน่งกล้องระยะใกล้, แกน Z แทนแกนความลึกของกล้อง ซึ่งเป็นแนวเส้นสมมติผ่านศูนย์กลางเลนส์ให้ชี้ไปยังกลางลำต้นหรือตั้งฉากกับผิวลำต้น ดังแสดงในรูปที่ 2.10, แกน Y มีทิศทางชี้ขึ้นจากระดับพื้น, และแกน X เป็นไปตามกฎมือขวา



รูปที่ 2.9: กรอบอ้างอิงกล้องระยะใกล้ (กล้อง 2)



รูปที่ 2.10: มุมมองด้านข้าง (ระนาบ YZ ตามกรอบอ้างอิงกล้องระยะใกล้) แสดงตำแหน่งมุมรับภาพของกล้องระยะใกล้ (กล้อง 2) กับพื้นที่รอยกรีด

2.3 การพัฒนาวิธีตรวจหาตำแหน่งรอยกรีดและถ้ำร่องน้ำยางของกล้องถ่ายภาพมุมกว้าง (กล้อง 1)

การพัฒนาวิธีตรวจหาตำแหน่งรอยกรีดและถ้ำร่องน้ำยางในภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง (กล้อง 1) ประกอบด้วยขั้นตอน การถ่ายภาพต้นยางพาราด้วยกล้องถ่ายภาพมุมกว้าง, การกำหนดตำแหน่งรอยกรีดและถ้ำร่องน้ำยางภายในภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง, วิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมและความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี, วิธีตรวจหารอยกรีดและถ้ำร่องน้ำยางในภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง ด้วยวิธีตรวจหาวัตถุชนิดโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันแบบ Faster-RCNN (Faster Region with Deep Convolutional Neural Network), และการคำนวณตำแหน่งเป้าหมายในภาพเพื่อการใช้งานกล้องระยะใกล้ (กล้อง 2)

2.3.1 การถ่ายภาพต้นยางพาราด้วยกล้องถ่ายภาพมุมกว้าง

2.3.1.1 อุปกรณ์ถ่ายภาพ

ชุดอุปกรณ์ถ่ายภาพประกอบไปด้วย กล้องถ่ายภาพสีและภาพความลึกโมเดล D415, ไฟส่องสว่าง โดยทั้งหมดถูกจัดวางบนขาตั้งกล้องเพื่อสะดวกต่อการเคลื่อนย้ายขณะถ่ายภาพในสวนยางพาราตามรูปที่ 2.11



รูปที่ 2.11: อุปกรณ์ถ่ายภาพกล้องถ่ายภาพมุมกว้าง

กล้องถ่ายภาพ กล้องอินเทลเรียลเซนส์โมเดล D415 จะถ่ายทั้งภาพสีและภาพความลึก โดยมีการตั้งค่าของโมดูลกล้องภายในดังนี้

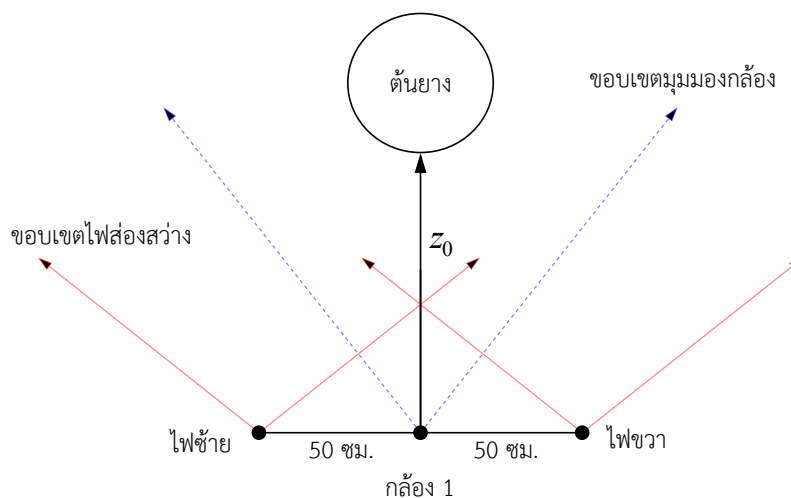
- **โมดูลกล้องถ่ายภาพสี** เลือกใช้งานการตั้งค่าเดิมจากผู้ผลิตกล้อง (Default Preset) ยกเว้นค่าสมดุลสี (White Balance) โดยให้ค่าสมดุลสีตรงกับอุณหภูมิสีของไฟส่องสว่าง 5500K และการวัดแสงของกล้องให้เลือกวัดแสง บริเวณพื้นที่ของลำต้นยางพาราหรือส่วนกลางภาพ เพื่อให้พื้นที่ว่างข้างลำต้นที่ปรากฏในภาพมีค่าน้อยกว่าบริเวณลำต้น
- **โมดูลกล้องถ่ายภาพความลึก** เปิดใช้งานอินฟราเรดโปรเจกเตอร์ภายในตัวกล้องเพื่อให้กล้องทำงานในโหมดแอคทีฟสเตอริโอ (Active IR Stereo) พร้อมเปิดการใช้งานฟิลเตอร์ประมวลผลภาพเพื่อลดสัญญาณรบกวน (Post Processing Smoothing Filters [27]) โดยอาศัยซอฟต์แวร์จากผู้ผลิตกล้อง¹ ซึ่งได้แก่ ฟิลเตอร์กรองสัญญาณรบกวนโดยเฉลี่ยความถี่ระหว่างเฟรมภาพ (Temporal Filter) และฟิลเตอร์เบลอภาพ (Spatial Smoothing Filter)

ไฟส่องสว่าง การใช้ไฟแบบโคมหรือแหล่งกำเนิดแสงแบบจุด (Blub) กับลำต้นยางพาราที่มีลักษณะเหมือนทรงกระบอกวางตรงในแนวตั้ง อาจทำให้มีบางพื้นที่บนลำต้นได้รับแสงมากเกินไปจนเกิดการคลิบของสัญญาณภาพ (Blown-out Highlights) หรือมีบางพื้นที่ที่ได้รับแสงน้อยเกินไปจนเป็นบริเวณมืดในภาพ เนื่องจากลักษณะการฉายแสงของแหล่งกำเนิดแสงแบบจุด บริเวณแนวกลางลำแสงจะมีปริมาณแสงมากกว่าบริเวณรอบข้างซึ่งไม่เหมาะกับการให้แสงกับลำต้น เพื่อลดปัญหาการกระจายแสงที่ไม่ทั่วถึงตามความสูงของลำต้น จึงเหมาะที่จะใช้ไฟแบบรางเส้นตรง เป็นไฟส่องสว่างสำหรับการถ่ายภาพด้วยกล้องถ่ายภาพมุมกว้าง

การจัดวางไฟส่องสว่างแบบรางเส้นตรงสองตำแหน่งที่มีความสว่างเท่ากันวางขนานทางข้างซ้ายและขวาของกล้อง ห่างจากกล้องข้างละประมาณ 50 เซนติเมตร ซึ่งไฟทั้งสองจะอยู่ในระนาบความลึกเดียวกันกับกล้อง เพื่อให้ขอบเขตมุมให้แสงส่องถึงต้นยางพาราในบริเวณส่วนโค้งด้านข้างของลำต้น, หลีกเลี่ยงการเกิดเงาของกล้องบนหน้ายาง, และเพื่อไม่ให้ลำแสงรบกวนหน้ากล้อง ดังแสดงในรูปที่ 2.12

ไฟส่องสว่างแบบราง (LED Strip, Light Emitting Diodes) ขนาดยาว 1 เมตร แสงมีอุณหภูมิสีช่วง 5500K ซึ่งมีสีใกล้เคียงแสงสว่างจากดวงอาทิตย์ในเวลากลางวัน ไฟทั้งสองถูกต่อเข้ากับวงจรหรี่ไฟเพื่อควบคุมความสว่างที่ไม่ใช่หลักการหรี่แบบกระพริบ (Non-PWM Dimmer) เพื่อหลีกเลี่ยงปัญหาแถบกระพริบสีดำในภาพอันเนื่องมาจากความถี่ในการรีเฟรชภาพของกล้องดิจิทัล กับการกระพริบของไฟส่องสว่างไม่สอดคล้องกัน (Flickering)

¹Intel® RealSense™ SDK 2.0, <https://github.com/IntelRealSense/librealsense>



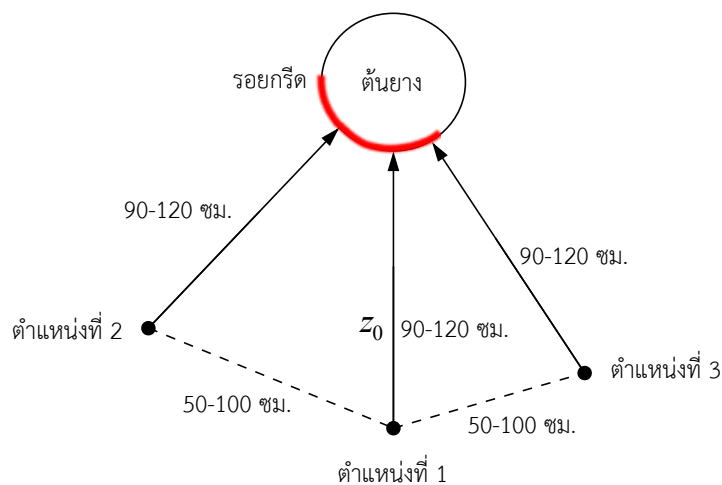
รูปที่ 2.12: (มุมมองจากด้านบน) แผนภาพแสดงตำแหน่งของอุปกรณ์ถ่ายภาพเมื่อวางหน้าต้นยางพารา แสดงตำแหน่งไฟส่องสว่างทั้งสอง (ไฟซ้าย, ไฟขวา) กับตำแหน่งกล้องถ่ายภาพมุมกว้าง (กล้อง 1)

2.3.1.2 ขั้นตอนการถ่ายภาพ

การวางกล้องถ่ายภาพมุมกว้างเป็นไปตามข้อ 2.2.3 โดยให้แกนความลึก (แกน Z) ชี้เข้าหาลำต้น ระยะห่างหน้ายางกับตำแหน่งกล้อง (z_0) อยู่ในช่วง 90-120 เซนติเมตรเพื่อรองรับความเหลี่ยมของต้นยางในแต่แนวแถว, กล้องวางสูงเหนือพื้นประมาณ 100 เซนติเมตร, และหันกล้องให้แกนความลึก หรือระนาบ XZ ขนานกับระนาบพื้นดิน แสดงในรูปที่ 2.8

การวางกล้อง ณ ตำแหน่งหนึ่งๆ จะถ่ายทั้งภาพสีและภาพความลึก เรียกภาพถ่ายต้นยางพาราในหนึ่งตำแหน่งนี้ว่า 1 ชุดภาพ การถ่ายภาพ 1 ชุดภาพ ทำโดยปรับไฟส่องสว่างให้ได้ความสว่างในช่วง 50-150 ลักซ์ (Lux) บนหน้ายางในตำแหน่งตามแนวกลางลำต้นที่ความสูง 50, 100, และ 150 เซนติเมตรจากพื้น การถ่ายภาพชุดต่อไปจะย้ายตำแหน่งอุปกรณ์ถ่ายภาพไปจากตำแหน่งเดิมประมาณ 50-100 เซนติเมตร โดยคงระยะระหว่างกล้องกับหน้ายางให้อยู่ในช่วง 90-120 เซนติเมตร (ตัวอย่างตามรูปที่ 2.13) และปรับค่าความสว่างให้อยู่ในช่วงเดิม การถ่ายภาพต้นยางพาราต้นหนึ่งๆ จะถ่ายภาพเป็นจำนวน 1-3 ชุดภาพ เพื่อเพิ่มความหลากหลายของลำต้นหรือรอยผิวเปลือกยางที่ปรากฏในชุดข้อมูล

ข้อมูลภาพถ่ายสำหรับการพัฒนาวิธีตรวจหารอยกริดและถ้วยรองน้ำยาง จำเป็นที่จะต้องถ่ายภาพให้มีรอยกริดและถ้วยรองน้ำยางปรากฏอยู่ในภาพ จึงกำหนดเงื่อนไขการวางกล้องถ่ายภาพที่ทำให้ขนาดของรอยกริดที่ปรากฏในภาพมีความกว้างมากกว่าหรือเท่ากับครึ่งหนึ่งของความกว้างของลำต้นที่ปรากฏในภาพ และกำหนดให้ต้นยางพาราทุกต้นใช้ถ้วยรองน้ำยางชนิดพลาสติกดำขนาด 0.5 หรือ 0.75 ลิตร ตัวอย่างตามรูปที่ 2.14



รูปที่ 2.13: (มุมมองจากด้านบน) แผนภาพแสดงการวางอุปกรณ์ถ่ายภาพเมื่อย้ายตำแหน่งไปจากเดิม ประมาณ 50-100 เซนติเมตร ซึ่งยังทำให้การถ่ายภาพรอยกรีดเป็นไปตามเงื่อนไข โดยหันกล้องเข้าหาลำต้นและยังคงระยะห่างระหว่างกล้องกับหน้ายางในช่วง 90-120 เซนติเมตร

2.3.1.3 ชุดข้อมูลภาพถ่าย

ภาพถ่ายสีและภาพความลึกของแต่ละชุดภาพ จะถูกประมวลผลโดยโปรแกรมถ่ายภาพเพื่อจัดเรียงพิกเซลให้ตรงกันทั้งภาพสีและภาพความลึก (Pixel-wise Aligned) โดยอาศัยซอฟต์แวร์ไลบรารีจากผู้ผลิตกล้อง¹ ซึ่งรายละเอียดภาพถ่าย 1 ชุดภาพเป็นดังนี้

- 1 ภาพถ่ายสี RGB ขนาดกว้าง 720 สูง 1280 พิกเซล ขนาดบิตพิกเซลรวม (Bit Depth) 24 บิต (8 บิตต่อช่องสัญญาณ จำนวน 3 ช่องสัญญาณ) รูปแบบการเก็บภาพแบบ PNG (Portable Network Graphics) ตัวอย่างในรูปที่ 2.15 (ก)
- 2 ภาพถ่ายโทนเทา (Grayscale) ขนาดกว้าง 720 สูง 1280 พิกเซล ขนาดบิตพิกเซล 8 บิต (1 ช่องสัญญาณ) โดยเลือกคลิป์สัญญาณของภาพให้แปลค่าระยะความลึกตั้งแต่ 50 ถึง 200 เซนติเมตรจากระนาบกล้องเท่านั้น ซึ่งกล้องจะแมปค่าความสว่างของแต่ละพิกเซลในช่วง 0-255 ให้อยู่ในช่วงระยะความลึก 50 ถึง 200 เซนติเมตร โดยค่าของพิกเซล 255 จะเท่ากับระดับความลึก 50 เซนติเมตร และค่าของพิกเซล 0 เท่ากับระดับความลึก 200 เซนติเมตรจากกล้องเพื่อกำจัดพื้นที่ในภาพที่อยู่นอกระยะ ทั้งนี้ได้ขยายภาพออกเป็น 3 ช่องสัญญาณหรือ 24 บิต โดยการทำซ้ำภาพ 8 บิต ให้กับแต่ละช่องสัญญาณ เพื่อบันทึกภาพในรูปแบบการเก็บภาพแบบ PNG ตัวอย่างตามรูปที่ 2.15 (ข)



รูปที่ 2.14: ตัวอย่างถ้วยรองน้ำยางชนิดพลาสติกสีดำขนาด 1000 และ 700 มิลลิลิตร



(ก) ภาพถ่ายสี



(ข) ภาพถ่ายความลึก

รูปที่ 2.15: ตัวอย่างภาพถ่ายจากกล้องถ่ายภาพมุมกว้างจำนวน 1 ชุดภาพ



(ก) สวนยางพาราแหล่งที่ 1



(ข) สวนยางพาราแหล่งที่ 2

รูปที่ 2.16: ภาพถ่ายตัวอย่างสวนยางพารา

ในวิทยานิพนธ์ฉบับนี้เก็บรวบรวมภาพถ่ายต้นยางพาราด้วยกล้องถ่ายภาพมุมกว้างจำนวนทั้งสิ้นรวม 500 ชุดภาพ จากต้นยางพารา 250 ต้น ประกอบไปด้วยยางพาราพันธุ์ RRIT-251 และพันธุ์ RRIM-600 จากสวนยางพารา 2 แหล่ง โดยเลือกเฉพาะต้นที่มีความเอียงไม่เกิน 15 องศาจากแนวตั้ง สวนยางพาราตัวอย่างมีรายละเอียดดังนี้

- 1 สวนยางพาราพันธุ์ RRIM-600 อายุปลูกระหว่าง 7-8 ปี เปิดกรีดแล้วนานประมาณ 4-5 เดือน เป็นหน้ายางเปลือกใหม่ ใช้ระบบกรีดผสม S/3 และ S/4 ในพื้นที่อำเภอนาโยง จังหวัดตรัง ตามรูปที่ 2.16 (ก) พิกัด GPS : 7.544821, 99.708918
- 2 สวนยางพาราพันธุ์ RRIT-251 อายุปลูกระหว่าง 8-10 ปี เปิดกรีดแล้วนานประมาณ 1-2 ปี เป็นหน้ายางเปลือกใหม่ ใช้ระบบกรีดผสม S/2 และ S/3 ในพื้นที่อำเภอนาโยง จังหวัดตรัง ตามรูปที่ 2.16 (ข) พิกัด GPS : 7.544923, 99.717024

2.3.2 การกำหนดตำแหน่งอ้างอิงของรอยกริดและถ้วยร่อนน้ำยางในภาพ (Ground Truth Annotation)

จุดประสงค์ของการตรวจหาแนวรอยกริดและถ้วยร่อนน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้าง เพื่อต้องการตำแหน่งในสามมิติสำหรับให้กล้องระยะใกล้ (กล้อง 2) เคลื่อนเข้าทำงานในขั้นถัดไป ทำให้การตรวจหาต้องการตำแหน่งสามมิติของรอยกริดและถ้วยร่อนน้ำยางที่ไม่ละเอียดมากนักหรือมีความคลาดเคลื่อนได้ เพราะกล้องระยะใกล้ทำงานที่ระยะความลึกในช่วง 20-40 เซนติเมตร จึงเหมาะที่จะเลือกใช้กรอบสี่เหลี่ยมล้อมรอบ (Bounding Box) สำหรับกำหนดตำแหน่งอ้างอิงของรอยกริดและถ้วยร่อนน้ำยางในภาพ มากกว่าการกำหนดตำแหน่งด้วยรูปแบบอื่นๆ เช่น การกำหนดวัตถุด้วยรูปร่างเรขาคณิตอื่น เช่น วงกลม, วงรี เป็นต้น, หรือการกำหนดพื้นที่ของวัตถุในภาพระดับพิกเซล (Pixel-wise Annotation) ที่ออกแบบมาสำหรับวัตถุที่มีรูปร่างปิดซึ่งไม่เหมาะกับลักษณะรอยกริดในภาพ

การใช้งาน “กรอบสี่เหลี่ยมล้อมรอบวัตถุ” หรือ “Bounding Box” ระบุตำแหน่งของวัตถุภายในภาพถูกนำเสนอในงานวิจัย Pascal VOC [16, 17] ประโยชน์เพื่อให้สามารถวัดผลการตรวจหาตำแหน่งของวัตถุในภาพได้ ทั้งการระบุชนิดของวัตถุ (Classification) และการระบุตำแหน่งของวัตถุ (Localization) ซึ่งอัลกอริทึมการตรวจหาตำแหน่งวัตถุภายในภาพที่มีโดยส่วนมากให้เอาต์พุตเป็นกรอบสี่เหลี่ยมเช่นกัน อีกทั้งการกำหนดตำแหน่งวัตถุในการสร้างชุดข้อมูลนั้นทำได้ง่าย ดังเช่นชุดข้อมูล COCO-Stuff [7] เป็นต้น โดยการสร้างชุดข้อมูลรูปภาพสำหรับพัฒนาวิธีตรวจหาวัตถุ ผู้ที่สร้างชุดข้อมูลนั้นๆ จะเป็นผู้กำหนดกรอบสี่เหลี่ยมให้ล้อมรอบวัตถุเป้าหมายในแต่ละภาพ ซึ่งเรียกว่า “กรอบสี่เหลี่ยมอ้างอิง” หรือ “Ground Truth Bounding Box” โดยถือว่าการกรอบสี่เหลี่ยมที่ผู้สร้างชุดข้อมูลกำหนดนั้นเป็นตำแหน่งที่ถูกต้อง อ้างอิงจากความสามารถของสายตามนุษย์ที่มี Precision-Recall ดีที่สุดเมื่อเปรียบเทียบกับอัลกอริทึมการหาขอบในภาพอื่น [4] ตัวอย่างจาก [55]

การกำหนดกรอบสี่เหลี่ยมล้อมรอบวัตถุอ้างอิงโดยทั่วไปผู้สร้างชุดข้อมูลจะกำหนดตำแหน่งกรอบสี่เหลี่ยมด้วยโปรแกรมกราฟิกที่สามารถใช้เมาส์วาดกรอบสี่เหลี่ยมภายในภาพ เพื่อวาดกรอบสี่เหลี่ยมซ้อนทับกับวัตถุที่ต้องการในภาพได้ โดยผู้สร้างชุดข้อมูลสามารถวาดกรอบสี่เหลี่ยมล้อมรอบวัตถุหนึ่งๆ ให้มีลักษณะตามที่ใช้โปรแกรมเห็นสมควร เช่น ให้กรอบสี่เหลี่ยมให้ซ้อนทับกับขอบของวัตถุได้อย่างพอดี หรือให้กรอบสี่เหลี่ยมล้อมรอบวัตถุเป้าหมายโดยมีระยะห่างพอประมาณ เป็นต้น กรอบสี่เหลี่ยมล้อมรอบจากการวาดด้วยซอฟต์แวร์แต่ละกรอบ จะถูกแปลงไปเป็นชุดตัวเลข 4 จำนวนที่กำกับขนาดและตำแหน่งของกรอบนั้น $\{u_b, v_b, w_b, h_b\}$ ซึ่งเป็นพิกัดพิกเซลและขนาดของกรอบสี่เหลี่ยมในภาพ ประกอบด้วยตำแหน่งคอลัมน์มุมซ้ายบน (u_b), ตำแหน่งแถวมุมซ้ายบน (v_b), ความกว้างของกรอบในหน่วยพิกเซล (w_b), และความสูงของกรอบในหน่วยพิกเซล (h_b) เป็นต้น ซึ่งวิทยานิพนธ์ฉบับนี้ได้เลือกใช้โปรแกรม Microsoft VoTT [56] เป็นเครื่องมือกำหนดตำแหน่งกรอบสี่เหลี่ยมล้อมรอบอ้างอิงของรอยกริด และถ้วยร่อนน้ำยางในภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง

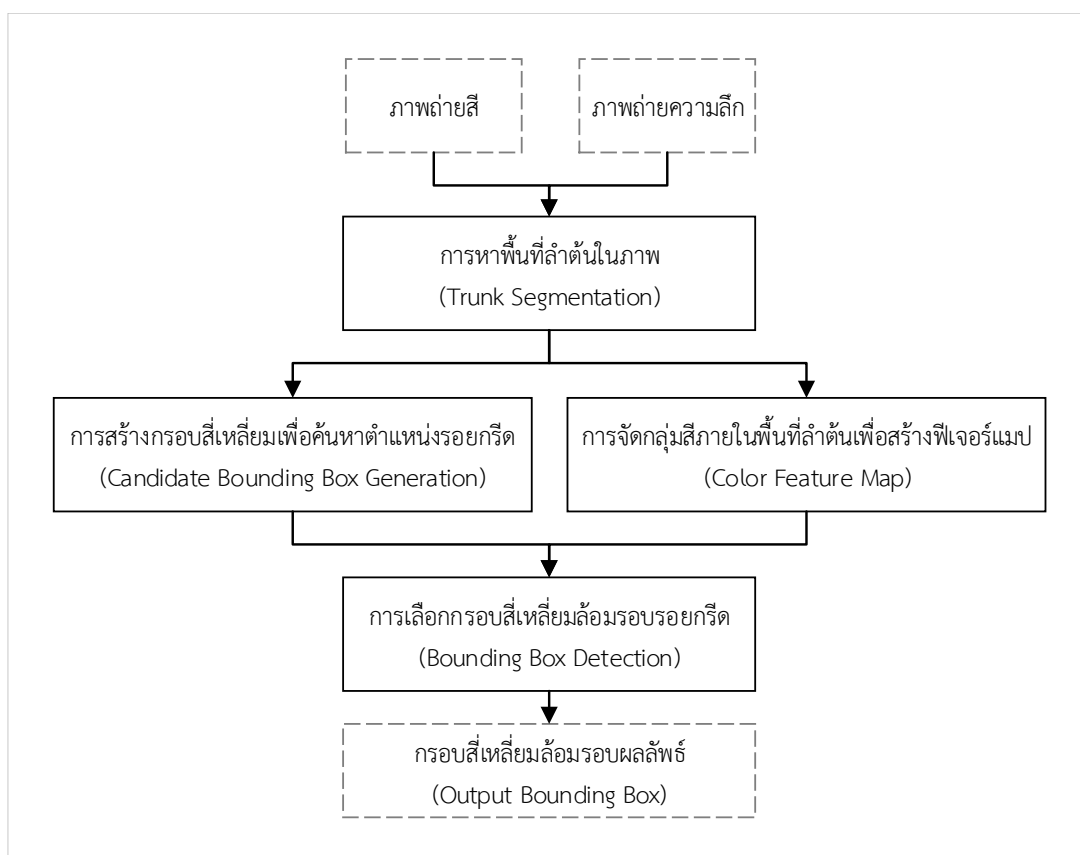


รูปที่ 2.17: ตัวอย่างการวาดกรอบสี่เหลี่ยมล้อมรอบรอยกรีดและถ้วยรองน้ำยาง

การกำหนดกรอบสี่เหลี่ยมล้อมรอบรอยกรีดอ้างอิงในภาพทำได้โดย วางมุมซ้ายบนของกรอบสี่เหลี่ยมล้อมรอบที่ตำแหน่งซ้ายสุดของแนวรอยกรีดที่ปรากฏในภาพ และมุมขวาล่างของกรอบสี่เหลี่ยมล้อมรอบอยู่ที่ตำแหน่งขวาสุดของแนวรอยกรีดที่ปรากฏในภาพ และกำหนดกรอบสี่เหลี่ยมล้อมรอบถ้วยรองน้ำยางอ้างอิงโดยให้แต่ละด้านของกรอบสี่เหลี่ยมซ้อนทับขอบถ้วยรองน้ำยางที่ปรากฏในภาพ ดังแสดงในรูปที่ 2.17 การกำหนดกรอบสี่เหลี่ยมอ้างอิงให้แต่ละชุดภาพให้ผู้กำหนดตำแหน่งวาดกรอบสี่เหลี่ยมล้อมรอบในภาพสีเท่านั้น เพราะผู้ใช้โปรแกรมสามารถแยกแยะตำแหน่งของแนวรอยกรีดและถ้วยรองน้ำยางในภาพถ่ายสีได้ดีกว่าภาพความลึก อ้างอิงจากลักษณะของประสาทสัมผัสสายตาของมนุษย์ สามารถสังเกตเห็นความไม่ต่อเนื่องบริเวณขอบภาพตามแนวรอยกรีดในภาพสีได้ดีกว่าในภาพโทนเทา [4, 29]

2.3.3 วิธีตรวจหารอยกริดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี

การตรวจหาตำแหน่งรอยกริดในภาพโดยใช้สีของบริเวณพื้นที่รอยกริดเก่าเป็นลักษณะบ่งชี้ (พีเจอร์) ในการรู้จำรอยกริด แบ่งออกเป็น 4 ขั้นตอนย่อย ตามกรอบสี่เหลี่ยมเส้นทึบในรูปที่ 2.18 ประกอบด้วย การหาพื้นที่ลำต้นในภาพ, การสร้างกรอบสี่เหลี่ยมเพื่อค้นหาตำแหน่งรอยกริด, การจัดกลุ่มสีภายในพื้นที่ลำต้นเพื่อสร้างพีเจอร์แมป, และการเลือกกรอบสี่เหลี่ยมล้อมรอบรอยกริด



รูปที่ 2.18: แผนผังวิธีการตรวจหารอยกริดบนภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี

2.3.3.1 การหาพื้นที่ลำต้นในภาพ (Trunk Segmentation)

การหาพื้นที่ลำต้นสามารถทำได้โดยวิธีการเลือกค่าพิกเซลแบบกำหนดช่วง (Thresholding) ของภาพถ่ายความลึกโทนเทา เนื่องจากภาพความลึกโทนเทาจากกล้องถ่ายภาพมุมกว้าง ที่มีค่าพิกเซล 255 ถึง 0 แทนระยะความลึกตั้งแต่ 50-200 เซนติเมตรจากกล้อง ทำให้พื้นที่นอกระยะมีค่าพิกเซลใกล้เคียงศูนย์

กำหนดให้เซต $C_1 : \{(u, v) \in U_1 \times V_1\}$ แทนพิกัดตำแหน่งพิกเซลภายในภาพ (Pixel Coordinate) โดย $U_1 = \{1, 2, 3, \dots, w_1\}$ แทนตำแหน่งตามแกนนอนของภาพ, $V_1 = \{1, 2, 3, \dots, h_1\}$ แทนตำแหน่งตามแกนตั้งของภาพ โดยมี $(u, v) = (1, 1)$ เป็นพิกัดพิกเซลมุมซ้ายบนของภาพและ $(u, v) = (w_1, h_1)$ เป็นพิกัดพิกเซลมุมขวาล่างของภาพ และกำหนดให้ภาพโทนเทาใดๆ (I) ที่มีขนาดบิตพิกเซล 8 บิต ขนาดความกว้างและสูง (Resolution) เท่ากับ $w_1 \times h_1$ แทนด้วยสมการที่ 2.2 ซึ่งเป็นเซตของความสัมพันธ์ระหว่างพิกัดตำแหน่งภาพ $(u, v) \in C_1$ กับค่าของความสว่าง ณ ตำแหน่งดังกล่าว เขียนแทนด้วยฟังก์ชัน $p(u, v) \in \{0, 1, 2, \dots, 255\}$

$$I : \{(u, v, p(u, v)) \mid (u, v) \in C_1\} \quad (2.2)$$

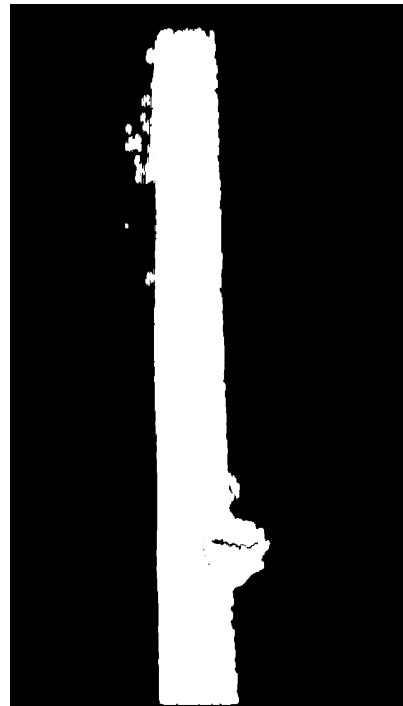
จากสมการ 2.2 กำหนดให้ $I_{D1} : \{(u, v, p_{D1})\}$ (ตัวห้อย D1) แทนภาพถ่ายความลึกจากกล้องถ่ายภาพมุมกว้าง สามารถสร้างภาพขาวดำของพื้นที่ลำต้น ($I_M : \{(u, v, p_M)\}$) ที่มีค่าเฉพาะช่วงความลึก 50-200 เซนติเมตรได้จากการเลือกค่าของพิกเซล (p_{D1}) ด้วยค่า T_d ซึ่งมีค่าใกล้เคียง 0 ตามสมการ 2.3

$$p_M(u, v) := \begin{cases} 1 & ; p_{D1}(u, v) \geq T_d \\ 0 & ; p_{D1}(u, v) < T_d \end{cases} \quad (2.3)$$

ผลลัพธ์ของการหาพื้นที่ลำต้นในภาพถ่ายความลึกในรูปแบบที่ 2.19 แสดงด้วยภาพขาวดำของพื้นที่ลำต้นรูปที่ 2.20 ทั้งนี้การหาพื้นที่ลำต้นในภาพประโยชน์เพื่อลดพื้นที่การค้นหารอยกริดในภาพ ซึ่งทำให้พื้นที่ค้นหาจะเหลือเฉพาะภายในพื้นที่ลำต้นเท่านั้น การใช้งานภาพขาวดำของพื้นที่ลำต้นจำกัดพื้นที่ลำต้นในภาพถ่ายสีทำการคูณค่าของพิกเซล แบบพิกเซลต่อพิกเซล ณ ตำแหน่งพิกัดเดียวกัน (Pixel-wise Multiplication) ระหว่างภาพถ่ายสีกับภาพขาวดำของพื้นที่ลำต้น (Trunk Mask Image) ตัวอย่างผลลัพธ์ของการหาพื้นที่ลำต้นในภาพถ่ายสีรูปที่ 2.21 แสดงในรูปที่ 2.22



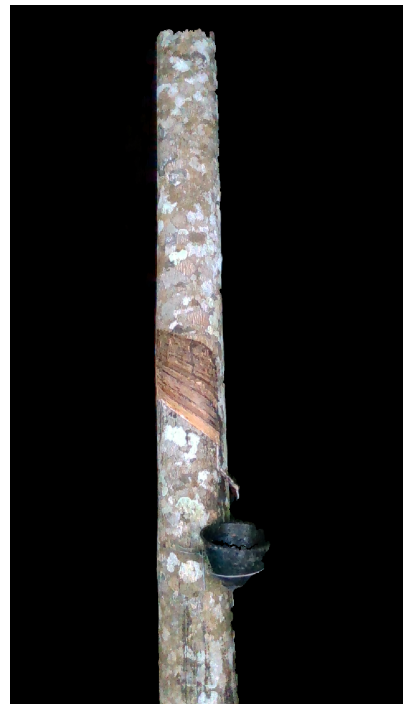
รูปที่ 2.19: ภาพถ่ายความถี่กิโลโทนเทา (I_{D1})



รูปที่ 2.20: ภาพขาวดำของพื้นที่ลำต้น (I_M)



รูปที่ 2.21: ภาพถ่ายสีจากกล้องถ่ายภาพมุมกว้าง



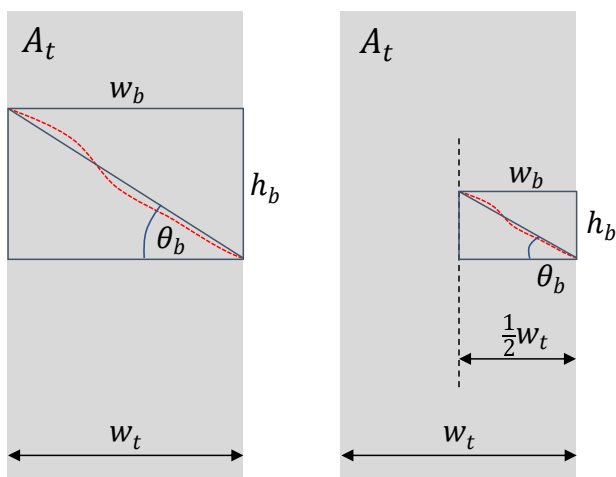
รูปที่ 2.22: ภาพถ่ายสีเมื่อผ่านการหาพื้นที่ลำต้น

2.3.3.2 การสร้างกรอบสี่เหลี่ยมเพื่อค้นหาตำแหน่งรอยกริด (Candidate Bounding Box Generation)

วิธีตรวจสอบรอยกริดให้เอาต์พุตเป็นกรอบสี่เหลี่ยมล้อมรอบรอยกริด โดยในระหว่างการตรวจหา จะมีขั้นตอนการสร้างกรอบสี่เหลี่ยมล้อมรอบที่เป็นไปได้ในพื้นที่ลำต้น เพื่อการค้นหากรอบผลลัพธ์ของรอยกริดในขั้นตอนถัดไป

การสร้างกรอบสี่เหลี่ยมที่เป็นไปได้ จะอาศัยภาพขาวดำของพื้นที่ลำต้นเพื่อให้ตำแหน่งของกรอบอยู่ภายในพื้นที่ลำต้น โดยการกำหนดขนาดของกรอบสี่เหลี่ยมที่เป็นไปได้ อาศัยเงื่อนไขความกว้างของรอยกริดที่ปรากฏในภาพ ที่มากกว่าหรือเท่ากับครึ่งหนึ่งของความกว้างของลำต้นที่ปรากฏในภาพ และจากขอบเขตของการศึกษาที่ให้ตัวอย่างสวยงามพาราที่ใช้ระบบกริดแบบ $S/2$, $S/3$ และ $S/4$ ซึ่งส่งผลให้ความกว้างที่น้อยที่สุดของกรอบสี่เหลี่ยมล้อมรอบรอยกริดที่เป็นไปได้ จะเท่ากับครึ่งหนึ่งของความกว้างลำต้นที่ปรากฏในภาพเช่นกัน อีกทั้งสัมพันธ์กับระบบการกริดแบบ $S/2$ ที่ความกว้างของแนวรอยกริดมีขนาดเท่ากับครึ่งหนึ่งของเส้นรอบวงรอบลำต้น และความกว้างที่มากที่สุดของกรอบสี่เหลี่ยมที่เป็นไปได้จะเท่ากับความกว้างเฉลี่ยของลำต้นที่ปรากฏในภาพ

ระบบกริดลงแบบวนขวารอบลำต้นที่แนวรอยกริดเฉียงทำมุม 30 ถึง 45 องศากับระดับพื้นดิน และทอดยาวจากตำแหน่งที่อยู่สูงกว่าทางซ้ายไปยังตำแหน่งต่ำกว่าทางขวา ซึ่งจากการวางกล่องที่กำหนดให้แกนนอนของภาพขนานกับพื้นดิน ทำให้แนวรอยกริดที่ปรากฏในภาพทำมุมเฉียงตั้งแต่ 30 ถึง 45 องศากับแกนนอนของภาพเช่นกัน และจากการกำหนดกรอบสี่เหลี่ยมล้อมรอบรอยกริดอ้างอิง ที่ให้มุมซ้ายบนของกรอบอยู่ที่ตำแหน่งซ้ายบนสุดของรอยกริดที่ปรากฏในภาพ และมุมขวาล่างของกรอบอยู่ที่ตำแหน่งขวาล่างสุดของรอยกริดที่ปรากฏในภาพ ดังนั้นสามารถประมาณตำแหน่งของรอยกริด คือเส้นทแยงมุมที่เชื่อมระหว่างมุมซ้ายบนและมุมขวาล่างของกรอบสี่เหลี่ยม ดังแสดงในรูปที่ 2.23



รูปที่ 2.23: แสดงกรอบสี่เหลี่ยมล้อมรอบรอยกริดภายในพื้นที่ลำต้น (A_t) ที่ขนาดความกว้างมากที่สุด ($w_b = w_t$) และที่ขนาดความกว้างน้อยที่สุด ($w_b = \frac{1}{2}w_t$) โดยแนวรอยกริดจริง (เส้นโค้งประทแยงมุม) ถูกประมาณด้วยเส้นทแยงมุมของกรอบสี่เหลี่ยม

ความกว้างลำต้นที่ปรากฏในภาพโดยเฉลี่ย (w_t) คำนวณได้จากสมการ 2.4 และความสัมพันธ์ระหว่างมุมของเส้นทแยงมุมที่ทำกับด้านกว้างของกรอบสี่เหลี่ยม (θ_b) กับความกว้าง (w_b) และความยาว (h_b) ของกรอบสี่เหลี่ยมล้อมรอบที่เป็นไปได้เป็นไปตามสมการ 2.5 ดังแสดงในรูปที่ 2.23

$$w_t = \frac{1}{h_1} \sum_{(u_i, v_i) \in C_1} p_M(u_i, v_i) \quad (2.4)$$

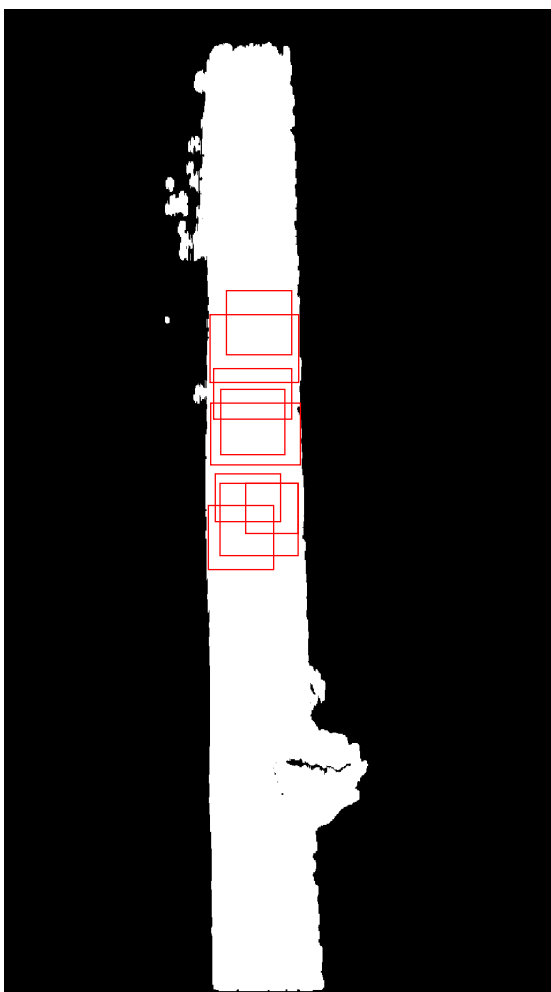
$$\begin{aligned} \{w_b\} &= \left\{ \left[\frac{1}{2} w_t, w_t \right] \right\} \\ \{\theta_b\} &= \left\{ [30^\circ, 45^\circ] \right\} \end{aligned}$$

$$h_b = w_b \tan \theta_b \quad (2.5)$$

กำหนดให้ $\{w_b\}, \{h_b\}$, และ $\{\theta_b\}$ เป็นเซตของความกว้าง, ความสูง, และมุมที่เส้นทแยงมุมทำกับด้านกว้างของกรอบสี่เหลี่ยม ของกรอบสี่เหลี่ยมที่เป็นไปได้ทั้งหมดในแต่ละภาพ โดย $w_b \in \{w_b\}$, $h_b \in \{h_b\}$, $\theta_b \in \{\theta_b\}$ ทั้งนี้หากเลือกใช้ค่า θ_b ที่มีความละเอียดสูง จะทำให้ค่าของ h_b เมื่อถูกประมาณค่าเป็นจำนวนเต็มจะไม่ต่างกันเกิน 1 พิกเซล ซึ่งความแตกต่างของความกว้างหรือความยาวของกรอบสี่เหลี่ยมล้อมรอบที่ 1 พิกเซล อาจไม่มีนัยสำคัญต่อการตรวจหาตำแหน่งของรอยกริด นอกจากนี้จำนวนสมาชิกของ $\{\theta_b\}$ จะส่งผลต่อจำนวนกรอบสี่เหลี่ยมล้อมรอบที่เป็นไปได้ทั้งหมด ซึ่งส่งผลต่อปริมาณการคำนวณในการค้นหากรอบสี่เหลี่ยมล้อมรอบรอยกริดผลลัพธ์ ตัวอย่างกรอบสี่เหลี่ยมล้อมรอบที่เป็นไปได้บางส่วนในพื้นที่ลำต้นแสดงในรูปที่ 2.24

จำนวนกรอบสี่เหลี่ยมล้อมรอบเริ่มต้น (n_b) สำหรับค่า w_b, h_b คู่หนึ่ง เมื่อนำไปกำหนดกรอบสี่เหลี่ยมเริ่มต้นด้วยการเลื่อน ด้วยระยะ n_s พิกเซล เท่ากับ

$$n_b \approx \frac{(w_t - w_b)(h_1 - h_b)}{n_s^2} \quad (2.6)$$



รูปที่ 2.24: ภาพแสดงกรอบสี่เหลี่ยมที่เปลี่ยนที่เป็นไปได้ภายในพื้นที่ลำต้นที่มีอัตราส่วนความกว้างต่อความสูง (Aspect Ratio) และขนาดต่างๆ

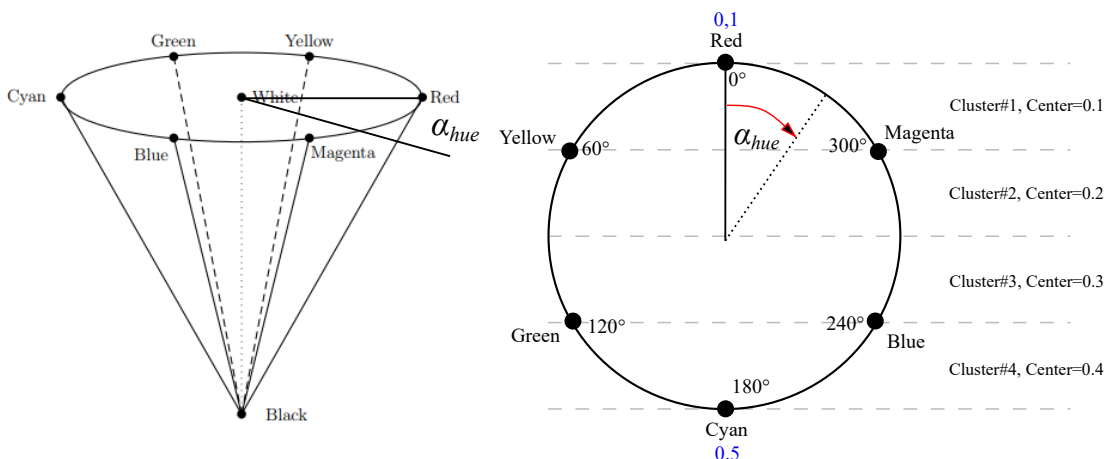
2.3.3.3 การจัดกลุ่มสีภายในพื้นที่ลำต้นเพื่อสร้างพีเจอร์แมป (Color Feature Map)

จากการสังเกตภาพถ่ายรอยกรีดพบว่าลักษณะพื้นที่บริเวณรอยกรีดเก่าในหน้ายางที่ถูกเปิดกรีดมีพื้นผิว (Texture) เป็นรอยเส้นขนานกับรอยกรีดล่าสุด และเป็นเนื้อไม้ชั้นถัดจากเปลือกชั้นนอกของลำต้น ซึ่งทั้งพื้นผิวและสีบริเวณรอยกรีดเก่ามีความสม่ำเสมอมากกว่าพื้นที่หน้ายางที่ยังไม่ผ่านการกรีด โดยเนื้อไม้บริเวณรอยกรีดเก่าจะมีสีโทนแดงมากกว่าเปลือกที่ยังไม่ผ่านการกรีด และแนวรอยกรีดจะเป็นแนวแบ่งชัดเจนระหว่างพื้นที่รอยกรีดเก่ากับพื้นที่ที่ยังไม่กรีด ดังแสดงในรูปที่ 2.25

ดังนั้นสีเนื้อไม้บริเวณรอยกรีดสามารถใช้ระบุตำแหน่งของรอยกรีดบนหน้ายางได้ด้วยการประมวลผลภาพเพื่อหาบริเวณสีโทนแดง แต่เนื่องจากภาพสีจากกล้องอยู่ในรูปแบบภาพโมเดลสี RGB ค่าของพิกเซล $p(u, v)$ ของแต่ละช่องสัญญาณ แสดงถึงความสว่างหรือการสะท้อนแสงของสีที่พิกเซลนั้นๆ (Luminance) ซึ่งไม่เหมาะต่อการหาความแตกต่างของสี [35] จึงเลือกแปลงโมเดลสีไปยังโมเดลสีอื่นที่เหมาะสมต่อการแยกความแตกต่างของสีซึ่งได้แก่ โมเดลสี HSV (Hue-Saturation-Value) และโมเดลสี $L^*a^*b^*$



รูปที่ 2.25: ภาพวาดหน้ายางและลำต้นยางพาราแสดงพื้นที่รอยกรีดเก่าและตำแหน่งของแนวรอยกรีด โดยแนวรอยกรีดคือแนวแบ่งระหว่าง เปลือกยางที่ยังไม่ผ่านการกรีดกับพื้นที่รอยกรีดเก่า



รูปที่ 2.26: ทรงกรวยอธิบายโมเดลสี HSV [35] และการจัดกลุ่มค่าสีออกเป็น 4 กลุ่ม

การจัดกลุ่มสีในภาพโมเดลสี HSV ภาพในโมเดลสี HSV (Hue-Saturation-Value) แตกต่างจากภาพในโมเดลสี RGB ที่กำหนดให้ค่าของสี (Hue), ค่าความเข้มของสี (Saturation) อยู่ในช่องสัญญาณเดียวซึ่งแยกออกจากความสว่าง (Values หรือ Brightness) ทำให้ภาพในช่องสัญญาณสี (Hue) เหมาะต่อการนำไปวิเคราะห์สี ความแตกต่างของค่าสีมากกว่าภาพโมเดลสี RGB [62, 73]

ค่าของพิกเซล ($p(u, v)$) ของภาพในโมเดลสี HSV ถูกอธิบายด้วยทรงกรวย [35] ค่าสีถูกแทนด้วยมุมภายในวงกลมหน้าตัดของทรงกรวย ซึ่งค่าของมุมถูกใช้แทนสีต่างๆ ที่มีความต่อเนื่องกันตามความยาวคลื่น ได้แก่ แดง, เหลือง, เขียว ลำดับอิงตามสเปกตรัมของสีของแสง (Visible Light Spectrum) แต่ต่างกันตรงค่ามุมที่ 0 และ 360 องศา โทนสีถูกวนกลับมาที่เดิมเพื่อใช้แทนค่าของสีแดง-ม่วง ซึ่งโดยปกติค่ามุมเหล่านี้จะถูกแปลงไปเป็นค่าของพิกเซลภายในภาพซึ่งเป็นจำนวนจริงบนเส้นจำนวน ซึ่งจะทำให้เสียความต่อเนื่องของค่าสีบริเวณโทนแดง-ม่วง เนื่องจากมุมที่ 0 องศาถูกแมปไปยังค่าพิกเซลที่ต่ำที่สุด $\min(p(u, v))$ แต่มุมที่ 360 องศาจะถูกแปลงเป็นค่าพิกเซลสูงสุด $\max(p(u, v))$ เช่นในภาพขนาดบิตพิกเซล 8 บิต ค่าพิกเซลสำหรับมุมที่ 0 องศาเท่ากับ 0 และค่าพิกเซลสำหรับมุมที่ 360 องศาจะเท่ากับ 255

การแมปค่ามุมของวงกลมหน้าตัดทรงกรวยไปเป็นค่าพิกเซลของภาพในโมเดลสี HSV ยกตัวอย่างตามสมการที่ 2.7 ซึ่งจะแปลงค่ามุมของหน้าตัดทรงกรวย α_{hue} ที่อยู่ในช่วง 0 ถึง 360 องศา ไปเป็นจำนวนจริงตั้งแต่ 0 ถึง 1 ค่าของสีโทนแดงอยู่ในบริเวณค่าของพิกเซลเมื่อเข้าใกล้ 0 ซึ่งเขียนแทนด้วย $p(u, v) \rightarrow 0$ และค่าของพิกเซลเมื่อเข้าใกล้ 0 เขียนแทนด้วย $p(u, v) \rightarrow 1$

$$p(u, v) = \frac{\alpha_{hue}}{360} \quad (2.7)$$

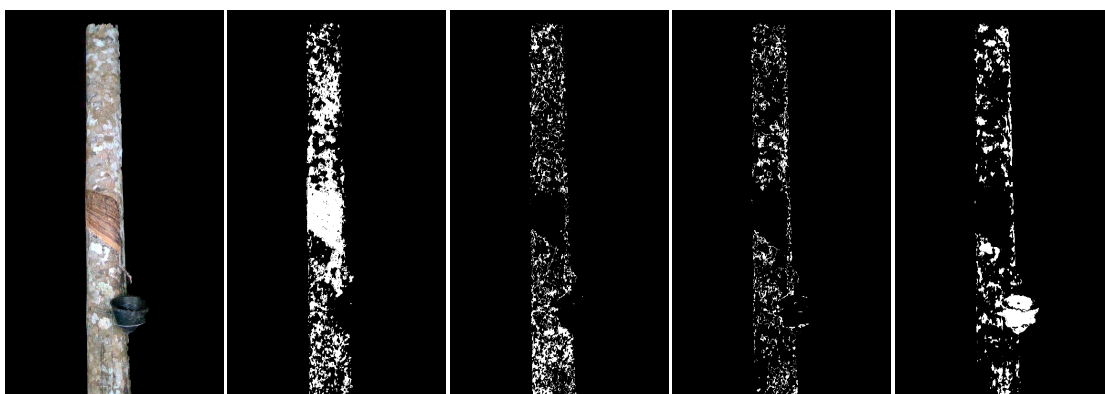
เพื่อแก้ปัญหาความไม่ต่อเนื่องของค่าตัวเลขบริเวณโทนสีแดงของโมเดลสี HSV จึงเปลี่ยนวิธีการแปลงค่ามุมของวงล้อสี (Hue) ไปเป็นค่าพิกเซลของภาพตามของสมการ 2.7 ข้างต้น ให้เกิดการวกกลับสองจุด เพื่อให้สามารถจัดกลุ่มของสีโดยใช้วิธีจัดกลุ่มแบบค่ากลาง K กลุ่ม (K-means Clustering Algorithm หรือ K-Nearest Neighbors) บนค่าพิกเซล $p(u, v)$ ได้ กล่าวคือทำให้กลุ่มของสีที่มีค่า $p(u, v) \rightarrow 0$ และ $p(u, v) \rightarrow 1$ ในสมการ 2.7 เป็นกลุ่มเดียวกัน โดยสร้างเงื่อนไขการแปลงค่าตามสมการ 2.8 เมื่อกำหนดให้

$p(u, v)$ มีค่าตามสมการ 2.7

$$\hat{p}(u, v) := \begin{cases} p(u, v) & ; p(u, v) < 0.5 \\ 1 - p(u, v) & ; p(u, v) \geq 0.5 \end{cases} \quad (2.8)$$

การแยกสีโทนแดงทำโดยจัดกลุ่มค่าสี ($\{\hat{p}(u, v)\}$) ภายในพื้นที่ลำต้นด้วยวิธีจัดกลุ่มแบบค่ากลาง K กลุ่ม ตาม 5 โทนสีหลักของวงล้อสี ซึ่งได้แก่ สีแดง, สีเหลือง, สีเขียว, สีนํ้าเงิน และสีม่วง แต่เนื่องจากโทนสีของเนื้อไม้บริเวณรอยกรีดเก่าอยู่ในช่วงโทนสีแดง (ม่วงแดง-แดง-เหลืองแดง) ค่าสีในภาพเมื่อผ่านการแปลงค่าของสมการ 2.8 ที่ควบรวมโทนสีม่วงแดง-แดง-แดงเหลือง ให้อยู่ในกลุ่มเดียวกัน ทำให้สามารถแบ่งการจัดกลุ่มสีออกเป็น 4 กลุ่ม ได้แก่ ม่วงแดง-แดง-เหลืองแดง, ม่วง-เหลือง, นํ้าเงินม่วง-เขียวเหลือง, และนํ้าเงิน-นํ้าเงินเขียว-เขียว ทำให้สามารถกำหนดค่ากลางของกลุ่มเริ่มต้นจำนวน 4 กลุ่ม (Initial Centroids $K=4$) ซึ่งเป็นค่าคงที่ ได้แก่ 0.1, 0.2, 0.3 และ 0.4 ดังแสดงในรูปที่ 2.26 จากนั้นเลือกเฉพาะกลุ่มที่มีค่าอยู่ในโทนสีแดงซึ่งเป็นกลุ่มที่มีค่ากลางใกล้เคียง 0

ตัวอย่างผลการจัดกลุ่มสีแสดงในรูปที่ 2.27 จากผลการทดลองแม้วิธีแปลงค่าตามสมการ 2.8 สามารถแก้ปัญหาเชิงตัวเลขของการแมปค่ามวงล้อสีไปยังค่าพิกเซลเพื่อการจัดกลุ่มได้ แต่ได้สร้างปัญหาตามมาหลายประเด็น เช่น กลุ่มของสีมีการปนกันมากขึ้น เพราะการทับซ้อนของค่ามวงเนื่องจากสมการ 2.8 ทำให้สีสองสีที่แตกต่างกัน (ตรงข้ามกันตามแนวตั้งในวงล้อสี) ในวงล้อสีมีค่าพิกเซลเดียวกัน ซึ่งส่งผลไปยังการจัดกลุ่มของค่าสีเป็น 4 กลุ่มย่อย นอกจากนี้ภาพผลลัพธ์จะเป็นภาพขาวดำของภาพขาวดำของกลุ่มพิกเซลที่มีสีโทนแดงเพียงกลุ่มเดียวเท่านั้น ซึ่ง 3 กลุ่มที่เหลือเป็นการประมวลผลที่ไม่จำเป็น



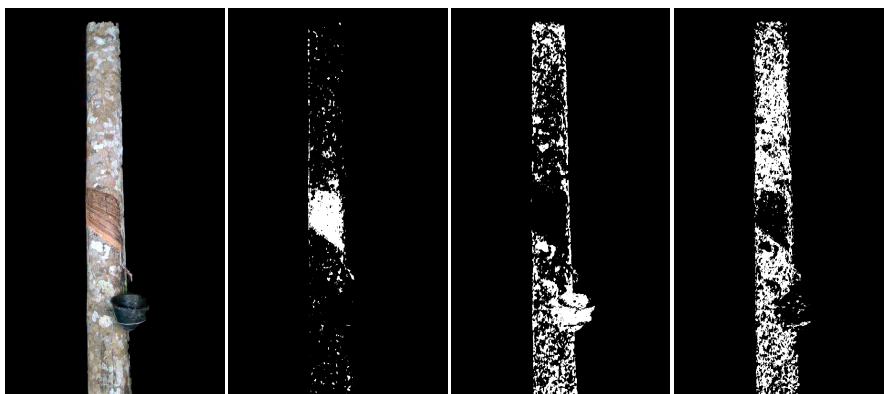
รูปที่ 2.27: ตัวอย่างการจัดกลุ่มค่าสีออกเป็น 4 กลุ่ม ของภาพในโมเดลสี HSV เมื่อถูกตัดแปลงค่าด้วยสมการที่ 2.8 ตามรูปที่ 2.26, โดยรูปทางซ้ายแสดงภาพสี RGB ต้นฉบับที่ถูกจำกัดพื้นที่ลำต้นจากรูปที่ 2.22 และกลุ่มภาพทางขวาแสดงภาพขาวดำของกลุ่มสีที่มีค่ากลางใกล้เคียงกับค่ากลางเริ่มต้น ซึ่งได้แก่ 0.1, 0.2, 0.3 และ 0.4 เรียงตามลำดับ

การจัดกลุ่มสีในภาพโมเดลสี $L^*a^*b^*$ โมเดลสี $L^*a^*b^*$ เหมาะกับการจัดกลุ่มสีโทนแดงมากกว่าโมเดลสี RGB เช่นกัน เนื่องจากช่องสัญญาณภาพ a^* ของโมเดลสี $L^*a^*b^*$ มีความสัมพันธ์โดยตรงกับสีโทนแดง-เขียว ซึ่งต่างจากช่องสัญญาณสี (Hue) ของโมเดลสี HSV ที่รวมทุกย่านสีเอาไว้ [35, 62] ทำให้การจัดกลุ่มค่าสีในช่องสัญญาณ a^* ก็เพียงพอต่อการแยกพื้นที่โทนสีแดง การจัดกลุ่มสีอาศัยวิธีจัดกลุ่มแบบค่ากลาง K กลุ่ม จัดกลุ่มสีในช่องสัญญาณ a^* เป็น 3 กลุ่ม ($K=3$) ตามสีที่เป็นสีโทนแดง, สีโทนเขียว และกลุ่มสีโทนกลางที่อยู่ระหว่างโทนแดง-เขียว พื้นที่ของแต่ละกลุ่มถูกแปลงเป็นภาพขาวดำ ผลลัพธ์ตัวอย่างแสดงในรูปที่ 2.28 โดยกำหนดค่ากลางของกลุ่มเริ่มต้นจากค่าทางสถิติของค่าพิกเซลภายในภาพช่องสัญญาณ a^* ดังนี้ กำหนดเครื่องหมาย \circ แทนการคูณค่าพิกเซลต่อพิกเซลที่ตรงตำแหน่งกัน (Element-wise Multiplication)

$$\text{กลุ่มสีโทนแดง} : \max(\{p_{a^*} \circ p_M\})$$

$$\text{กลุ่มสีโทนกลาง} : 0$$

$$\text{กลุ่มสีโทนเขียว} : \min(\{p_{a^*} \circ p_M\})$$



รูปที่ 2.28: ตัวอย่างการจัดกลุ่มสีออกเป็น 3 กลุ่มในช่องสัญญาณ a^* ของโมเดลสี CIELAB โดยรูปทางซ้ายมือแสดงภาพสีตัวอย่างต้นฉบับ RGB (รูปที่ 2.22) และภาพทางขวามือแสดงกลุ่มสีที่มีโทนสีแดงมาก, ปานกลาง และ โทนสีเขียวมาก ตามลำดับ ซึ่งใช้วิธีจัดกลุ่มแบบค่ากลาง K กลุ่ม ในการจัดกลุ่มเช่นเดียวกันกับการจัดกลุ่มสีในช่องสัญญาณ Hue ของโมเดลสี HSV

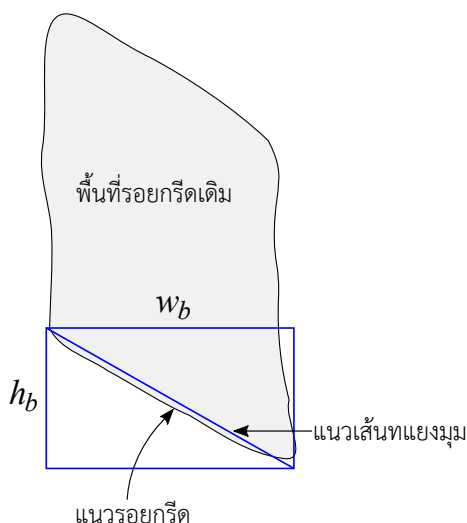
ภาพขาวดำของพื้นที่กลุ่มโทนสีแดง อาจพิจารณาได้ว่าภาพขาวดำ (Binary) ของพิกเซลกลุ่มสีโทนแดงจากการจัดกลุ่มสี คือพีเจอร์แมปจากการแยกสี (Color Feature Maps) เพื่อใช้ประโยชน์ในการตรวจหากรอบของรอยกริดต่อไป โดยพีเจอร์แมปแทนด้วย I_{map} ตามสมการที่ 2.9 ซึ่งได้แก่ ภาพขาวดำกลุ่มสีแดงจากการจัดกลุ่มสีด้วยภาพตามโมเดลสี HSV และภาพขาวดำกลุ่มสีโทนแดงจากการจัดกลุ่มสีในภาพตามโมเดลสี $L^*a^*b^*$

$$I_{\text{map}} : \{(u, v, p_{\text{map}}(u, v)) ; p_{\text{map}} \in \{0, 1\}\} \quad (2.9)$$

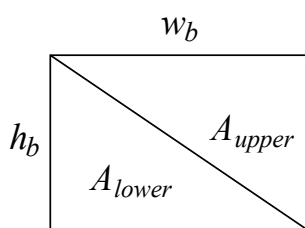
2.3.3.4 การเลือกกรอบสี่เหลี่ยมล้อมรอบรอยกรีด (Bounding Box Detection)

ขั้นตอนการเลือกกรอบสี่เหลี่ยมที่ตรงตำแหน่งรอยกรีดจะอาศัยภาพขาวดำของกลุ่มสีโทนแดงจากขั้นตอนการจัดกลุ่มสี โดยภาพขาวดำของกลุ่มสีโทนแดง (ตามสมการที่ 2.9 ในข้อย่อย 2.3.3.3) จะปรากฏตำแหน่งของแนวรอยกรีดเป็นแนวแบ่งระหว่างพื้นที่บริเวณรอยกรีดเก่าที่มีค่าพิกเซลเท่ากับ 1 กับพื้นที่ลำดับที่ยังไม่ถูกรีดซึ่งเป็นพื้นที่ที่มีค่าพิกเซลเท่ากับศูนย์

การกำหนดกรอบสี่เหลี่ยมล้อมรอบอ้างอิงของรอยกรีด ที่ประมาณแนวรอยกรีดในภาพด้วยเส้นทแยงมุมหลักของกรอบสี่เหลี่ยม ทำให้กรอบที่ตรงตำแหน่งรอยกรีดจะต้องมีพื้นที่บริเวณรอยกรีดเก่า และพื้นที่ที่ยังไม่ถูกรีดอยู่ภายในพื้นที่ของกรอบสี่เหลี่ยม แนวเส้นทแยงมุมระหว่างมุมซ้ายบนกับมุมขวาล่างของกรอบสี่เหลี่ยมจะเป็นแนวแบ่งพื้นที่ทั้งสอง โดยพื้นที่รอยกรีดเก่าจะอยู่ในพื้นที่สามเหลี่ยมบนภายในกรอบ (Upper Triangle, A_{upper}) และพื้นที่หน้ายางที่ยังไม่ถูกรีดจะอยู่ในพื้นที่สามเหลี่ยมล่างของกรอบ (Lower Triangle, A_{lower}) ดังแสดงในรูปที่ 2.29 และ 2.30



รูปที่ 2.29: ตัวอย่างกรอบสี่เหลี่ยมที่ตรงตำแหน่งรอยกรีด



รูปที่ 2.30: พื้นที่สามเหลี่ยมบนและล่างจากการแบ่งพื้นที่ภายในกรอบสี่เหลี่ยมด้วยเส้นทแยงมุม

ค่าคะแนนการตรวจหา คะแนนการตรวจหากรอบที่ตรงตำแหน่งรอยกริดแต่ละกรอบ (*boxScore*) กำหนดด้วยอัตราส่วนระหว่าง ความแตกต่างของพื้นที่สามเหลี่ยมบนกับพื้นที่สามเหลี่ยมล่างต่อพื้นที่ของกรอบสี่เหลี่ยมทั้งหมดที่วางอยู่บนภาพขาวดำของกลุ่มสีโทนแดง (I_{map}) ตามสมการที่ 2.10 ซึ่งค่าคะแนนจะอยู่ในช่วง $[-0.5, 0.5]$ โดยค่าที่มากกว่าแสดงถึงกรอบสี่เหลี่ยมนั้นตรงตำแหน่งรอยกริดมากกว่ากรอบที่มีค่าคะแนนน้อยกว่า เนื่อง nA_{upper} จะมีค่าสูงกว่า nA_{lower} บนพีเจอร์แมปในบริเวณรอยกริด ดังนั้นการตรวจหากรอบสี่เหลี่ยมที่ล้อมรอบรอยกริดจะค้นหากรอบที่มีค่าคะแนนที่สูงที่สุดจากกรอบสี่เหลี่ยมที่เป็นไปได้ทั้งหมด ที่ถูกสร้างในข้อย่อย 2.3.3.2

$$boxScore = \frac{nA_{upper} - nA_{lower}}{w_b \times h_b} \quad (2.10)$$

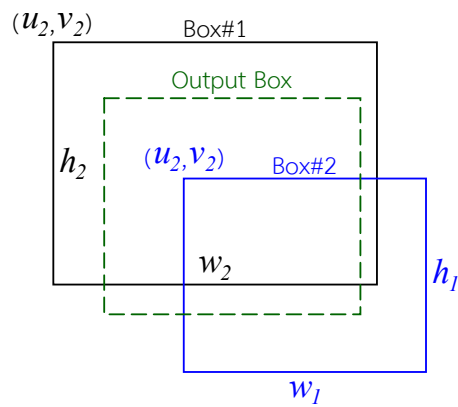
nA_{upper} : จำนวนพิกเซลมีค่าเท่ากับ 1 ใน I_{map} ภายในพื้นที่สามเหลี่ยมบนของกรอบสี่เหลี่ยม

nA_{lower} : จำนวนพิกเซลมีค่าเท่ากับ 1 ใน I_{map} ภายในพื้นที่สามเหลี่ยมล่างของกรอบสี่เหลี่ยม

การรวมกรอบ การสร้างกรอบสี่เหลี่ยมล้อมรอบที่เป็นไปได้ซึ่งมีหลากหลายขนาดความกว้างและความสูง ส่งผลต่อขนาดพื้นที่ของกรอบสี่เหลี่ยมนั้นๆ ปัญหาหนึ่งที่จะเกิดขึ้นเมื่อใช้ค่าคะแนนตามสมการ 2.10 เป็นตัวชี้วัดตำแหน่งรอยกริด คือ ค่าคะแนนของกรอบที่มีขนาดเล็กมีโอกาสมากกว่าค่าคะแนนของกรอบขนาดใหญ่กว่า ยกตัวอย่างดังในกรณีที่กรอบสี่เหลี่ยมสองขนาดตรงตำแหน่งของรอยกริด กรอบสี่เหลี่ยมที่มีขนาดใหญ่จะมีโอกาสที่จะคลุมพื้นที่ว่าง (พื้นที่ที่มีค่าพิกเซลเท่ากับศูนย์, $p(u, v) = 0$) ได้มากกว่ากรอบสี่เหลี่ยมที่มีขนาดเล็ก ส่งผลต่อค่าคะแนนที่ทำให้กรอบที่มีขนาดเล็กถูกเลือกเป็นกรอบสี่เหลี่ยมล้อมรอบรอยกริด ดังนั้นการเลือกกรอบสี่เหลี่ยมจึงต้องการวิธีเลือกกรอบสี่เหลี่ยมของรอยกริดโดยเลือกขนาดของกรอบให้เหมาะสมกับขนาดของรอยกริดที่ปรากฏในภาพ

จากข้อเท็จจริงที่ภาพถ่ายหน้ายางหนึ่งๆ จะมีรอยกริดเพียงตำแหน่งเดียว ส่งผลให้การเลือกกรอบสี่เหลี่ยมล้อมรอบรอยกริดจะต้องมีเพียงกรอบเดียวเช่นกัน ดังนั้นการเลือกกรอบสี่เหลี่ยมของรอยกริดจะเลือกกรอบที่มีค่าคะแนนที่มากที่สุดในแต่ละขนาดความกว้าง แทนการเลือกกรอบสี่เหลี่ยมที่มีค่าคะแนนสูงสุดโดยไม่พิจารณาขนาดของกรอบ กรอบสี่เหลี่ยมที่ตรงรอยกริดผลลัพธ์จะเป็นการรวมขนาดและตำแหน่งของกรอบสี่เหลี่ยมที่มีค่าคะแนนที่สูงที่สุดในแต่ละความกว้าง ผ่านการรวมแบบถ่วงน้ำหนัก (Weighting) ด้วยค่าคะแนน

ตัวอย่างการรวมกรอบแสดงในรูปที่ 2.31 สมมติให้กรอบที่ 1 (Box#1) ที่มีชุดเลขกำกับ $\{u_1, v_1, w_1, h_1\}$ มีค่าคะแนนเท่ากับ s_1 และกรอบที่ 2 (Box#2) ที่มีชุดเลขกำกับ $\{u_2, v_2, w_2, h_2\}$ มีค่าคะแนนเท่ากับ s_2 ทั้งสองเป็นกรอบสี่เหลี่ยมล้อมรอบที่มีค่าคะแนนมากที่สุดในแต่ละความกว้าง กรอบสี่เหลี่ยมผลลัพธ์ (Output Box) ที่เกิดจากการรวมกรอบที่ 1 และกรอบที่ 2 มีชุดเลขกำกับ $\{u_o, v_o, w_o, h_o\}$ สามารถคำนวณได้ดังนี้



รูปที่ 2.31: ภาพตัวอย่างแสดงกรอบสี่เหลี่ยมผลลัพธ์ (Output Box) ที่เกิดจากการรวมกรอบสี่เหลี่ยมที่มีค่าคะแนนสูงสุดในแต่ละขนาดความกว้าง

$$u_o = \frac{s_1 u_1 + s_2 u_2}{s_1 + s_2}, v_o = \frac{s_1 v_1 + s_2 v_2}{s_1 + s_2}$$

$$w_o = \frac{s_1 w_1 + s_2 w_2}{s_1 + s_2}, h_o = \frac{s_1 h_1 + s_2 h_2}{s_1 + s_2}$$

2.3.4 วิธีตรวจหารอยกริดและถ้วยรอน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยวิธีตรวจหาวัตถุชนิด โครงข่ายประสาทเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN (Faster Region with Deep-Convolutional Neural Network)

การพัฒนาวิธีตรวจหาวัตถุในรูปภาพด้วยการเรียนรู้จากชุดข้อมูล (Supervised Learning) โดยอาศัยโครงข่ายประสาทเชิงลึกแบบคอนโวลูชัน (Deep Convolutional Neural Network, Deep-CNN) เป็นที่นิยมมากขึ้นในปัจจุบัน เนื่องจากเป็นวิธีที่ให้ความแม่นยำและมีประสิทธิภาพมากกว่าวิธีตรวจหาวัตถุแบบเดิม (Traditional Machine Learning) ซึ่งเป็นวิธีที่ผู้พัฒนาออกแบบพีเจอร์ขึ้นเอง ซึ่งสำหรับปัญหาการตรวจหารอยกริดและถ้วยรอน้ำยางสามารถใช้วิธีตรวจหาวัตถุแบบโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันได้เช่นกัน

การพัฒนาวิธีตรวจหารอยกริดและถ้วยรอน้ำยางในภาพ ด้วยวิธีตรวจหาวัตถุในภาพชนิดโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN ผ่านการเรียนรู้แบบถ่ายโอน (Transfer Learning) ด้วยชุดข้อมูลภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง ประกอบไปด้วย การพิจารณาเลือกใช้เน็ตเวิร์กที่ผ่านการสอนมาก่อน (Pre-trained Network), การพิจารณาโครงสร้างของเน็ตเวิร์กตรวจหาวัตถุแบบโครงข่ายประสาทเชิงลึกแบบคอนโวลูชัน (Deep-CNN Object Detector Architecture) สำหรับตรวจหารอยกริดและถ้วยรอน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้าง, การกำหนดเลเยอร์ของ Faster R-CNN ให้กับ MobileNetV2, การสอนเน็ตเวิร์กแบบปรับละเอียด (Fine Tuning), การเตรียมข้อมูลภาพอินพุตสำหรับสอนเน็ตเวิร์กและการจัดกลุ่มทดลอง, และการเพิ่มจำนวนภาพของชุดข้อมูล (Image Augmentation)

2.3.4.1 การพิจารณาเลือกใช้เน็ตเวิร์กที่ผ่านการสอนมาก่อน

การพัฒนาวิธีตรวจหาวัตถุแบบโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันขั้นใหม่ทั้งหมดด้วยการเรียนรู้จากชุดข้อมูล มีข้อจำกัดที่สำคัญคือ การพัฒนาต้องการข้อมูลในปริมาณมากและหลากหลาย เพื่อให้เน็ตเวิร์กเรียนรู้และเพื่อความน่าเชื่อถือต่ออนุมานผลของการตรวจหาไปยังชุดข้อมูลอื่นที่ไม่ใช่สอนหรือทดสอบเน็ตเวิร์ก ซึ่งในกรณีภาพถ่ายหน้ายางด้วยกล้องถ่ายภาพมุมกว้างที่มีจำนวนทั้งสิ้น 500 ชุดภาพ ซึ่งมีปริมาณที่เหมาะสมต่อการสอนเน็ตเวิร์กด้วยวิธีการเรียนรู้แบบถ่ายโอน (Transfer Learning) โดยการปรับละเอียดค่าน้ำหนักของเน็ตเวิร์กที่ผ่านการสอนมาก่อน (Fine-Tuning) ประโยชน์เพื่อทำให้การเรียนรู้บนชุดข้อมูลใหม่สามารถทำได้ง่ายและเร็วขึ้น เนื่องจากข้อมูลมีปริมาณและความหลากหลายที่น้อยกว่า เมื่อเปรียบเทียบกับชุดข้อมูลในงานพัฒนาการตรวจหาตำแหน่งวัตถุแบบโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันอื่นๆ เช่น ImageNet [14], MS-COCO [7]

การเรียนรู้แบบถ่ายโอนเป็นวิธีปฏิบัติหนึ่งสำหรับการเรียนรู้ของโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันด้วยชุดข้อมูลเฉพาะ (Custom Dataset) ซึ่งเป็นข้อมูลใหม่ที่มีปริมาณข้อมูลไม่มากนัก โดยดัดแปลงโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันเดิมให้มีส่วนของเน็ตเวิร์กในชั้นต้นหรือกลุ่มของเลเยอร์ชั้นต้น (Low-Level Layers) เป็นเน็ตเวิร์กที่ผ่านการเรียนรู้ด้วยชุดข้อมูลขนาดใหญ่มาก่อน ทำหน้าที่สกัดพีเจอร์และเป็นฐานให้กับเน็ตเวิร์กในชั้นสูงต่อไป (High-Level Layers) โดยเน็ตเวิร์กในชั้นสูงกว่าเป็นชั้นที่ถูกสร้างขึ้นใหม่เพื่อรู้จำชนิดข้อมูลตามข้อมูลชุดใหม่ที่ผู้พัฒนาต้องการ เรียกชั้นของเน็ตเวิร์ก (Layers) ที่ผ่านการเรียนรู้ด้วยชุดข้อมูลขนาดใหญ่มาก่อนนี้ว่า “Pre-trained Network” หรือ “Pre-trained Layers” โดยขณะโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันที่ถูกดัดแปลงกำลังเรียนรู้ข้อมูลชุดใหม่ (Training) การสอนจะอนุญาตให้

เน็ตเวิร์กปรับเปลี่ยนค่าน้ำหนัก (Network Weight) ของกลุ่มเลเยอร์ชั้นลำดับสูง (High-Level Layers) ในอัตราส่วนที่มากกว่าค่าน้ำหนักของกลุ่มของเลเยอร์ชั้นลำดับต่ำกว่า (Low-Level Layers) ซึ่งเรียกรวมวิธีการเรียนรู้ด้วยการปรับค่าน้ำหนักแบบนี้ว่าการปรับละเอียด (Fine-Tuning)

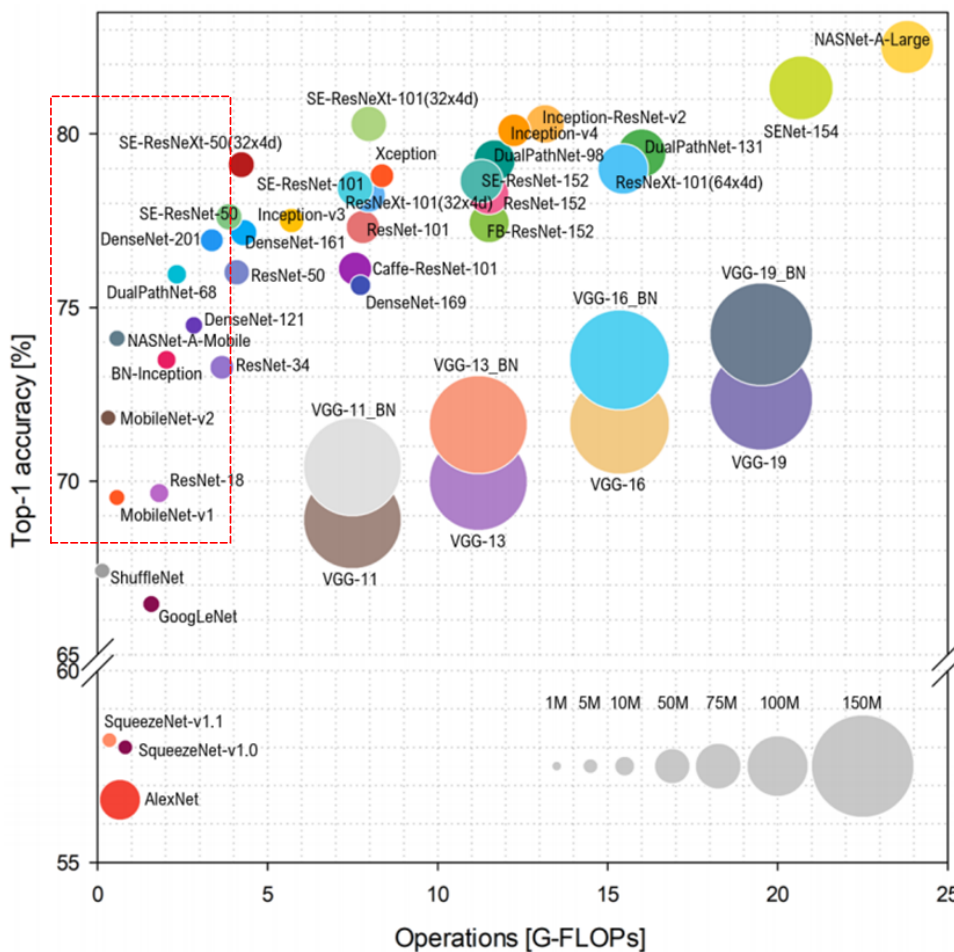
อ้างอิงจากข้อมูลปี 2018 [6] รูปที่ 2.32 แสดงกลุ่มของโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันที่ผ่านการสอนมาก่อน พลอตเรียงตามลำดับความแม่นยำในการรู้จำวัตถุ (Classification Accuracy) บนฐานข้อมูลขนาดใหญ่ ขนาดวงกลมแสดงถึงขนาดของเน็ตเวิร์กซึ่งแทนจำนวนพารามิเตอร์ภายในเน็ตเวิร์ก (Learn-able Parameters) และปริมาณการประมวลผลที่เน็ตเวิร์กต้องการขณะตรวจหา (Floating Point Operations per Second, FLOPs) โดยการเลือกเน็ตเวิร์กเพื่อตรวจหาตำแหน่งของรอยกริดและถ้วยรอนน้ำยางในภาพตามแนวคิดระบบหุ่นยนต์ทำสวนยางพาราอัตโนมัติ เหมาะที่จะเลือกเน็ตเวิร์กที่อยู่ในกลุ่มของเน็ตเวิร์กมีขนาดเล็กและมีปริมาณการประมวลผลที่ต่ำ สำหรับการนำไปใช้งานบนระบบคอมพิวเตอร์เคลื่อนที่ (Mobile Computing Platform) ซึ่งมีความสามารถในการประมวลผลที่จำกัด โดยจากภาพภายในกรอบเส้นประเป็นกลุ่มของเน็ตเวิร์กที่เหมาะสมต่อการนำมาใช้งาน โดยในวิทยานิพนธ์ฉบับนี้ได้เลือกใช้ MobileNetV2 [69] เนื่องจากในขณะทำงานวิจัย MobileNetV2 ที่ผ่านการสอนได้เผยแพร่สู่สาธารณะให้สามารถดาวน์โหลดเพื่อนำไปใช้ได้²

2.3.4.2 การพิจารณาโครงสร้างของเน็ตเวิร์กตรวจหาวัตถุแบบโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันสำหรับตรวจหารอยกริดและถ้วยรอนน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้าง

ลักษณะของเอาต์พุตของโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันสำหรับตรวจหาวัตถุในภาพ (Deep-CNN Object Detector) แตกต่างจากโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันสำหรับรู้จำประเภทวัตถุ (Deep-CNN Object Classifier) โดยเอาต์พุตของเน็ตเวิร์กสำหรับตรวจหาวัตถุในภาพจะให้ค่าตัวเลขกำกับกรอบสี่เหลี่ยมที่ล้อมรอบวัตถุเป้าหมาย (Object Bounding Box) และระบุชนิดของวัตถุภายในกรอบสี่เหลี่ยมนั้นได้ (Object Classification) ทำให้สามารถจำแนกโครงสร้างของโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันสำหรับตรวจหาตำแหน่งวัตถุในภาพออกเป็นสองประเภทตามลักษณะเอาต์พุต ได้แก่ เน็ตเวิร์กชนิดสองส่วน (Two-Stage Deep-CNN Object Detector) ยกตัวอย่างเช่น R-CNN [22], Fast-RCNN [20] หรือ Faster-RCNN เป็นต้น โดยโครงสร้างของเน็ตเวิร์กจะประกอบไปด้วย 2 เน็ตเวิร์กย่อยที่ใช้งานเลเยอร์จำนวนหนึ่งร่วมกัน (Shared Layers) โดยเน็ตเวิร์กย่อยหนึ่งทำหน้าที่ทำนายตำแหน่งกรอบสี่เหลี่ยมล้อมรอบ (Region Proposal Network) และอีกเน็ตเวิร์กย่อยหนึ่งทำหน้าที่ทำนายชนิดของวัตถุ (Detection Network) ภายในกรอบสี่เหลี่ยมล้อมรอบที่ทำนายจากจากอีกเน็ตเวิร์กหนึ่ง และเน็ตเวิร์กชนิดส่วนเดียว (One-Stage Deep-CNN Object Detector) ที่รวมทั้งสองส่วนเข้าไว้ในเน็ตเวิร์กเดียวกัน ตัวอย่างเช่น Single Shot Multibox Detector (SSD) [50], YOLO และ YOLO9000 [64, 65] เป็นต้น

ลักษณะโครงสร้างของเน็ตเวิร์กตรวจหาวัตถุส่งผลต่อความแม่นยำในการตรวจหาตำแหน่งวัตถุ โดยเน็ตเวิร์กแบบสองส่วนมีความแม่นยำในการตรวจหาตำแหน่งวัตถุที่มีขนาดเล็กสูงกว่าเน็ตเวิร์กแบบ

²PyTorch: pytorch.org/hub/pytorch_vision_mobilenet_v2,
TensorFlow: github.com/tensorflow/models/tree/master/research/slim/nets/mobilenet,
MATLAB: mathworks.com/help/deeplearning/ref/mobilenetv2.html



รูปที่ 2.32: ข้อมูลจาก [6] แผนภาพโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันที่ผ่านการสอนมาก่อน โดยในกรอบสี่เหลี่ยมเส้นประเป็นกลุ่มที่ของเน็ตเวิร์กที่เหมาะสมสำหรับระบบหุ่นยนต์ทำสวนยางพารา

ส่วนเดียว เนื่องจากเน็ตเวิร์กแบบส่วนเดียวถูกพัฒนาโดยควมรวมลอสมฟังก์ชัน (Loss Function) ของการสอนเพื่อการทำนายตำแหน่งและการทำนายประเภทเข้าด้วยกัน ทำให้เน็ตเวิร์กแบบส่วนเดียวสามารถสอนโดยการป้อนข้อมูลและผลลัพธ์แบบต้นถึงปลายได้ (End-to-End Training) แต่ไม่ได้เน้นให้เน็ตเวิร์กมีความแม่นยำในการทำนายตำแหน่งกรอบสี่เหลี่ยมให้ตรงกับตำแหน่งวัตถุมากนักเมื่อเทียบกับเน็ตเวิร์กแบบสองส่วน ยกตัวอย่างเช่น YOLO จะใช้ลักษณะการทำนายกรอบสี่เหลี่ยมด้วยการแบ่งกริดภายในภาพแบบหยาบ เพื่อทำนายตำแหน่งของกรอบสี่เหลี่ยมภายในกริดนั้นๆ ด้วยการใช้ลอสมฟังก์ชัน SSE (Sum of Squared Error) ทำให้เน็ตเวิร์กไม่สามารถทำนายตำแหน่งของวัตถุขนาดเล็กได้ดี เท่ากับเน็ตเวิร์กแบบสองส่วนที่มีเน็ตเวิร์กย่อยหนึ่งเรียนรู้ตำแหน่งที่เป็นไปได้ของวัตถุโดยเฉพาะเพื่อผลิตกรอบสี่เหลี่ยมล้อมรอบวัตถุนั้นๆ

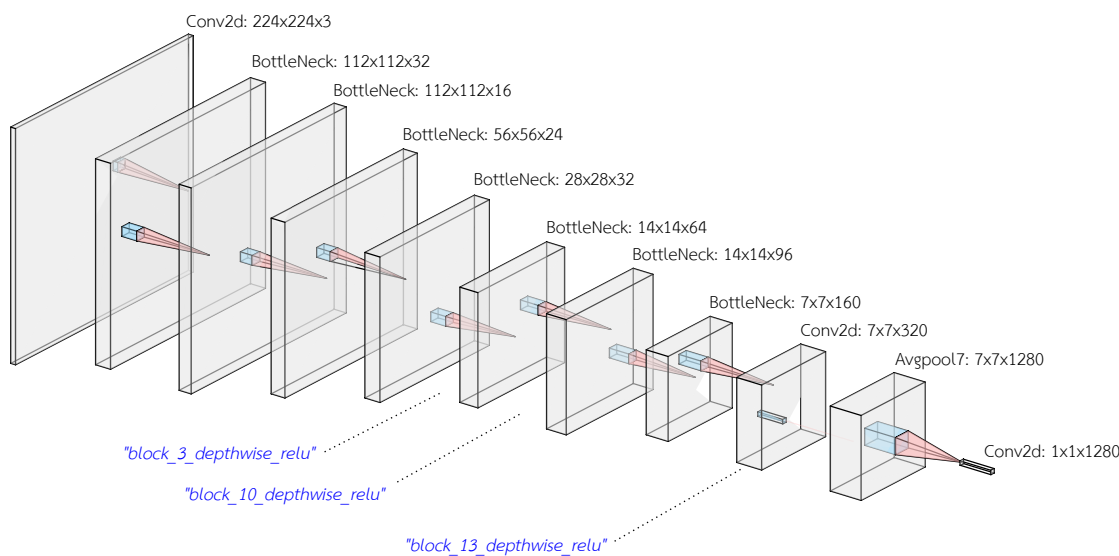
แม้ว่าปัญหาการตรวจหาวัตถุขนาดเล็กในภาพของโครงสร้างแบบ YOLO ถูกพัฒนาให้มีความแม่นยำมากขึ้นในการพัฒนาขั้นต่อมา เช่น YOLO9000 [65] ซึ่งมีการใช้กรอบสี่เหลี่ยมที่เป็นไปได้เริ่มต้น (Anchor Box) เช่นเดียวกับ Faster-RCNN และเปลี่ยนรูปแบบการคำนวณกรอบสี่เหลี่ยมผลลัพธ์ใหม่ ซึ่งทำนายโดยตรงจาก Fully Connected Layer ทำให้เน็ตเวิร์กจะทำนายผลของการตรวจหาบนพีเจอร์เลเยอร์

ที่มีขนาดเล็ก (Resolution) ซึ่งยังคงให้ผลการตรวจหาวัตถุที่มีขนาดเล็กด้อยกว่าเน็ตเวิร์กแบบสองส่วน เช่นเดียวกับโครงสร้างแบบ SSD ที่ยังคงให้ผลการตรวจหาวัตถุขนาดเล็กในภาพด้อยกว่า Faster-RCNN ในกรณีที่ใช้ภาพอินพุตที่ขนาดเท่ากัน

ดังนั้นวิทยานิพนธ์นี้จึงเลือกใช้โครงข่ายประสาทเชิงลึกแบบคอนโวลูชันสำหรับตรวจหาวัตถุที่มีโครงสร้างแบบสองส่วนชนิด Faster-RCNN สำหรับการตรวจหาวัตถุและถ้วยรองน้ำยาง เนื่องจากขนาดของกรอบสี่เหลี่ยมล้อมรอบวัตถุและถ้วยรองน้ำยางในภาพมีขนาดเล็กเมื่อเทียบกับขนาดภาพ และการพัฒนาตัวตรวจหาวัตถุแบบโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันเพียงเพื่อต้องการเปรียบเทียบเฉพาะความแม่นยำกับวิธีตรวจหาวัตถุในข้อย่อย 2.3.3

2.3.4.3 การกำหนด Faster R-CNN ร่วมกับ MobileNetV2

โครงสร้าง Faster-RCNN ประกอบไปด้วยโครงข่ายประสาทแบบคอนโวลูชันย่อยจำนวน 2 เน็ตเวิร์ก ซึ่งใช้งานเลเยอร์คอนโวลูชันร่วมกันบางส่วน (Shared Layers) โดยเน็ตเวิร์กย่อยหนึ่งทำหน้าที่ทำนายพื้นที่ที่คาดว่า มีวัตถุเป้าหมายอยู่ภายใน เรียกว่า “เน็ตเวิร์กทำนายพื้นที่ที่สนใจ” (“Region Proposal Network”, คำย่อ RPN) และอีกเน็ตเวิร์กย่อยหนึ่งทำหน้าที่ทำนายชนิดวัตถุภายในพื้นที่จาก RPN เรียกว่า “เน็ตเวิร์กตรวจหาวัตถุ” (Detection Network)



รูปที่ 2.33: แผนภาพแสดงขนาดของข้อมูล (รูปภาพที่อยู่ในรูปแบบเมทริกซ์สามมิติ) หรือเรียกว่าเทนเซอร์ (Tensors) ภายในเน็ตเวิร์ก MobileNetV2 [69] โดยขนาดของกล่องสี่เหลี่ยมแสดงถึงขนาดเทนเซอร์ (กว้าง \times ยาว \times ลึก) ซึ่งภายในและระหว่างกล่องมีชั้นประมวลผลคอนโวลูชันแบบ Inverted Residuals และ Linear Bottlenecks หรือคอนโวลูชันเลเยอร์แบบปกติ (Conv2d)

โครงสร้างอย่างง่ายของ MobileNetV2 แสดงในรูปที่ 2.33 ซึ่งรูปทรงกล่องสี่เหลี่ยมแสดงถึงขนาดเทนเซอร์ (รูปภาพที่ผ่านการคอนโวลูชันในแต่ละชั้นของเน็ตเวิร์ก) ที่เปลี่ยนแปลงไปภายในเน็ตเวิร์ก การ

กำหนดตำแหน่งของเลเยอร์ที่ใช้ร่วมกันของสองเน็ตเวิร์กย่อยให้ทำงานร่วมกับ MobileNetV2 มีรายละเอียดดังนี้

การกำหนดเน็ตเวิร์กทำนายพื้นที่ที่สนใจ (RPN) โดยทั่วไป RPN สามารถสร้างได้จากเน็ตเวิร์กที่มีเฉพาะเลเยอร์คอนโวลูชัน (Fully Convolutional Network, คำย่อ FCN) โดยมีส่วนของเอาต์พุตเลเยอร์ 2 ประเภท ประเภทหนึ่งเป็นเลเยอร์ที่ทำหน้าที่ให้ค่าตัวเลข 4 จำนวน (Regression Layer) ซึ่งกำกับตำแหน่งและขนาดของกรอบสี่เหลี่ยมล้อมรอบพื้นที่ที่สนใจในพิกัดของเทนเซอร์ (ROI Coordinates) และอีกประเภทหนึ่งเป็นเลเยอร์สำหรับทำนายชนิดวัตถุภายในกรอบสี่เหลี่ยมที่สร้างจาก RPN (Classification Layer หรือ Softmax) โดยจะทำนายผลเพียง 2 คำตอบคือ มีวัตถุหรือไม่มีวัตถุ (เป็นพื้นหลัง) ดังแสดงในกรอบสี่เหลี่ยมเส้นประในรูปที่ 2.34 โดยเอาต์พุตพื้นที่ที่สนใจ (ROI) จะถูกป้อนไปยังเลเยอร์พูลิงพื้นที่ที่สนใจ (ROI Pooling) ซึ่งเป็นเลเยอร์ที่อยู่ในลำดับสูงกว่าต่อไป

“ROI Pooling” เป็นเลเยอร์เฉพาะของ Faster-RCNN ซึ่งรับอินพุตตัวเลข 4 จำนวนที่ระบุตำแหน่งพื้นที่ที่สนใจ และรับอินพุตเป็นเทนเซอร์จากเลเยอร์ติดกันในอันดับต่ำกว่า เลเยอร์พูลิงพื้นที่ที่สนใจจะทำหน้าที่แปลงพิกัดของกรอบสี่เหลี่ยมจากเทนเซอร์อินพุตให้มีขนาด (Spatial Dimension) ตรงกับขนาดเทนเซอร์ฝั่งเอาต์พุตด้วยการพูลิง (Pooling) เฉพาะพื้นที่ภายในกรอบสี่เหลี่ยมอินพุต ซึ่งพื้นที่ดังกล่าวจะเป็นเทนเซอร์ที่มีขนาดและความลึก ณ ลำดับเลเยอร์ภายในเน็ตเวิร์กขณะนั้น

การสร้าง RPN ในรูปแบบการเรียนรู้แบบถ่ายโอนด้วยเลเยอร์จาก MobileNetV2 ทำได้โดยกำหนดให้เลเยอร์คอนโวลูชันส่วนต้นของ RPN (Low-level Convolutional Layers) เป็นเลเยอร์ที่มีค่าน้ำหนักจาก MobileNetV2 เดิม จากนั้นเพิ่มเลเยอร์คอนโวลูชันส่วนปลายขึ้นใหม่ (High-Level Convolutional Layers) และกำหนดเอาต์พุตเลเยอร์ขึ้นใหม่

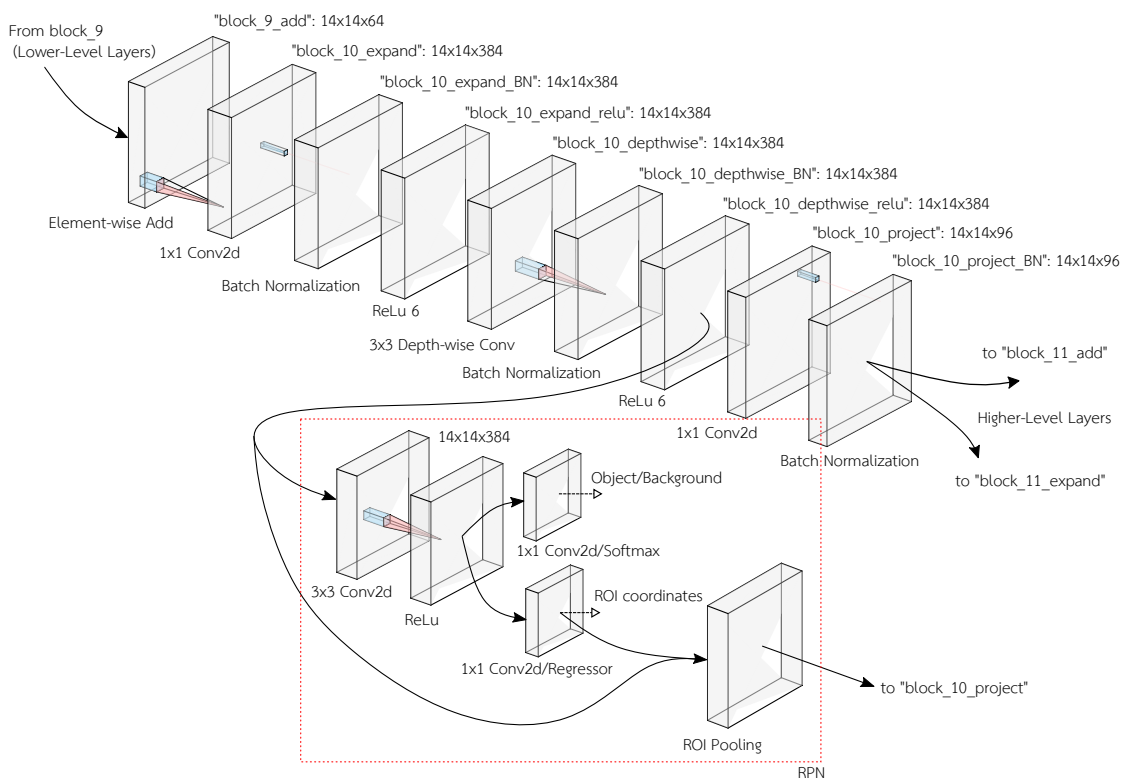
การจัดเรียงเลเยอร์ภายใน MobileNetV2 มีโครงสร้างเฉพาะแบบบล็อก (Block) ซึ่งแต่ละบล็อกประกอบไปด้วยเลเยอร์ที่จัดเรียกกันแบบ Inverted Residuals กับ Linear BottleNecks โดยในแต่ละบล็อก ขนาดความลึกของเทนเซอร์จะถูกขยายด้วยโครงสร้างเลเยอร์แบบ Inverted Residual เพื่อดำเนินการคอนโวลูชันในแนวความลึก (Depth-wise Convolution) และในส่วนปลายของบล็อก เทนเซอร์จะถูกคอนโวลูชันให้มีขนาดความกว้างกับความยาว (Spatial Dimension) ลดลงแต่มีความลึกที่มากขึ้น เมื่อเปรียบเทียบขนาดจากฝั่งอินพุตของบล็อก (Linear BottleNecks) เป็นผลให้หากแทรกเลเยอร์ส่วนปลายของ RPN ไปยังตำแหน่งส่วนกลางหรือส่วนท้ายของบล็อก จะส่งผลต่อปริมาณฟิลเตอร์คงเดิมใน MobileNetV2 และปริมาณพารามิเตอร์ของคอร์เนลใน MobileNetV2 ที่มีผลต่อความสามารถในการเรียนรู้ (Learnable Parameters) โดยหากเพิ่มเลเยอร์คอนโวลูชันใหม่ในตำแหน่งกลางบล็อก จะทำให้จำนวนฟิลเตอร์เท่ากับความลึกของเมทริกซ์สามมิติ (Tensor) ณ ตำแหน่งนั้น ซึ่งจะมีจำนวนมากกว่า เมื่อเปรียบเทียบกับกรณีการเพิ่มในตำแหน่งปลายของบล็อก ส่งผลให้เน็ตเวิร์กโดยรวมมีโอกาสเรียนรู้ปัญหาใหม่เฉพาะชุดข้อมูลได้ดีกว่า เพราะมีจำนวนพารามิเตอร์ใหม่เกิดขึ้นจำนวนมาก

การแทรกกลุ่มของเลเยอร์ RPN ให้กับ MobileNetV2 สามารถเลือกตำแหน่งแทรกในเลเยอร์อันดับใดๆ ภายในเน็ตเวิร์กได้อย่างอิสระ ยกตัวอย่างเช่นระหว่างเลเยอร์ “block_10_depthwise_relu” กับ “block_10_project” โดยเอาต์พุตของเลเยอร์พูลิงพื้นที่ที่สนใจ จะถูกป้อนให้กับเลเยอร์ “block_10_project”

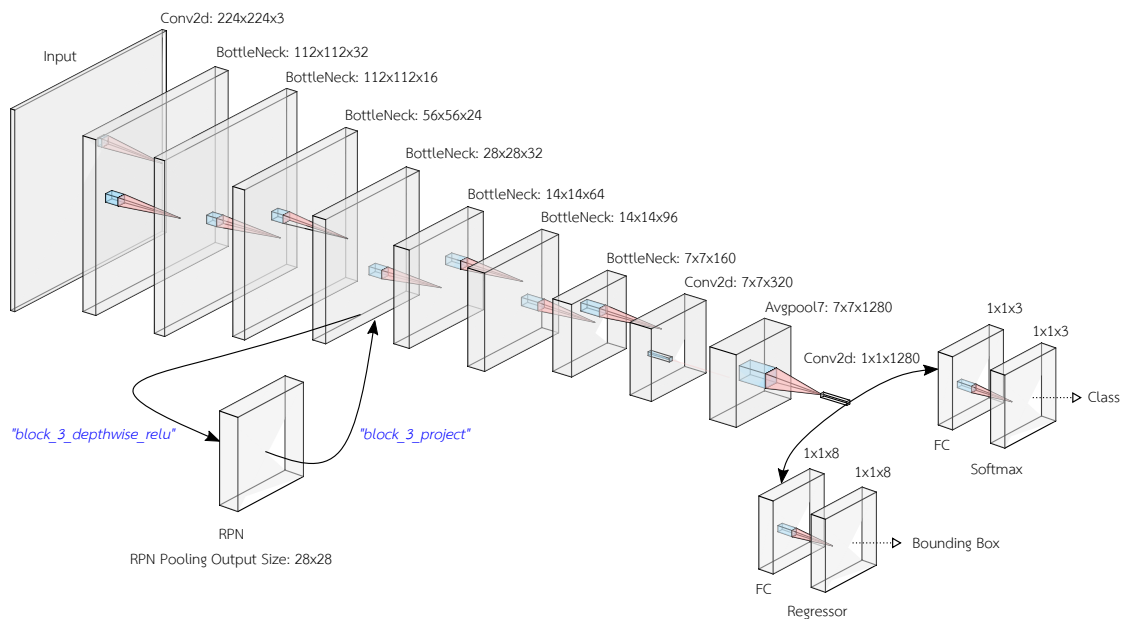
ตามรูปที่ 2.34 ซึ่งทำให้ RPN จะประมวลผลเทนเซอร์ที่มีขนาดเท่ากับ $14 \times 14 \times 384$ ในกรณีที่มีอินพุตภาพเท่ากับ 224×224 พิกเซล หรือสามารถแทรกกลุ่มของเลเยอร์ RPN ในเลเยอร์ลำดับอื่นๆ ภายใน MobileNetV2 โดยหากเลือกลำดับเลเยอร์ที่มีขนาดกว้างและสูงของเทนเซอร์ใหญ่กว่า จะส่งผลให้ RPN สามารถตรวจหาตำแหน่งของวัตถุขนาดเล็กได้ดีกว่าตำแหน่งที่มีขนาดเทนเซอร์เล็กกว่า เนื่องจากการคอนโวลูชัน ขนาดความกว้างและความสูงของเทนเซอร์จะหดสั้นลงในขณะที่ความลึกจะเพิ่มสูงขึ้น อาจส่งผลให้พื้นที่ในเทนเซอร์ที่มีวัตถุขนาดเล็กสูญหายไปก่อนในชั้นเลเยอร์อันดับสูงขึ้น (Loss of Information) ตัวอย่างการกำหนดเลเยอร์ส่วนปลายของ RPN ในเลเยอร์ลำดับต่างๆ ใน MobileNetV2 ที่มีขนาดเทนเซอร์เท่ากับ 28×28 , 14×14 และ 7×7 แสดงในรูปที่ 2.35, 2.36, และ 2.38 ตามลำดับ

ตัวอย่างการเพิ่มกลุ่มของเลเยอร์ส่วนปลายของ RPN ที่ตำแหน่งกลาง Block_10 ในเลเยอร์ “block_10_depthwise_relu” ที่มีขนาดเทนเซอร์เท่ากับ $14 \times 14 \times 384$ แสดงในรูปที่ 2.34 ประกอบไปด้วยเลเยอร์ใหม่ที่น่าหนักยังไม่ผ่านการเรียนรู้ใดๆ เพื่อการเรียนรู้วัตถุชนิดรอยกритและถ้วยรองน้ำยาง ได้แก่ เลเยอร์คอนโวลูชันขนาด 3×3 (3×3 Conv2d) ตามด้วย Rectified Linear Unit (ReLU) และเลเยอร์คอนโวลูชันขนาด 1×1 (Conv2d) เพื่อลดขนาดความลึกของเทนเซอร์ให้กับแต่ละเอาต์พุตเลเยอร์ โดยมีเอาต์พุตเลเยอร์ ได้แก่ เลเยอร์สำหรับทำนายตำแหน่งกรอบสี่เหลี่ยมล้อมรอบ (Bounding Box Regressor), เลเยอร์สำหรับบ่งชี้ว่ากรอบสี่เหลี่ยมมีวัตถุอยู่หรือไม่ (Softmax Layer) และเลเยอร์ ROI Pooling ซึ่งทำหน้าที่แมปพื้นที่ที่สนใจไปยังขนาดเลเยอร์ลำดับถัดไปในเน็ตเวิร์ก

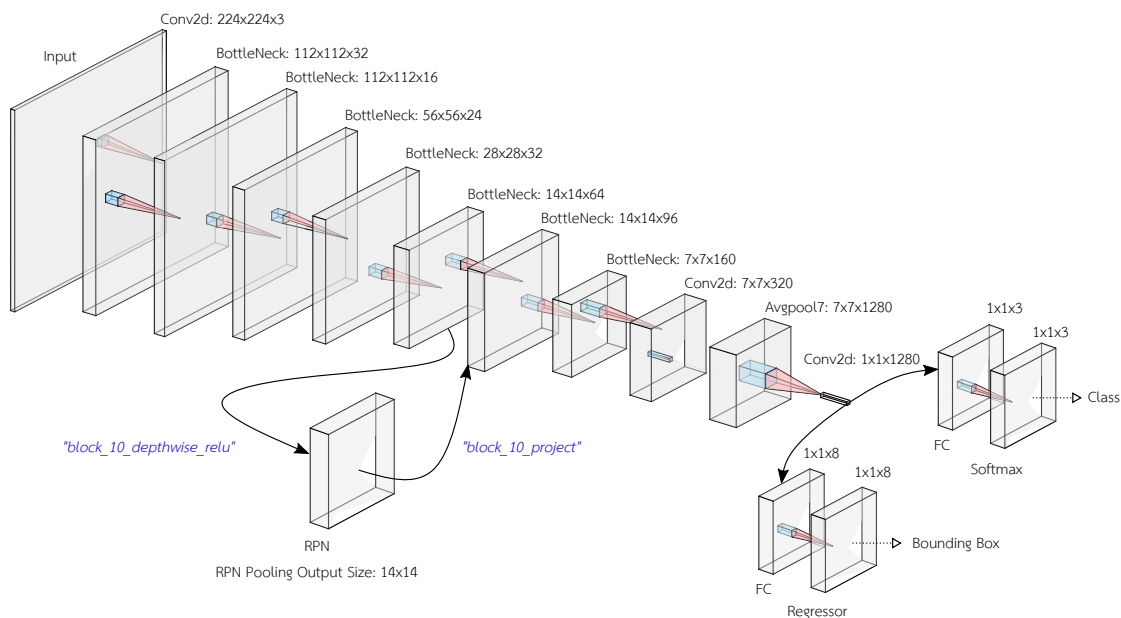
ทั้งนี้การเลือกตำแหน่งเลเยอร์ใน MobileNetV2 เพื่อแทรกกลุ่มของเลเยอร์ส่วนปลายของ RPN สำหรับการตรวจหารอยกритและถ้วยรองน้ำยาง จะต้องพิจารณาขนาดของกรอบสี่เหลี่ยมล้อมรอบผลลัพธ์ที่จะถูกผลิตโดย RPN ซึ่งสัมพันธ์กับขนาดของเทนเซอร์ (Spatial Dimension) ณ ตำแหน่งเลเยอร์ที่แทรกนั้น และสัมพันธ์กับขนาดของวัตถุที่สนใจซึ่งคือรอยกритและถ้วยรองน้ำยาง เนื่องจากขนาดของรอยกритและถ้วยรองน้ำยางมีขนาดเล็กเมื่อเทียบกับขนาดของภาพ หากเลือกแทรกกลุ่มของเลเยอร์ส่วนปลายของ RPN ไปยังตำแหน่งบล็อกที่มีขนาดความกว้างและสูงของเทนเซอร์ที่เล็กเกินไป จะส่งผลให้ RPN อาจไม่สามารถระบุตำแหน่งของกรอบสี่เหลี่ยมขนาดเล็กในเทนเซอร์นั้นๆ ได้ ทั้งนี้การทดลองได้ปรับเปลี่ยนตำแหน่งการแทรก RPN ใน MobileNetV2 เพื่อสังเกตความแม่นยำในการตรวจหารอยกритและถ้วยรองน้ำยาง



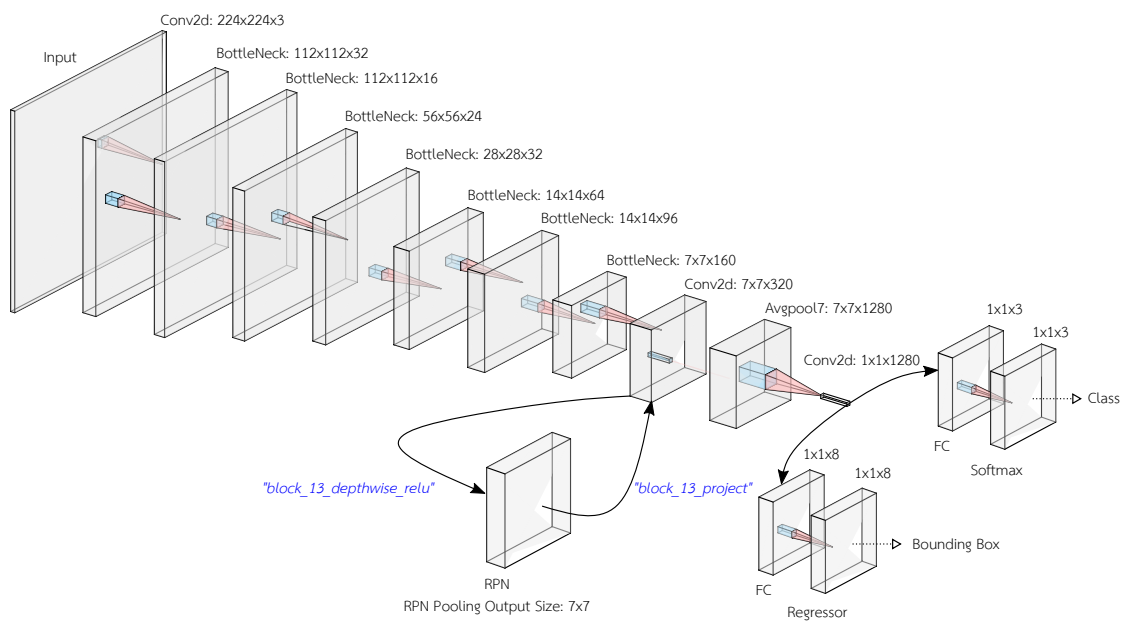
รูปที่ 2.34: โครงสร้างเลเยอร์แบบ Inverted Residuals และ Linear Bottlenecks ของ MobileNetV2 จำนวน 1 บล็อก แสดง ณ ตำแหน่ง “block_10” ซึ่งรับอินพุตจาก “block_9” และแสดงเลเยอร์ส่วนปลายของ RPN ได้แก่ เลเยอร์คอนโวลูชัน (Conv2d), ReLU, เลเยอร์ทำนายประเภทวัตถุ (Softmax), เลเยอร์ทำนายตำแหน่งกรอบสี่เหลี่ยม (Regressor) และเลเยอร์พูลิงพื้นที่ที่สนใจ (ROI Pooling) สำหรับคำนวณหาเทนเซอร์เฉพาะในพื้นที่กรอบสี่เหลี่ยมที่สนใจ เพื่อป้อนให้ชั้นของเน็ตเวิร์กในลำดับสูงกว่าใน “block_11”



รูปที่ 2.35: ภาพรวมเน็ตเวิร์กตรวจหาตำแหน่งวัตถุแบบ Faster-RCNN ที่สร้างบน MobileNetV2 สำหรับตรวจหากรอยกริดและถ้วยร่อนน้ำยาง โดย RPN ทำงานบนพีเจอร์แมปขนาดกว้าง 28x28



รูปที่ 2.36: ภาพรวมเน็ตเวิร์กตรวจหาตำแหน่งวัตถุแบบ Faster-RCNN ที่สร้างบน MobileNetV2 สำหรับตรวจหากรอยกริดและถ้วยร่อนน้ำยาง โดย RPN ทำงานบนพีเจอร์แมปขนาดกว้าง 14x14



รูปที่ 2.37: ภาพรวมเน็ตเวิร์กตรวจหาตำแหน่งวัตถุแบบ Faster-RCNN ที่สร้างบน MobileNetV2 สำหรับตรวจหารอยกริดและถ้วยร่อนน้ำยาง โดย RPN ทำงานบนพีเจอร์แมปขนาดกว้าง 7x7

เน็ตเวิร์กตรวจหาวัตถุ (Detection Network) เลเยอร์ใหม่ส่วนปลายของ RPN ซึ่งมีเลเยอร์พูลิงพื้นที่ที่สนใจป้อนเอาต์พุตผ่านต่อไปยังเลเยอร์ลำดับสูงกว่าที่เหลือใน MobileNetV2 เพื่อให้พีเจอร์ระดับสูง (High-Level Features) ของแต่ละพื้นที่ที่สนใจจาก RPN จะถูกสกัด (Extraction) และรู้จำรอยกริดและถ้วยรองน้ำยางในเลเยอร์ที่เพิ่มใหม่ในส่วนปลายของเน็ตเวิร์ก ดังนั้นสามารถสร้างเน็ตเวิร์กตรวจหาวัตถุ (Detection Network) ได้โดยทำการเพิ่มกลุ่มของเลเยอร์ในเลเยอร์ลำดับบนสุดของเน็ตเวิร์ก MobileNetV2 โดยแยกเอาต์พุตออกเป็น 2 กลุ่ม (Two-Tails Output) ดังนี้

1. เพิ่มเลเยอร์แบบรวมอินพุต (Fully Connected Layer, คำย่อ FC) ซึ่งจะแมปขนาดของเทนเซอร์จากเลเยอร์คอนโวลูชันชั้นปลายสุดของเน็ตเวิร์ก MobileNetV2 ไปยังจำนวนเอาต์พุตที่กำหนด ซึ่งเลเยอร์กลุ่มที่หนึ่งจะมีจำนวน 3 เอาต์พุตสำหรับจำแนกพื้นที่ที่สนใจว่าเป็นรอยกริดหรือถ้วยรองน้ำยางหรือพื้นหลัง ซึ่งอาศัยเลเยอร์ Softmax โดยมีฟังก์ชันคำนวณความผิดพลาดแบบครอสเอนโทรปี (Cross Entropy Loss) สำหรับการสอนเน็ตเวิร์ก
2. เพิ่มเลเยอร์แบบรวมอินพุต (FC) ที่มีจำนวน 8 เอาต์พุต เพื่อใช้ทำนายกรอบสี่เหลี่ยมของรอยกริดและถ้วยรองน้ำยาง (ประเภทวัตถุละ 4 จำนวน) ซึ่งอาศัยเลเยอร์ประมาณค่าเชิงตัวเลข (Regressor Layer) โดยมีฟังก์ชันคำนวณความผิดพลาดแบบกำหนดเงื่อนไข (Smooth-L1 Loss) สำหรับการสอนเน็ตเวิร์ก

2.3.4.4 การสอนเน็ตเวิร์กแบบปรับละเอียด (Fine Tuning)

Faster-RCNN ที่สร้างบน MobileNetV2 ได้เพิ่มกลุ่มเลเยอร์ส่วนปลายของ RPN และเปลี่ยนเอาต์พุตเลเยอร์ลำดับสูงสุดของเน็ตเวิร์กเพื่อใช้ทำนายรอยกริดและถ้วยรองน้ำยาง ทำให้เลเยอร์ในลำดับตั้งแต่ชั้นอินพุต (Input Layer) จนถึงเลเยอร์ที่ให้ค่าเอาต์พุตเข้าสู่เลเยอร์ส่วนปลายของ RPN จะถูกใช้งานร่วมกันระหว่างเน็ตเวิร์กทำนายพื้นที่ที่สนใจ และเน็ตเวิร์กตรวจหาวัตถุ

การสอนเน็ตเวิร์กจะเป็นไปตามวิธีที่งาน Faster-RCNN ต้นฉบับนำเสนอ ซึ่งจะแบ่งการสอนเน็ตเวิร์กย่อยทั้งสองแยกกันด้วยวิธีการสอนแบบสลับ 4 ขั้นตอน (4-Step Alternate Training) [66] ได้แก่

1. สอนเน็ตเวิร์กทำนายพื้นที่ (RPN Training) โดยแต่ละเลเยอร์ใช้ค่าน้ำหนักเดิมของ MobileNetV2 ในขณะที่เลเยอร์ที่เพิ่มใหม่มีค่าน้ำหนักเริ่มต้นแบบสุ่ม และกำหนดอัตราการเรียนรู้ (Learning Rate) ของเลเยอร์ที่เพิ่มใหม่มีค่ามากกว่าเลเยอร์ MobileNetV2 เดิม 10 เท่า
2. สอนเน็ตเวิร์กตรวจหาวัตถุโดยใช้พื้นที่จากข้อ 1 (RCNN Training) ให้และแต่ละเลเยอร์ใช้ค่าน้ำหนักเดิมของ MobileNetV2 ในขณะที่เลเยอร์ที่เพิ่มใหม่มีค่าน้ำหนักเริ่มต้นแบบสุ่ม และกำหนดอัตราการเรียนรู้ของเลเยอร์ที่เพิ่มใหม่มีค่ามากกว่าเลเยอร์ MobileNetV2 เดิม 10 เท่า
3. สอนเน็ตเวิร์กทำนายพื้นที่อีกครั้ง โดยใช้ค่าน้ำหนักเลเยอร์จากข้อ 2 และไม่ให้เลเยอร์ที่ใช้งานร่วมกันระหว่างสองเน็ตเวิร์กมีการปรับเปลี่ยนค่าน้ำหนัก
4. สอนเน็ตเวิร์กตรวจหาวัตถุอีกครั้ง โดยไม่ให้เลเยอร์ที่ใช้งานร่วมกันของสองเน็ตเวิร์กมีการปรับเปลี่ยนค่าน้ำหนัก

และได้ปรับเปลี่ยนรายละเอียดของวิธีการสอนเน็ตเวิร์กเพิ่มเติมดังนี้

การสอนเน็ตเวิร์กทำนายพื้นที่ที่สนใจ (RPN Training) ในขั้นตอนการสอนที่ 1 และ 3 โดยทั่วไปในขั้นตอนการสอนเน็ตเวิร์กการกำหนดพื้นที่ที่สนใจเริ่มต้นให้กับ RPN จะกำหนดกรอบสี่เหลี่ยมล้อมรอบเริ่มต้น (“Anchors”) ทั้งทั้งภาพในหลายอัตราส่วน (Aspect Ratio) และขนาด (Scale) โดยกำหนดกรอบสี่เหลี่ยมที่มีอัตราส่วนซ้อนทับ (Intersection Over Union, IoU) กับกรอบสี่เหลี่ยมล้อมรอบอ้างอิงของวัตถุ ที่มีค่าในช่วง 0.0-0.5 สำหรับใช้เป็นพื้นที่พื้นหลังสำหรับการสอน (Negative Training Samples) และอัตราส่วนซ้อนทับในช่วง 0.5-1.0 สำหรับใช้เป็นพื้นที่ของวัตถุสำหรับการสอน (Positive Training Samples) จากนั้นเลือกกรอบสี่เหลี่ยมสำหรับสอนเน็ตเวิร์กโดยการสุ่มจากรอบที่กำหนดขึ้นข้างต้นทั้งหมดภายในภาพตามจำนวนที่ต้องการด้วยอัตราส่วนจำนวนกรอบพื้นหลังต่อจำนวนกรอบวัตถุ 1:1 ไปจนถึง 3:1

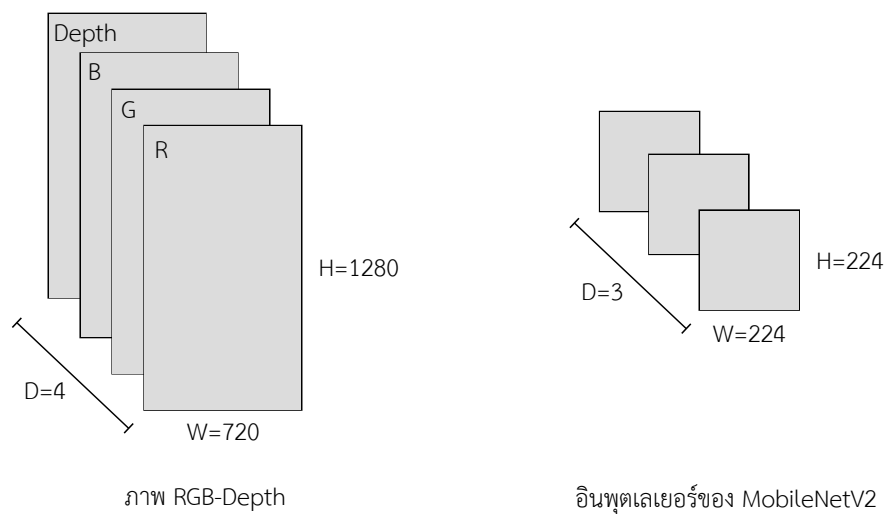
วิธีการสุ่มกรอบสี่เหลี่ยมเหมาะสำหรับข้อมูลรูปภาพที่มีวัตถุหลากหลายขนาด กระจายอยู่ทั่วบริเวณของภาพ และมีปริมาณชนิดวัตถุที่ไม่แน่นอนในแต่ละภาพ แต่สำหรับภาพถ่ายต้นยางพาราจากกล้องถ่ายภาพมุมกว้าง วิธีดังกล่าวอาจทำให้เกิดปัญหาสุ่มได้กรอบ (Anchors) ที่เป็นพื้นหลังมากกว่ากรอบสี่เหลี่ยมรอยกริดหรือถ้วยรองน้ำยาง เนื่องจากการกำหนดกรอบสี่เหลี่ยมอ้างอิงของรอยกริดและถ้วยรองน้ำยาง ในแต่ละภาพมีจำนวนกรอบสี่เหลี่ยมอ้างอิงรวมกันมากที่สุดเพียง 2 กรอบเท่านั้น และเป็นสัดส่วนพื้นที่ที่น้อยเมื่อเทียบกับพื้นที่พื้นหลัง ดังนั้นจึงบังคับจำนวนของกรอบสี่เหลี่ยมสำหรับการสอนให้มีจำนวนของกรอบพื้นหลังและกรอบของวัตถุเท่ากันในแต่ละภาพ รวม 200 กรอบต่อภาพ ประกอบไปด้วยกรอบพื้นหลัง 100 กรอบ รอยกริดและถ้วยรองน้ำยางรวมกัน 100 กรอบ และเปลี่ยนการกำหนดอัตราส่วนซ้อนทับระหว่างกรอบสี่เหลี่ยมอ้างอิง (Ground Truth Bounding Box) กับกรอบสี่เหลี่ยมเริ่มต้นโดยอัตราส่วนในช่วง 0.0-0.3 จะถูกใช้เพื่อระบุเป็นพื้นที่พื้นหลัง และกรอบที่มีอัตราส่วนการซ้อนทับในช่วง 0.7-1.0 ใช้เพื่อระบุเป็นพื้นที่วัตถุ ส่งผลให้กรอบสี่เหลี่ยมเริ่มต้นที่มีอัตราส่วนซ้อนทับกับกรอบสี่เหลี่ยมอ้างอิงในช่วง 0.3-0.7 ไม่ถูกนำไปใช้สอนเน็ตเวิร์ก เป็นผลให้เน็ตเวิร์กสามารถให้คะแนนการตรวจหากรอบสี่เหลี่ยมของวัตถุได้สูงกว่ากรอบอื่นๆ ซึ่งเป็นผลดีกับปัญหาการตรวจหาวัตถุเพียงจำนวน 1 กรอบต่อชนิดต่อภาพ

การสอนเน็ตเวิร์กตรวจหาวัตถุ (Detection Network Training) เน็ตเวิร์กตรวจหาวัตถุจะใช้เลเยอร์ทั้งหมดของ MobileNetV2 โดยมีเลเยอร์ในส่วนต้นใช้งานร่วมกันกับ RPN และตั้งแต่อินพุตเลเยอร์จนถึงพีเออร์เลเยอร์ของ RPN ทำให้เลเยอร์ถัดจากเลเยอร์พูลิงจนถึงเลเยอร์ชั้นบนสุดในส่วนปลายไม่ถูกใช้งานร่วมกัน จึงกำหนดให้เลเยอร์เดิมของ MobileNetV2 ที่ไม่ใช้ร่วมกับ RPN มีอัตราการเรียนรู้ที่น้อยกว่าเลเยอร์ที่เพิ่มใหม่ส่วนปลายของเน็ตเวิร์กตรวจหาวัตถุ 10 เท่า ในการสอนขั้นตอนที่ 4

2.3.4.5 การเตรียมข้อมูลภาพสำหรับสอนและทดสอบเน็ตเวิร์ก

โครงข่ายประสาทเชิงลึกแบบคอนโวลูชันสำหรับการตรวจหาวัตถุที่มีเผยแพร่โดยทั่วไปต่างเรียนรู้และทดสอบด้วยฐานข้อมูลภาพถ่ายสี RGB ที่มีชนิดของวัตถุจำนวนมาก จึงทำให้เน็ตเวิร์กเหล่านั้นต่างเรียนรู้ฟีเจอร์ต่างๆ บนพื้นฐานของข้อมูลภาพสีและจำนวนช่องสัญญาณอินพุตของเน็ตเวิร์ก (Input Channel หรือความลึกของเมทริกซ์สามมิติ, อักษรย่อ D) เท่ากับ 3 ตามช่องสัญญาณภาพ แดง-เขียว-น้ำเงิน (RGB) เช่นเดียวกับ MobileNetV2 ที่มีขนาดอินพุตเท่ากับ $224 \times 224 \times 3$: กว้าง \times สูง \times ลึก (W \times H \times D)

ภาพจากกล้องถ่ายภาพมุมกว้างมีทั้งภาพถ่ายสีและภาพถ่ายความลึก (RGB-Depth) มีขนาดเท่ากับ $720 \times 1280 \times 4$: W \times H \times D จึงต้องแปลงข้อมูลให้มีขนาดตรงกับขนาดอินพุตของเน็ตเวิร์ก ซึ่งโดยทั่วไปภาพสี RGB จะถูกเปลี่ยนขนาด (Resize) ให้มีขนาดเท่ากับอินพุตในด้านกว้าง (W) ด้านสูง (H) และมีจำนวน 3 ช่องสัญญาณ (D) แต่ภาพถ่ายต้นแบบพาราชนิด RGB-Depth มีขนาดช่องสัญญาณมากกว่าขนาดของอินพุตดังแสดงในรูปที่ ทำให้การใช้งานเน็ตเวิร์กเพื่อประมวลผลภาพ RGB-Depth จึงต้องการการประมวลผลภาพเพิ่มเติมก่อนอินพุตภาพให้กับเน็ตเวิร์ก



รูปที่ 2.38: รูปวาดเปรียบเทียบขนาดของภาพถ่ายสีและภาพถ่ายความลึก (RGB-Depth) และขนาดของอินพุตเลเยอร์ของ MobileNetV2

การจัดสรรภาพอินพุตและกลุ่มทดลอง ภาพย่อยในแต่ละช่องสัญญาณของภาพถ่ายสี RGB เป็นภาพโทนเทา (Grayscale) ที่มีขนาดพิกเซล 8 บิต ซึ่งเป็นภาพประเภทเดียวกับภาพถ่ายความลึก ดังนั้นภาพถ่ายต้นแบบพาราแบบ RGB-D จึงอยู่ในรูปแบบภาพโทนเทาจำนวน 4 ช่องสัญญาณที่มีขนาดพิกเซล (Bit Depth) เท่ากัน

หากให้เน็ตเวิร์กรองรับภาพขนาด 4 ช่องสัญญาณด้วยการขยายจำนวนช่องสัญญาณอินพุตจาก 3 เป็น 4 ให้กับเน็ตเวิร์กสามารถทำได้ แต่อาจทำให้เน็ตเวิร์กต้องการการเรียนรู้ใหม่ในจำนวนครั้งที่ยากพอ

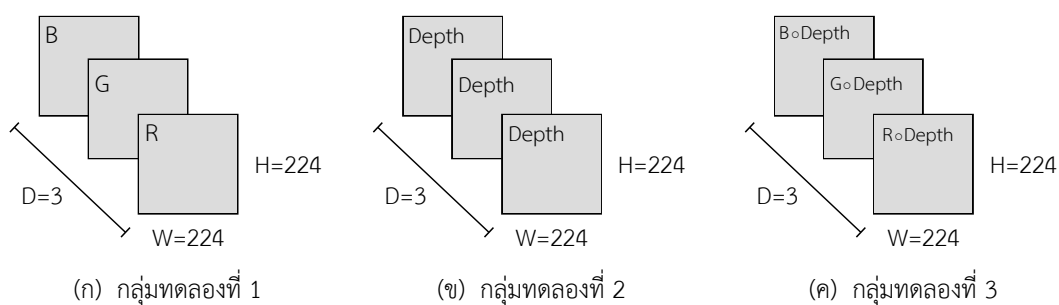
หรือเรียนรู้ใหม่ทั้งหมด เพื่อที่จะทำให้ค่าพารามิเตอร์ที่เพิ่มขึ้นในเลเยอร์ลำดับต้นมีค่าที่เป็นประโยชน์ต่อการตรวจหาวัตถุ วิธีดังกล่าวนี้อาจส่งผลกระทบต่อค่าน้ำหนักเดิมของเน็ตเวิร์กที่ถูกสอนมาแล้วบนฐานข้อมูลขนาดใหญ่ ซึ่งส่งผลกระทบต่อความสามารถในการสลัดพีเจอร์พื้นฐาน เนื่องจากการเรียนรู้โดยรวมของเน็ตเวิร์กในแต่ละเลเยอร์จะอาศัยอัตราการเปลี่ยนแปลงค่าน้ำหนักที่ส่งมาจากเลเยอร์ลำดับสูงกว่า (Backward Propagation) อีกวิธีหนึ่งคือการประมวลผลเพิ่มเติมเพื่อลดขนาดด้านความลึก (D) ให้ตรงกับขนาดอินพุตของเน็ตเวิร์ก

ดังนั้นวิทยานิพนธ์นี้เลือกวิธีประมวลผลภาพก่อนจะอินพุตภาพให้กับเน็ตเวิร์ก โดยภาพจากกล้องถ่ายภาพมุมกว้างจำนวน 500 ชุดภาพ ที่มีขนาดกว้าง 720 พิกเซล สูง 1280 พิกเซล จะถูกปรับขนาดให้กว้างเท่ากับ 224 พิกเซล และสูง 224 พิกเซล (ไม่รักษาอัตราส่วนความกว้างต่อความยาวเดิม) ตามขนาดอินพุตมาตรฐานของ MobileNetV2 และกรอบสี่เหลี่ยมอ้างอิงรอยกริดและถัวยรอนน้ำยางจะถูกย่อขนาดให้ตรงกับตำแหน่งของวัตถุในภาพขนาด 224x224 พิกเซลเช่นกัน จากนั้นจัดกลุ่มการทดลองตามช่องสัญญาณภาพเป็น 3 กลุ่มดังนี้

กลุ่มทดลองที่ 1 ภาพถ่ายสี RGB ประกอบด้วยช่องสัญญาณภาพสีแดง (R), สีเขียว (G), และสีน้ำเงิน (B) ตามรูปที่ 2.39 (ก)

กลุ่มทดลองที่ 2 ภาพถ่ายความลึกทั้ง 3 ช่องสัญญาณ ตามรูปที่ 2.39 (ข)

กลุ่มทดลองที่ 3 ภาพถ่ายสี RGB โดยแต่ละช่องสัญญาณภาพสีจะคูณด้วยภาพถ่ายความลึกแบบพิกเซลต่อพิกเซลที่ตรงตำแหน่งกัน ตามรูปที่ 2.39 (ค)

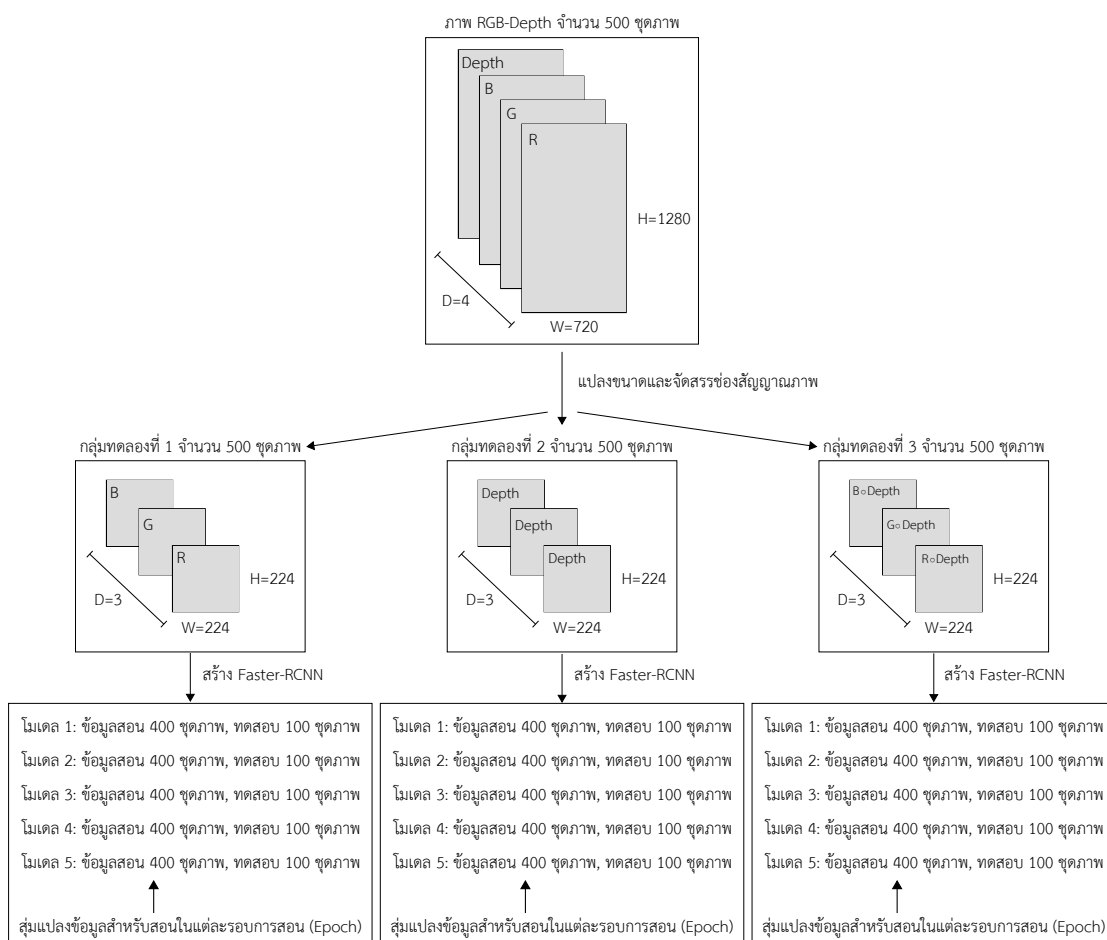


รูปที่ 2.39: กลุ่มทดลองจัดสรรภาพ RGB-Depth ให้ตรงกับขนาดอินพุตเลเยอร์ของ MobileNetV2

การแบ่งกลุ่มข้อมูลในแต่ละกลุ่มทดลองเพื่อสอนและทดสอบเน็ตเวิร์ก ชุดข้อมูลภาพถ่ายจากกล้องถ่ายภาพมุมกว้างจำนวน 500 ชุดภาพ เพื่อประเมินประสิทธิภาพของ Faster-RCNN ด้วยปริมาณข้อมูลภาพจำนวนที่ไม่มากนักดังกล่าว จึงเหมาะที่จะสอนและวัดผล Faster-RCNN ด้วยวิธีการแบ่งข้อมูลทดสอบ K กลุ่ม (K-Fold Cross Validation) ดังนั้นจึงกำหนดให้แต่ละกลุ่มทดลองใช้วิธีการสอนและทดสอบเน็ตเวิร์กด้วยวิธีการแบ่งข้อมูลทดสอบ K=5 กลุ่ม โดยในแต่ละกลุ่มทดลองจะแบ่งข้อมูลภาพออกเป็น 5 กลุ่มย่อย กลุ่มละ 100 ชุดภาพ ซึ่งใช้สร้าง Faster-RCNN จำนวน 5 โมเดล ซึ่งแต่ละโมเดลจะถูกสอนโดย 4 กลุ่มข้อมูลย่อยที่ไม่ซ้ำกัน และใช้ 1 กลุ่มข้อมูลย่อยที่เหลือเพื่อวัดผลของเน็ตเวิร์ก

การเพิ่มจำนวนภาพของชุดข้อมูลขณะสอนเน็ตเวิร์ก (Image Augmentation) ข้อมูลภาพจำนวน 500 ชุดภาพ เมื่อแบ่งเป็นกลุ่มข้อมูลสำหรับสอน 400 ภาพ มีปริมาณที่อาจทำให้ Faster-RCNN ที่สร้างขึ้นมีความเฉพาะเจาะจงกับชุดภาพมากเกินไป (Overfitting) ดังนั้นจึงเพิ่มความหลากหลายให้กับข้อมูลภาพถ่ายในขณะสอนเน็ตเวิร์ก โดยกำหนดให้ในแต่ละรอบการสอนเน็ตเวิร์กจนครบข้อมูลสำหรับสอนทุกภาพ (Epoch) ในแต่ละรอบถัดไปจะทำการแปลงภาพทุกภาพให้เปลี่ยนไปจากเดิมในสองรูปแบบ ได้แก่

- 1 สุ่มหมุนภาพ (Rotation) ด้วยมุมขนาด ± 15 องศา เพื่อเพิ่มความหลากหลายของลักษณะเอียงของรอยกริดและถัวยรอน้ำยาง เพื่อจำลองสถานการณ์ในกรณีที่ดินยาง, รอยกริด, หรือถัวยรอน้ำยางมีความเอียงมากกว่าปกติ
- 2 สุ่มเพิ่มหรือลดค่าความสว่างเฉลี่ยในภาพ (Exposure Compensation) $\pm 10\%$ เพื่อจำลองสถานการณ์เมื่อสภาพแสงขณะถ่ายภาพ มีความผันผวนหรือผลจากแสงรบกวนภายนอก และเมื่อระยะห่างของต้นยางกับกล้องถ่ายภาพมากขึ้นหรือลดลง

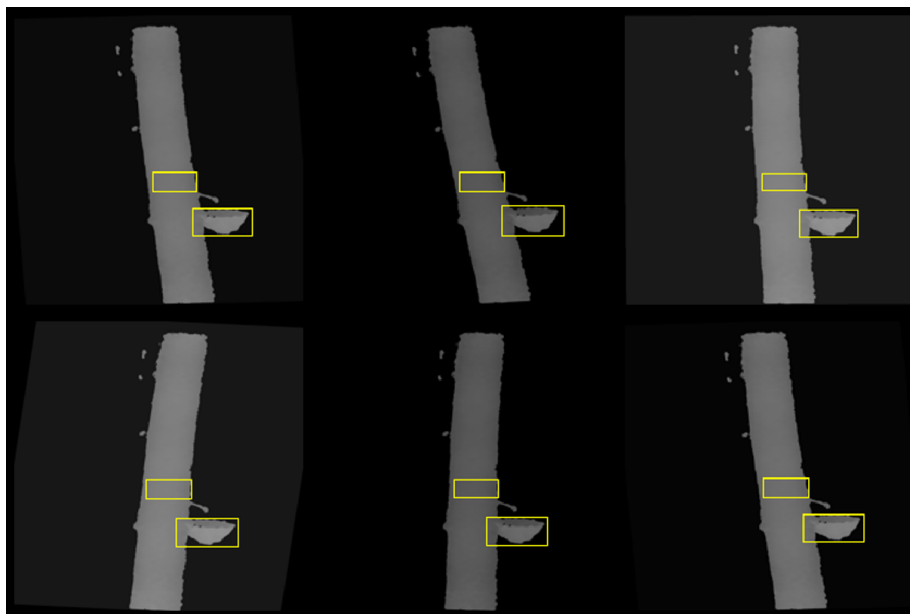


รูปที่ 2.40: การเตรียมข้อมูลภาพสำหรับสอนและทดสอบเน็ตเวิร์ก

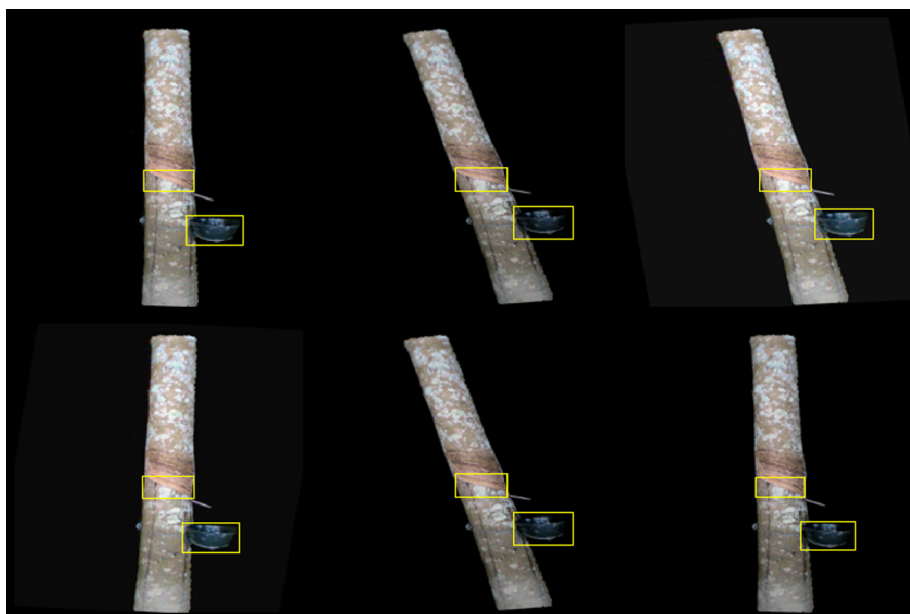
แผนภาพสรุปการแบ่งกลุ่มทดลองและการเตรียมภาพสำหรับสอนและทดสอบเน็ตเวิร์กแสดงในรูปที่ 2.40 ตัวอย่างภาพอินพุตสำหรับการสอน Faster-RCNN ของกลุ่มทดลองที่ 1 แสดงในรูปที่ 2.41, กลุ่มทดลองที่ 2 รูปที่ 2.42, และกลุ่มทดลองที่ 3 แสดงในรูปที่ 2.43



รูปที่ 2.41: ตัวอย่างการแปลงข้อมูลภาพในกลุ่มทดลองที่ 1 จากภาพตัวอย่างภาพถ่ายสี RGB 1 ภาพ เมื่อถูกแปลงขนาดให้เท่ากับ 224x224 พิกเซล และเมื่อสุ่มแปลงภาพด้วยการหมุนและขดเซยความสว่างใน 6 รูปแบบที่ไม่ซ้ำกัน พร้อมทั้งแสดงกรอบสี่เหลี่ยมล้อมรอบอ้างอิงของรอยกรีดและถ้วยรองน้ำยางในภาพ



รูปที่ 2.42: ตัวอย่างการแปลงข้อมูลภาพในกลุ่มทดลองที่ 2 จากภาพตัวอย่างภาพถ่ายความลึก 1 ภาพ เมื่อถูกแปลงขนาดให้เท่ากับ 224x224 พิกเซล และเมื่อสุมแปลงภาพด้วยการหมุนและชัดเจนความสว่างใน 6 รูปแบบที่ไม่ซ้ำกัน พร้อมทั้งแสดงกรอบสี่เหลี่ยมล้อมรอบอ้างอิงของรอยกริดและถ้วยรองน้ำยางในภาพ



รูปที่ 2.43: ตัวอย่างการแปลงข้อมูลภาพในกลุ่มทดลองที่ 3 จากภาพตัวอย่างภาพ 1 ชุดภาพ เมื่อถูกแปลงขนาดให้เท่ากับ 224x224 พิกเซล และเมื่อสุมแปลงภาพด้วยการหมุนและชัดเจนความสว่างใน 6 รูปแบบที่ไม่ซ้ำกัน พร้อมทั้งแสดงกรอบสี่เหลี่ยมล้อมรอบอ้างอิงของรอยกริดและถ้วยรองน้ำยางในภาพ

2.3.5 ตำแหน่งในสามมิติของรอยและกริดด้วยร่องน้ำอย่างอ้างอิงกับกล้องถ่ายภาพมุมกว้าง

ตำแหน่งกรอบล้อมรอบรอยกริดและด้วยร่องน้ำอย่างอยู่ในระบบพิกัดภาพ $C_1 : \{U_1 \times V_1\}$ ซึ่งระบบหุ่นยนต์ต้องการตำแหน่งของรอยกริดและด้วยร่องน้ำอย่างในระบบพิกัดฉากสามมิติ เพื่ออ้างอิงถึงตำแหน่งของหุ่นยนต์ ทำให้ต้องแปลงตำแหน่งรอยกริดและด้วยร่องน้ำอย่างในพิกัดภาพเป็นตำแหน่งในพิกัดฉาก

2.3.5.1 การแปลงตำแหน่งในพิกัดภาพเป็นตำแหน่งในพิกัดฉากสามมิติของกล้องถ่ายภาพมุมกว้าง

กำหนดให้ตำแหน่งฐานของหุ่นยนต์ คือตำแหน่งกล้องถ่ายภาพมุมกว้าง โดยกล้องถ่ายภาพมุมกว้างได้กำหนดให้พิกัดภาพ คือระบบพิกัดภาพของโมดูลกล้องถ่ายภาพสี ทำให้ทั้งภาพถ่ายสีและภาพถ่ายความลึกต่างอ้างอิงกับระบบพิกัดของโมดูลกล้องถ่ายภาพสี

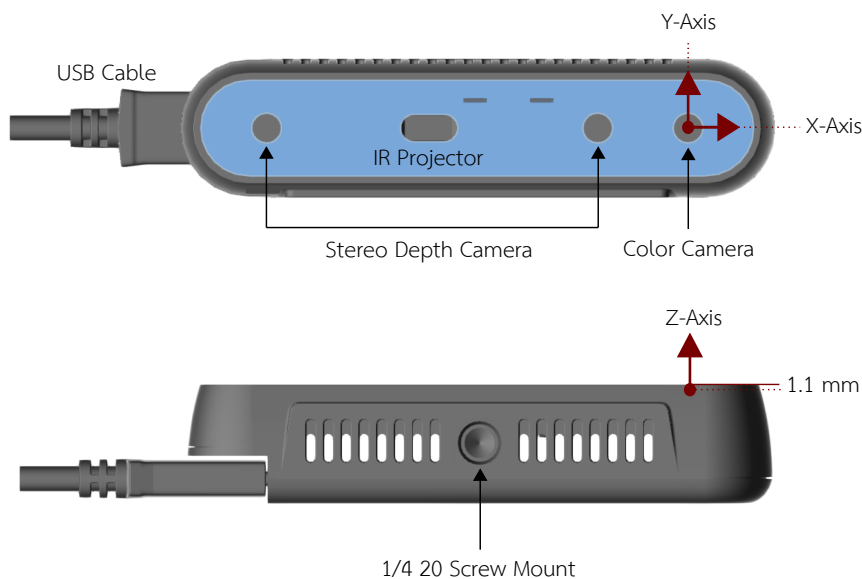
ตำแหน่งจุดกำเนิดของกล้องถ่ายภาพมุมกว้าง อยู่ ณ ตำแหน่งกลางหน่วยรับภาพสีของโมดูลกล้องถ่ายภาพสี ซึ่งแกนอ้างอิงของระบบพิกัดฉากของกล้องถ่ายภาพสียังคงเป็นไปตามข้อกำหนด ตามข้อย่อย 2.2.3 ในรูปที่ 2.7 ซึ่งจากเอกสารอ้างอิงของกล้องโมเดล D415 กำหนดให้ระนาบความลึกเท่ากับศูนย์ ($Z=0$) อยู่ ณ ตำแหน่งถัดจากหน้ากระจกภายนอก (Cover Glass) เข้าไปด้านในตัวกล้อง D415 เป็นระยะ 1.1 มิลลิเมตร [10] โดยหากจะนำไปใช้งานจริงกับระบบหุ่นยนต์จะต้องแปลงค่าพิกัดนี้อีกครั้งหนึ่งเพื่อเทียบกับตำแหน่งจุดกำเนิดกับจุดยึดกล้อง (Screw Mount) ดังแสดงในรูปที่ 2.44

การคำนวณตำแหน่งพิกัดภาพเป็นระยะทางในพิกัดฉากของกล้องโมเดล D415 ใช้หลักการโปรเจกชันของกล้องรูเข็ม (Pinhole Camera) ซึ่งทางผู้ผลิตกล้องได้ให้ค่าของพารามิเตอร์สำหรับแปลงพิกัด (Intrinsic Parameters) ได้แก่ ทางยาวโฟกัสในแนวแกน X (f_x), ทางยาวโฟกัสในแนวแกน Y (f_y) และตำแหน่งจุดฉาย (Projection Point), (pp_x, pp_y) กับทั้งโมดูลกล้องถ่ายภาพสี (Color Camera Module) และโมดูลกล้องถ่ายภาพความลึก (Depth Module) ของกล้องโมเดล D415 โดยสำหรับภาพถ่ายสีที่ขนาด 1280×720 พิกเซล ในกรณีที่กล้องวางตัวที่ปกติ (แกนนอนของกล้องตามปกติขนานกับระดับพื้น ดังรูปที่ 2.44) พารามิเตอร์สำหรับแปลงพิกัดมีค่าดังนี้

$$f_x = 924.004, f_y = 923.654$$

$$pp_x = 612.835, pp_y = 356.136$$

สมการที่ 2.11 คือความสัมพันธ์ระหว่างตำแหน่งในพิกัดภาพ ในกรณีที่กล้องวางตัวตามปกติ (Landscape Orientation) ซึ่งอยู่ในรูป (u, v, p_{D1}) โดย $p_{D1}(u, v)$ คือ ค่าความลึกในระยะ 50 ถึง 200 เซนติเมตร ณ ตำแหน่ง (u, v) ใดๆ ในภาพ ซึ่งมีค่าเป็นจำนวนเต็มตามค่าพิกเซลขนาด 8 บิต กับตำแหน่งในพิกัดฉากสามมิติซึ่งอยู่ในรูป (x, y, z) เทียบกับจุดกำเนิดของกรอบอ้างอิงในรูปที่ 2.44 ในหน่วยเซนติเมตร

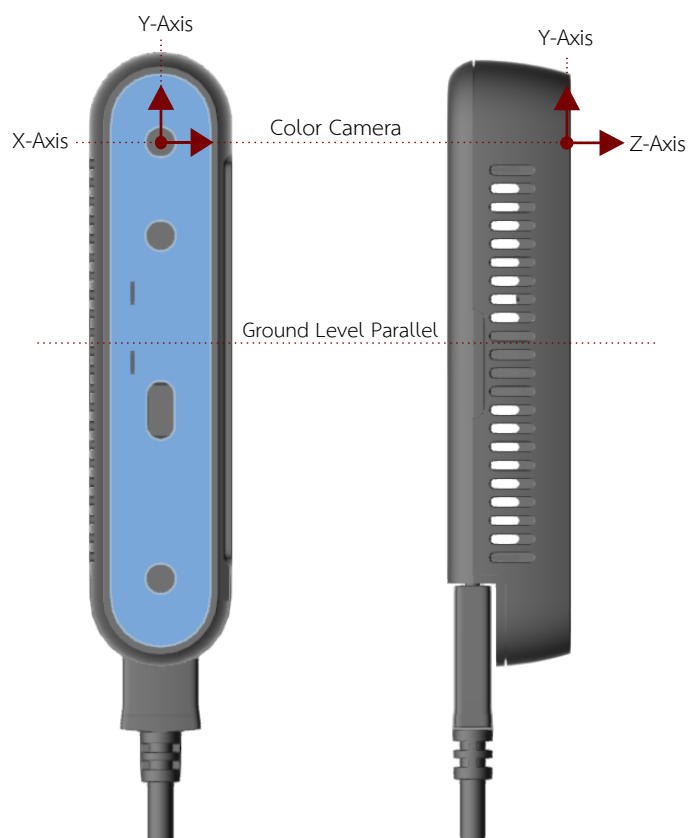


รูปที่ 2.44: ภาพสามมิติโมเดลกล้อง D415 [10, 11] แสดงตำแหน่งกล้องถ่ายภาพสี (Color Camera) และกล้องถ่ายภาพความลึก (Stereo Depth Camera) ที่อยู่ภายใน และกรอบอ้างอิงพิกัดฉากที่มีจุดกำเนิด ณ ตำแหน่งกลางกล้องถ่ายภาพสี โดยกรอบอ้างอิงมีทิศทางของแกนตามเดิมจากผู้ผลิตกล้อง และระนาบความลึกศูนย์ ($Z=0$) อยู่ถัดจากกระจกหน้าเลนส์ (Cover Glass) เข้าไปในตัวกล้องเป็นระยะ 1.1 มิลลิเมตร

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{f_x} & 0 & 0 \\ 0 & \frac{1}{f_y} & 0 \\ 0 & 0 & -\frac{150}{255} \end{bmatrix} \begin{bmatrix} u \\ v \\ p_{D1}(u, v) \end{bmatrix} + \begin{bmatrix} -\frac{pp_x}{f_x} \\ -\frac{pp_y}{f_y} \\ 200 \end{bmatrix} \quad (2.11)$$

เนื่องจากลักษณะการวางกล้องถ่ายภาพมุมกว้างที่กำหนดให้แกนนอนตามปกติของกล้องกลับเป็นแกนตั้ง (Portrait Orientation) หรือกล่าวคือมีการสลับระหว่างแกน X และแกน Y แสดงดังรูปที่ 2.45 ทำให้ในการคำนวณการแปลงค่าระหว่างพิกัดตำแหน่งภาพเป็นพิกัดสามมิติในสมการที่ 2.11 จึงต้องมีการสลับค่าระหว่างพารามิเตอร์สำหรับแปลงพิกัดที่กำหนดค่าสำหรับแกน X และ Y ซึ่งเป็นไปตามสมการที่ 2.12

$$\begin{bmatrix} y \\ x \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{f_x} & 0 & 0 \\ 0 & \frac{1}{f_y} & 0 \\ 0 & 0 & -\frac{150}{255} \end{bmatrix} \begin{bmatrix} u \\ v \\ p_{D1}(u, v) \end{bmatrix} + \begin{bmatrix} -\frac{pp_x}{f_x} \\ -\frac{pp_y}{f_y} \\ 200 \end{bmatrix} \quad (2.12)$$



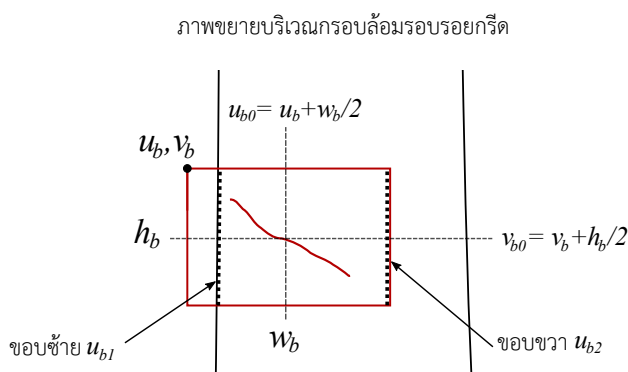
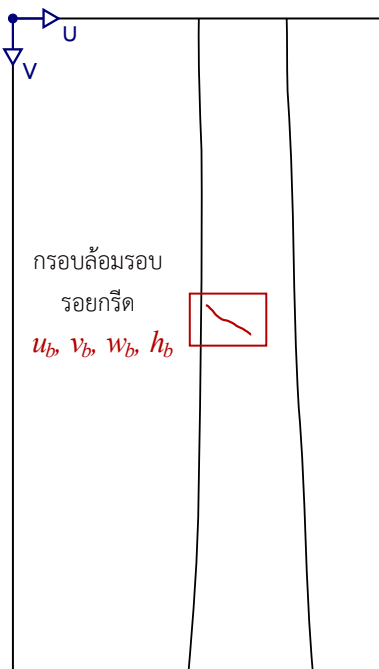
รูปที่ 2.45: ภาพโมเดลกล้อง D415 [11] เมื่อกล้องวางตัวในแนวตั้ง (Portrait Orientation) โดยแสดงกรอบอ้างอิง มีจุดกำเนิด ณ ตำแหน่งกลางกล้องถ่ายภาพสี (Color Camera) โดยแกน Y มีทิศทางพุ่งขึ้นจากระนาบพื้น, แกน X ขนานกับระดับพื้น และแกนความลึก Z ชี้เข้าหาต้นยางพารา

2.3.5.2 ตำแหน่งของรอยกริดในสามมิติอ้างอิงกับตำแหน่งกล้องถ่ายภาพมุมกว้าง

การตรวจหารอยกริดให้ผลลัพธ์เป็นกรอบสี่เหลี่ยมล้อมรอบ ทำให้ตำแหน่งต่างๆ ภายในกรอบสี่เหลี่ยมนั้นจึงเป็นตำแหน่งที่อาจเป็นรอยกริด ซึ่งหากในภาพถ่ายความลึก ณ ตำแหน่งใดๆ ภายในกรอบสี่เหลี่ยมมีค่าไม่เท่ากับศูนย์ จะเป็นตำแหน่งที่สามารถแปลงเป็นตำแหน่งในพิกัดฉากได้ ด้วยสมการที่ 2.12 โดยตำแหน่งที่ได้จะเป็นตำแหน่งในพิกัดสามมิติอ้างอิงกับตำแหน่งของกล้องถ่ายภาพมุมกว้าง

ลักษณะของกรอบสี่เหลี่ยมผลลัพธ์การตรวจหารอยกริดอาจมีได้หลายรูปแบบ เช่นกรอบสี่เหลี่ยมผลลัพธ์มีขนาดใหญ่มากกว่าขนาดลำต้น, หรือมีขนาดเล็กกว่าและอยู่ภายในพื้นที่ลำต้น โดยจากข้อเท็จจริงที่ว่ารอยกริดจะอยู่ในพื้นที่ลำต้นเท่านั้น ทำให้สามารถหาตำแหน่งภายในกรอบสี่เหลี่ยมได้ในหลายรูปแบบ ยกตัวอย่างในรูปที่ 2.46 ซึ่งสมมติให้ (u_b, v_b, w_b, h_b) คือกรอบสี่เหลี่ยมผลลัพธ์การตรวจหารอยกริด, (u_{b0}, v_{b0}) คือตำแหน่งพิกเซลกึ่งกลางของกรอบสี่เหลี่ยม, ตำแหน่งขอบของลำต้นที่อยู่ภายในกรอบสี่เหลี่ยม (ขอบซ้ายตามตัวอย่างในรูปที่ 2.46), ตำแหน่งขอบของกรอบสี่เหลี่ยมที่เป็นพื้นที่ลำต้นที่ปรากฏในภาพ, ซึ่งสามารถคำนวณหาได้โดยอาศัยภาพขาวดำของพื้นที่ลำต้น (ขอบขวาตามตัวอย่างในรูปที่ 2.46) เช่น คอลัมน์ของขอบลำต้นที่อยู่ภายในกรอบโดยเฉลี่ย (ขอบซ้าย, u_{b1}), คอลัมน์ของขอบของกรอบสี่เหลี่ยมโดยเฉลี่ย (ขอบขวา, u_{b2}) เป็นต้น ซึ่งตำแหน่งต่างๆ เหล่านี้สามารถใช้เพื่อคำนวณการวางกล้องระยะใกล้ได้

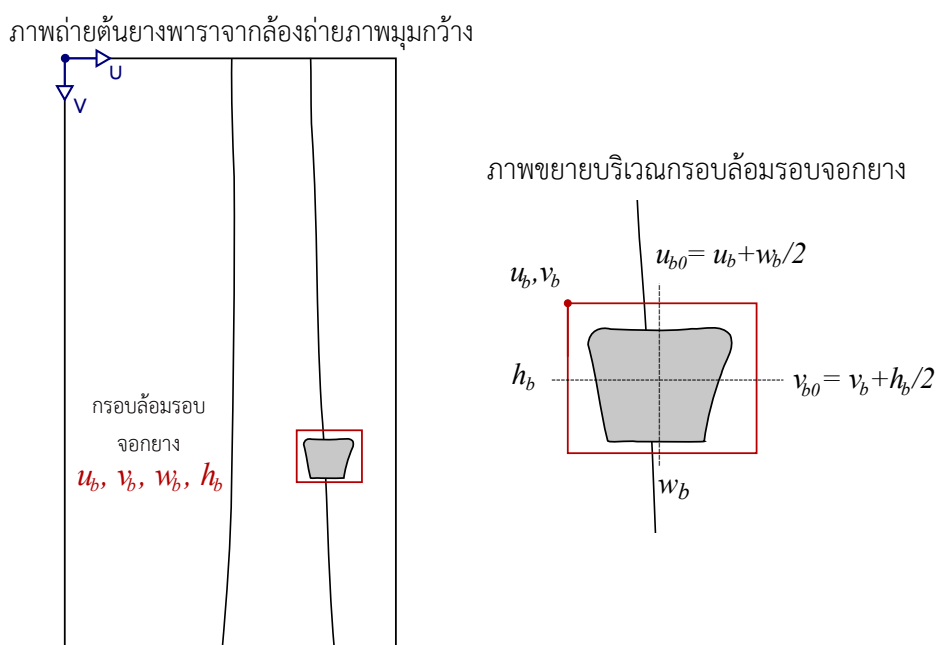
ภาพถ่ายต้นยางพาราจากกล้องถ่ายภาพมุมกว้าง



รูปที่ 2.46: แสดงตำแหน่งภายในกรอบสี่เหลี่ยมล้อมรอบรอยกริด และ (u_{b0}, v_{b0}) คือจุดกึ่งกลางของกรอบสี่เหลี่ยมในพิกัดภาพ

2.3.5.3 ตำแหน่งของถ้วยรองน้ำยางในสามมิติอ้างอิงกับตำแหน่งกล้องถ่ายภาพมุมกว้าง

ตำแหน่งถ้วยรองน้ำยางในพิกัดฉากสามมิติอ้างอิงกับกล้องถ่ายภาพมุมกว้างสามารถคำนวณได้ในลักษณะเดียวกันกับกรณีรอยกรีด โดยหากค่าพิกเซลภายในกรอบสี่เหลี่ยมในภาพถ่ายความลึกมีค่าไม่เท่ากับศูนย์ จะสามารถคำนวณหาตำแหน่งในสามมิติได้ด้วยสมการ 2.12 แต่ทั้งนี้กรอบสี่เหลี่ยมล้อมรอบผลลัพธ์จากการตรวจหาถ้วยรองน้ำยาง ไม่ได้รับรองว่ากรอบจะซ้อนทับถ้วยรองน้ำยางในลักษณะใด เช่นกรอบอาจครอบคลุมถ้วยรองน้ำยางทั้งหมด หรือกรอบซ้อนทับกับตำแหน่งถ้วยรองน้ำยางเฉพาะบางส่วน ดังนั้นจึงประมาณตำแหน่งของถ้วยรองน้ำยางด้วย จุดกึ่งกลางของกรอบสี่เหลี่ยมล้อมรอบ ซึ่งเป็นตำแหน่งที่มีแนวโน้มที่มีถ้วยรองน้ำมากกว่าบริเวณอื่นภายในกรอบสี่เหลี่ยม ตัวอย่างตามรูปที่ 2.47



รูปที่ 2.47: แสดงตำแหน่งกรอบล้อมรอบถ้วยรองน้ำยางในภาพ และ (u_{b0}, v_{b0}) คือจุดกึ่งกลางของกรอบสี่เหลี่ยมล้อมรอบในพิกัดภาพ

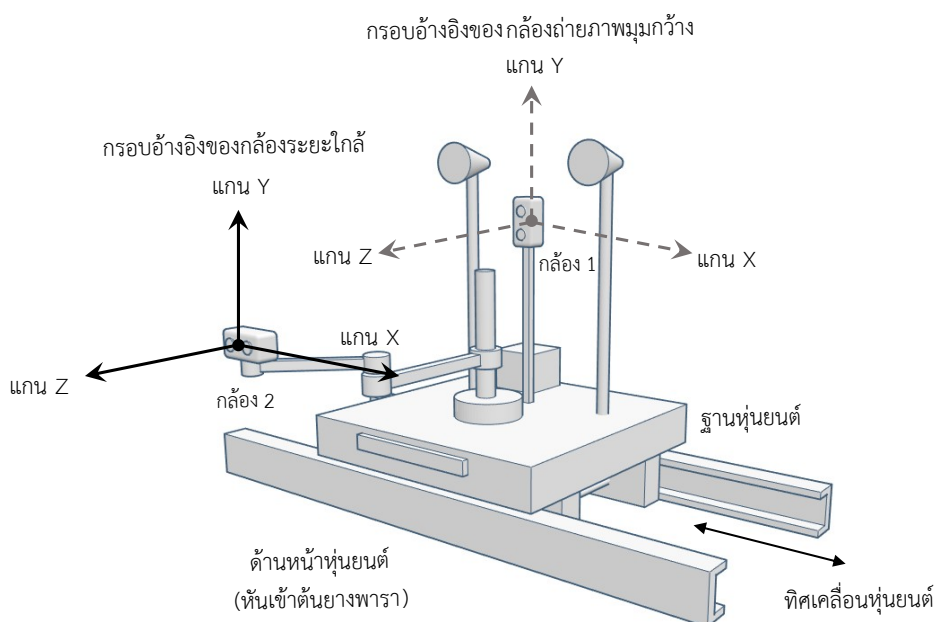
2.4 การตรวจหาแนวรอยกรีดด้วยกล้องระยะใกล้ (กล้อง 2)

กล้องระยะใกล้ (กล้อง 2) มีเป้าหมายเพื่อหาตำแหน่งรอยกรีดอย่างละเอียด หรือ “แนวรอยกรีด” ซึ่งเป็นเส้นทางในสามมิติของรอยกรีด โดยแนวรอยกรีดจะเป็นตำแหน่งในสามมิติที่อ้างอิงกับกรอบอ้างอิงของกล้องระยะใกล้ ดังนั้นผลการตรวจหาแนวรอยกรีดของกล้องระยะใกล้ จะให้ผลลัพธ์ต่างจากกรอบสี่เหลี่ยมล้อมรอบรอยกรีดผลลัพธ์การตรวจหาของกล้องถ่ายภาพมุมกว้าง และการถ่ายภาพรอยกรีดด้วยกล้องระยะใกล้จะให้มุมมองภาพต่างจากกล้องถ่ายภาพมุมกว้าง ทำให้ต้องพัฒนาวิธีตรวจหาแนวรอยกรีดเฉพาะสำหรับกล้องระยะใกล้

2.4.1 การถ่ายภาพแนวรอยกรีดด้วยกล้องระยะใกล้

กำหนดให้ระบบหุ่นยนต์ใช้งานกล้องโมเดล D435 เป็นกล้องระยะใกล้ โดยจากการอธิบายในข้อย่อที่ 2.2.2 กล้องโมเดล D435 สามารถเข้าใกล้วัตถุได้ใกล้ที่สุดที่ 280 มิลลิเมตร ที่ขนาดความละเอียดของภาพ 1280x720 พิกเซล, มีมุมมองภาพตามปกติของทุกโมดูลกล้องในระยะใกล้สามารถครอบคลุมแนวรอยกรีดบนหน้ายางได้ทั้งหมด โดยไม่ต้องพึ่งพาการปรับมุมกล้องตั้งเช่นกรณีกล้องถ่ายภาพมุมกว้าง, และกล้องมีขนาดเล็กเหมาะกับการติดตั้งในบริเวณปลายแขนหุ่นยนต์

กรอบอ้างอิงของกล้องระยะใกล้ที่ตำแหน่งปลายแขนหุ่นยนต์แสดงในรูปที่ 2.48 โดยการกำหนดทิศทางของแกนกรอบอ้างอิงกล้องระยะใกล้เริ่มต้น เป็นไปตามการวางกรอบอ้างอิงของกล้องถ่ายภาพมุมกว้างโดย แกน Z ของกล้องระยะใกล้จะชี้เข้าหาลำต้นยางพารา, แกน Y มีทิศชี้ขึ้นตามแนวตั้ง, และแกน X เป็นไปตามกฎมือขวา ซึ่งรายละเอียดการปรับมุมกล้องระยะใกล้เพื่อถ่ายภาพรอยกรีดจะอธิบายในหัวข้อถัดไป



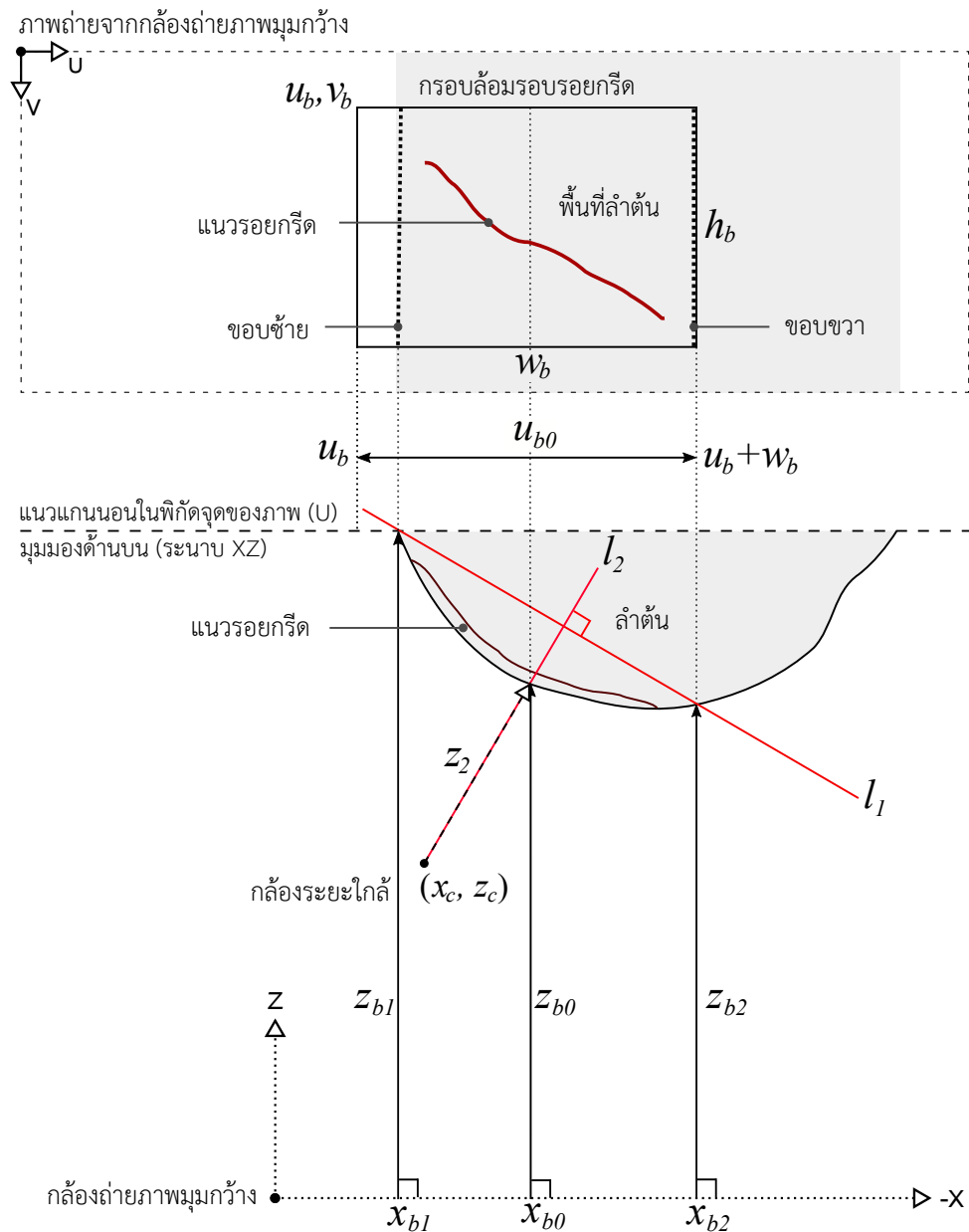
รูปที่ 2.48: แสดงกรอบอ้างอิงของกล้องระยะใกล้ในตำแหน่งปลายแขนหุ่นยนต์

2.4.1.1 การหาตำแหน่งและมุมของกล้องระยะใกล้สำหรับถ่ายภาพแนวรอยกริด

การวางกล้องเพื่อถ่ายภาพรอยกริดระยะใกล้ทำโดย กำหนดให้กล้องระยะใกล้โมเดล D435 วางห่างเป็นระยะตั้งฉากจากหน้ายางในช่วง 20 ถึง 45 เซนติเมตร ซึ่งเป็นช่วงที่กล้องให้ค่าความผิดพลาดต่ำที่สุดประมาณ 0.1 มิลลิเมตร [27] ซึ่งจะต้องปรับการตั้งค่าเพิ่มเติมเพื่อให้โมดูลกล้องถ่ายภาพความลึกสามารถทำงานได้ที่ระยะห่างน้อยที่สุดที่ 20 เซนติเมตร และกำหนดให้กล้องหันเข้าหารอยกริดเพื่อให้มุมรับภาพของกล้องครอบคลุมแนวรอยกริด ซึ่งส่งผลให้แกนความลึกของกล้องจะต้องตั้งฉากกับผิวหน้ายาง หรือกล่าวคือภาพถ่ายของกล้องระยะใกล้จะต้องมีลำต้นปรากฏในบริเวณกลางภาพ

การคำนวณตำแหน่งการวางกล้องระยะใกล้ อาศัยตำแหน่งต่างๆ ในกรอบสี่เหลี่ยมล้อมรอบรอยกริดในภาพจากกล้องถ่ายภาพมุมกว้าง โดยตำแหน่งต่างๆ ภายในกรอบเมื่อแปลงเป็นตำแหน่งในพิกัดฉากสามมิติแสดงในรูปที่ 2.49 ซึ่งได้แก่ตำแหน่งขอบซ้ายของพื้นที่ลำต้นที่ปรากฏในกรอบสี่เหลี่ยม, ตำแหน่งขอบขวาของพื้นที่ลำต้นที่ปรากฏในกรอบสี่เหลี่ยม เพื่อใช้ประมาณมุมของเส้นสัมผัสผิวลำต้นในบริเวณตำแหน่งกึ่งกลางกรอบสี่เหลี่ยมล้อมรอบรอยกริด (เส้นตรง l_1 ในรูปที่ 2.49) โดยอัลกอริทึม 2.4.1 จะคำนวณเส้นสัมผัสผิวลำต้นในตำแหน่งกึ่งกลางกรอบสี่เหลี่ยมล้อมรอบรอยกริด (x_{b0}, y_{b0}, z_{b0}) ซึ่งแนวแกนความลึกของกล้องระยะใกล้ (l_2) จะต้องผ่านจุดกึ่งกลางของกรอบสี่เหลี่ยมและตั้งฉากกับเส้นสัมผัสผิวลำต้นดังกล่าว และห่างจากผิวหน้ายางเป็นระยะตั้งฉากจากหน้ายางในช่วง 20 ถึง 45 เซนติเมตร (z_2)

การบังคับให้แกนความลึกของกล้องตั้งฉากกับผิวของหน้ายางจะต้องปรับมุมก้ม-เงยของกล้องเช่นกัน ซึ่งสามารถทำได้โดยปรับมุมก้มเงยของกล้องให้ระนาบรับภาพของกล้องระยะใกล้ (ระนาบ XY) ขนานกับพื้นผิวหน้ายางดังแสดงในรูปที่ 2.50 ด้วยวิธีอ่านค่าความลึกของลำต้นที่ปรากฏในภาพถ่ายความบริเวณส่วนบนกลางภาพ (z_{2top}) และบริเวณลำต้นส่วนล่างกลางภาพ (z_{2bot}) แล้วปรับมุมก้ม-เงยให้ความลึกทั้งสองมีค่าประมาณใกล้เคียงกัน ทั้งนี้สามารถประมาณค่าความลึกพื้นลำต้นตามแนวกลางภาพทั้งหมดมีค่าใกล้เคียงกับความลึกของลำต้นกลางภาพ ($z_2 \approx z_{2top} \approx z_{2bot}$)



รูปที่ 2.49: ตำแหน่งและการหันกล้องระยะใกล้เพื่อถ่ายภาพรอยกรีด, (x_c, y_c, z_c) , โดยแสดงความสัมพันธ์ระหว่างตำแหน่งในระนาบ XZ ของกรอบอ้างอิงของกล้องถ่ายภาพมุมกว้าง กับตำแหน่งในกรอบล้อมรอบรอยกรีดตามแกนนอนของภาพ (U) ภายในพิกัดภาพจากกล้องถ่ายภาพมุมกว้าง

อัลกอริทึม 2.4.1: ตำแหน่งและทิศทางการส่องระยะใกล้เพื่อถ่ายภาพพรอยกรีด

Input:

$\{u_b, v_b, w_b, h_b\}$: กรอบสี่เหลี่ยมล้อมรอบพรอยกรีด,

z_2 : ระยะความลึกของกล้องระยะใกล้กับลำต้น,

I_M : ภาพขาวดำของลำต้นยางพารา,

I_{D1} : ภาพถ่ายความลึกของกล้องถ่ายภาพมุมกว้าง

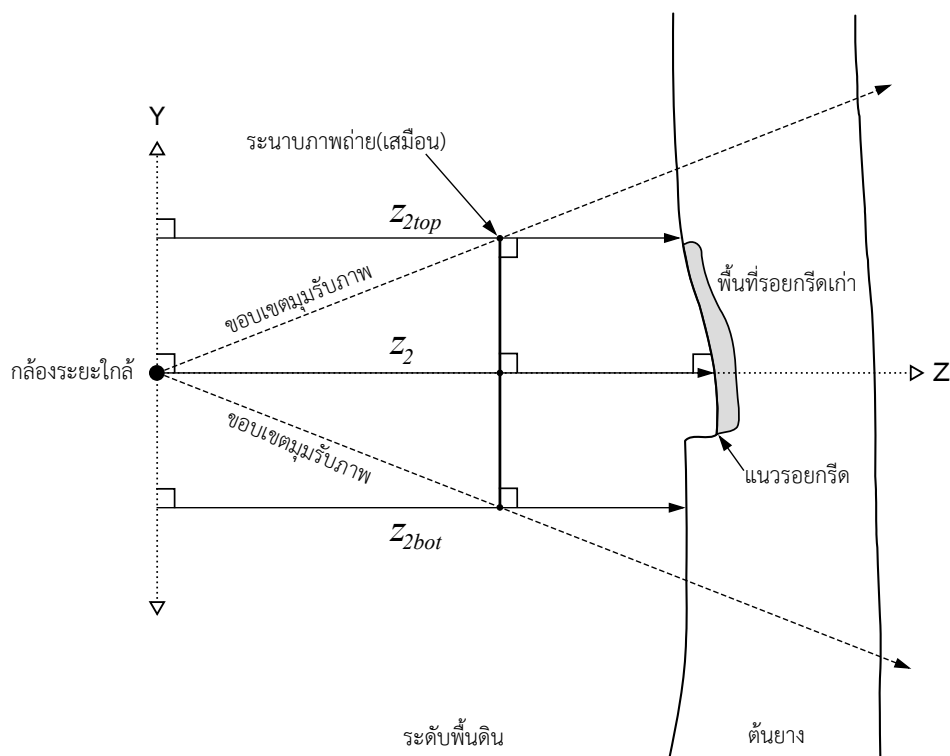
Output:

A : จุดวางกล้องระยะใกล้อ้างอิงกับตำแหน่งกล้องถ่ายภาพมุมกว้าง,

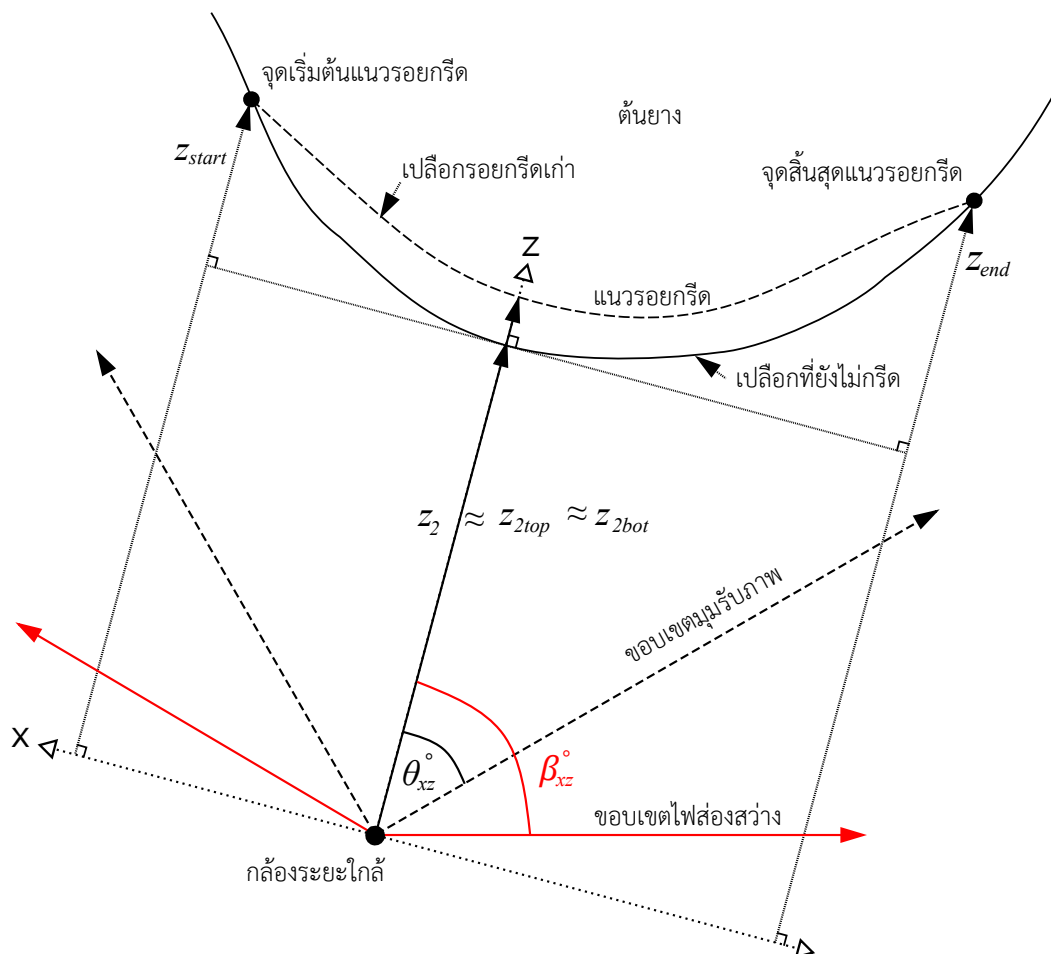
B : เวกเตอร์ทิศแกน Z ของกล้องระยะใกล้ภายในกรอบอ้างอิงกล้องถ่ายภาพมุมกว้าง

- 1 คำนวณหาตำแหน่งกึ่งกลางของกรอบล้อมรอบพรอยกรีด (u_{b0}, v_{b0}) โดย

$$v_{b0} = v_b + \frac{h_b}{2}, u_{b0} = u_b + \frac{w_b}{2}$$
 - 2 ความลึก (z_{b0}) คือความลึกบริเวณกึ่งกลางของกรอบล้อมรอบพรอยกรีด (u_{b0}, v_{b0})
 คำนวณจากภาพถ่ายความลึก I_{D1} , $z_{b0} = p_{D1}(u_{b0}, v_{b0})$
 - 3 แปลงตำแหน่งในพิกัดภาพ ไปเป็นตำแหน่งในพิกัดฉากสามมิติของกรอบอ้างอิงกล้อง
 ถ่ายภาพมุมกว้าง ด้วยสมการ 2.12, $(u_{b0}, v_{b0}, z_{b0}) \rightarrow (x_{b0}, y_{b0}, z_{b0})$
 - 4 คำนวณขอบซ้ายและขอบขวาของลำต้นภายในพื้นที่กรอบล้อมรอบพรอยกรีด ซึ่งอยู่
 ภายในพื้นที่ที่มีค่าพิกเซลเท่ากับ 1 ของภาพขาวดำลำต้น (I_M , $p_M = 1$) โดยขอบ
 ซ้าย คือ เซตของพิกเซลที่มีค่าพิกเซลเท่ากับ 1 ในตำแหน่งตามแกนนอนที่น้อยที่สุด
 ในแต่ละแถวภายในกรอบล้อมรอบพรอยกรีด และขอบขวา คือ เซตของพิกเซลที่มีค่า
 พิกเซลเท่ากับ 1 ในตำแหน่งตามแกนนอนที่มากที่สุดในแต่ละแถวภายในกรอบล้อม
 รอบพรอยกรีด
 - 5 คำนวณค่าเฉลี่ยตำแหน่งตามแกนนอนของขอบซ้าย (u_{b1}) และค่าเฉลี่ยตำแหน่งตาม
 แกนนอนของขอบขวา (u_{b2})
 - 6 คำนวณค่าความลึกเฉลี่ยของขอบซ้าย (z_{b1}) และค่าความลึกเฉลี่ยของขอบขวา (z_{b2})
 จากภาพถ่ายความลึก I_{D1}
 - 7 แปลงตำแหน่งในพิกัดภาพ ไปยังตำแหน่งในพิกัดฉากสามมิติด้วยสมการ 2.12,
 $(u_{b1}, z_{b1}) \rightarrow (x_{b1}, z_{b1})$ และ $(u_{b2}, z_{b2}) \rightarrow (x_{b2}, z_{b2})$
 - 8 บนระนาบ XZ ในกรอบอ้างอิงของกล้องถ่ายภาพมุมกว้าง, สร้างเส้นตรง l_1 ซึ่งลาก
 ผ่านจุด (x_{b1}, z_{b1}) และ (x_{b2}, z_{b2})
 - 9 บนระนาบ XZ ในกรอบอ้างอิงของกล้องถ่ายภาพมุมกว้าง, หาเส้นตรง l_2 ซึ่งตั้งฉาก
 กับ l_1 และลากผ่านจุด (x_{b0}, z_{b0})
 - 10 **A** = (x_c, y_{b0}, z_c) โดยจุด (x_c, z_c) อยู่บนเส้นตรง l_2 ห่างจากจุด (x_{b0}, z_{b0}) จาก
 ผิวต้นยางพารา เป็นระยะ z_2
 - 11 **B** = เวกเตอร์จากจุด (x_c, y_{b0}, z_c) ไปยัง (x_{b0}, y_{b0}, z_{b0})
-

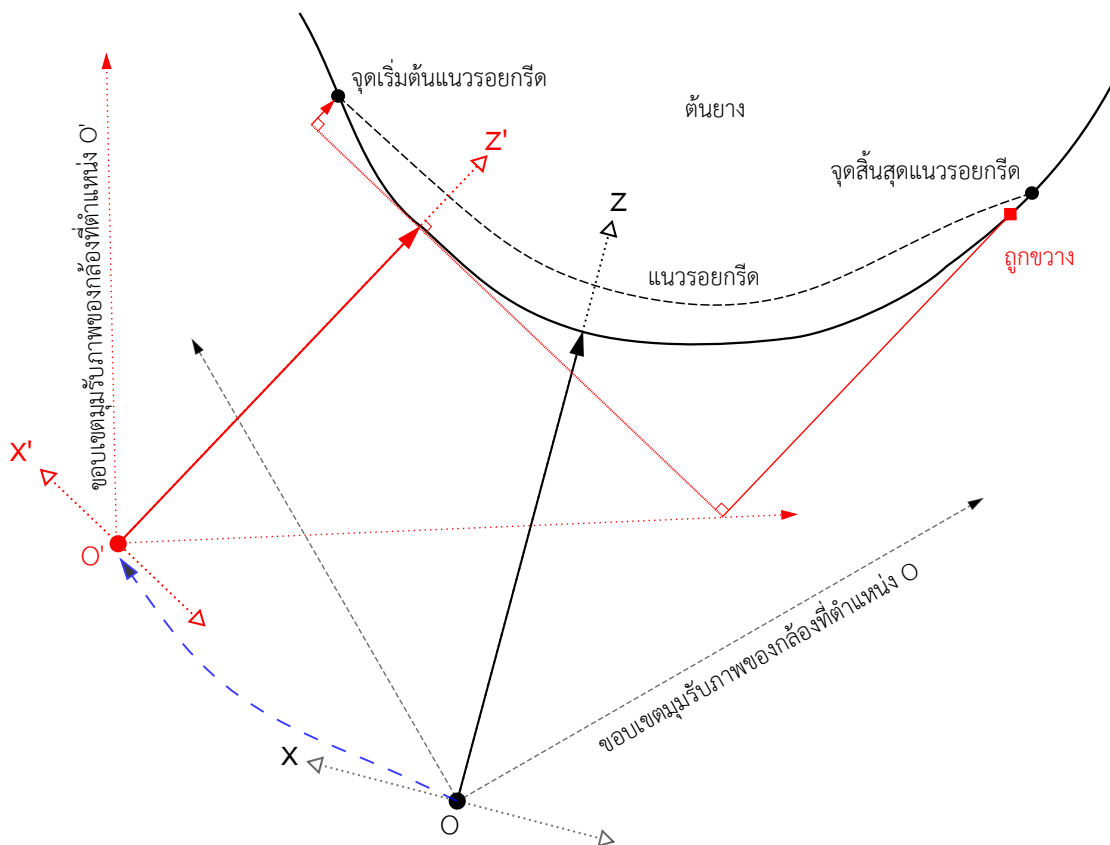


รูปที่ 2.50: แสดงมุมรับภาพของกล้องระยะใกล้ในระนาบ YZ เมื่อกำหนดให้แกนความลึกตั้งฉากกับผิวหน้ายาง ซึ่งทำให้ความลึกบริเวณลำต้นในภาพส่วนบน, ส่วนล่าง, และบริเวณกลางภาพ มีค่าประมาณเท่ากัน $z_{2top} \approx z_{2bot} \approx z_2$



รูปที่ 2.51: ตำแหน่งกล้องระยะใกล้ในระนาบ XZ ในกรณีที่มุมรับภาพครอบคลุมแนวรอยกรีดได้ทั้งหมด ซึ่งทำให้กล้องสามารถถ่ายภาพแนวรอยกรีดได้สมบูรณ์ โดยแสดงความลึกที่จุดเริ่มต้นแนวรอยกรีด (z_{start}) และความลึกจุดสิ้นสุดแนวรอยกรีด (z_{end}), และความลึกตำแหน่งกลางภาพ (z_2)

ตำแหน่งกล้องระยะใกล้ที่คำนวณได้จากอัลกอริทึม 2.4.1 ไม่ได้คำนึงถึงมุมรับภาพของกล้องระยะใกล้กับตำแหน่งของรอยกรีดบนหน้ายาง ทำให้ตำแหน่งการวางกล้องที่ได้ไม่ได้รับรองว่ากล้องระยะใกล้สามารถถ่ายภาพแนวรอยกรีดได้ครบตลอดทั้งแนว ซึ่งแนวรอยกรีดบางส่วนอาจถูกบดบังด้วยส่วนโค้งของลำต้นหรือไม่อยู่ในมุมรับภาพของกล้องระยะใกล้ ดังนั้นการสร้างชุดข้อมูลสำหรับพัฒนาวิธีตรวจหาแนวรอยกรีดของกล้องระยะใกล้ จึงกำหนดตำแหน่งการวางกล้องระยะใกล้ ที่ทำให้ภาพถ่ายแนวรอยกรีดที่ได้จะต้องมีแนวรอยกรีดปรากฏครบตลอดแนว หรือกล่าวคือการถ่ายภาพรอยกรีดจะต้องวางกล้องระยะใกล้ในตำแหน่งที่ทำให้แนวรอยกรีดไม่ถูกบดบังด้วยส่วนโค้งของลำต้น หรืออยู่นอกขอบเขตมุมรับภาพของกล้อง



รูปที่ 2.52: ตำแหน่งกล้องระยะใกล้ในระนาบ XZ เมื่อเคลื่อนจากจุด O (ตามรูปที่ 2.51) ไปยังตำแหน่ง O' ซึ่งเป็นตำแหน่งที่มุมรับภาพไม่ครอบคลุมแนวยกริตทั้งหมด ทำให้ไม่สามารถอ่านค่าความลึกของจุดสิ้นสุดแนวยกริตได้

จากแผนภาพตามรูปที่ 2.51 แสดงตัวอย่างตำแหน่งของกล้องระยะใกล้กับแนวยกริตในระนาบ XZ ซึ่งมุมรับภาพของกล้องครอบคลุมแนวยกริตได้ตลอดทั้งแนว เนื่องจากกล้องสามารถอ่านค่าความลึกของจุดเริ่มแนวยกริต (z_{start}) และความลึกของจุดสิ้นสุดแนวยกริตได้ (z_{end}) โดยจะเรียกภาพถ่ายลักษณะนี้ว่า “ภาพถ่ายแนวยกริตสมบูรณ์” และหากเคลื่อนกล้องระยะใกล้จากตำแหน่งตามรูปที่ 2.51 ครอบอ้างอิงตำแหน่งที่มีจุดกำเนิดที่ O ไปยังตำแหน่ง O' ตามรูปที่ 2.52 จะทำให้มุมรับภาพของกล้องไม่ครอบคลุมแนวยกริตได้กริตทั้งหมด กล้องจะถ่ายภาพแนวยกริตได้ไม่ครบตลอดทั้งแนว เรียกภาพถ่ายแนวยกริตลักษณะนี้ว่า “ภาพถ่ายแนวยกริตไม่สมบูรณ์”

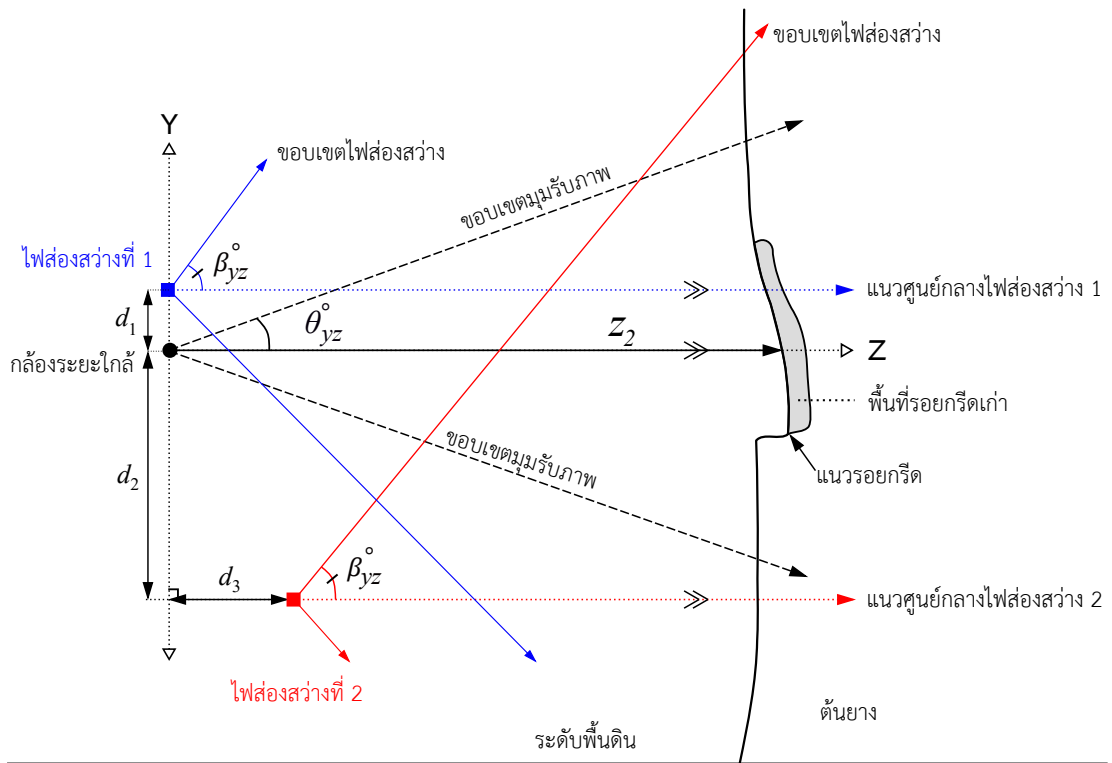
2.4.1.2 ระบบไฟส่องสว่างของกล้องระยะใกล้สำหรับถ่ายภาพรอยกริต

การใช้งานกล้องระยะใกล้ที่ติดตั้งในส่วนปลายแขนหุ่นยนต์เป็นผลให้ไม่สามารถใช้ไฟส่องสว่างชุดเดิมของกล้องถ่ายภาพมุมกว้างที่ตำแหน่งฐานได้ เนื่องจากการเคลื่อนปลายแขนหุ่นยนต์เข้าใกล้หน้ายางจะทำให้เกิดเงาในภาพจากกล้องระยะใกล้เมื่อใช้ไฟส่องสว่างชุดเดิม นอกจากนี้กล้องระยะใกล้จะหันเข้าหารอยกริตตามแนวรัศมีของลำต้นตามอัลกอริทึม 2.4.1 แต่แสงจากไฟส่องสว่างชุดเดิมจะส่องเข้าหาหน้ายางจากตำแหน่งของกล้องถ่ายภาพมุมกว้าง

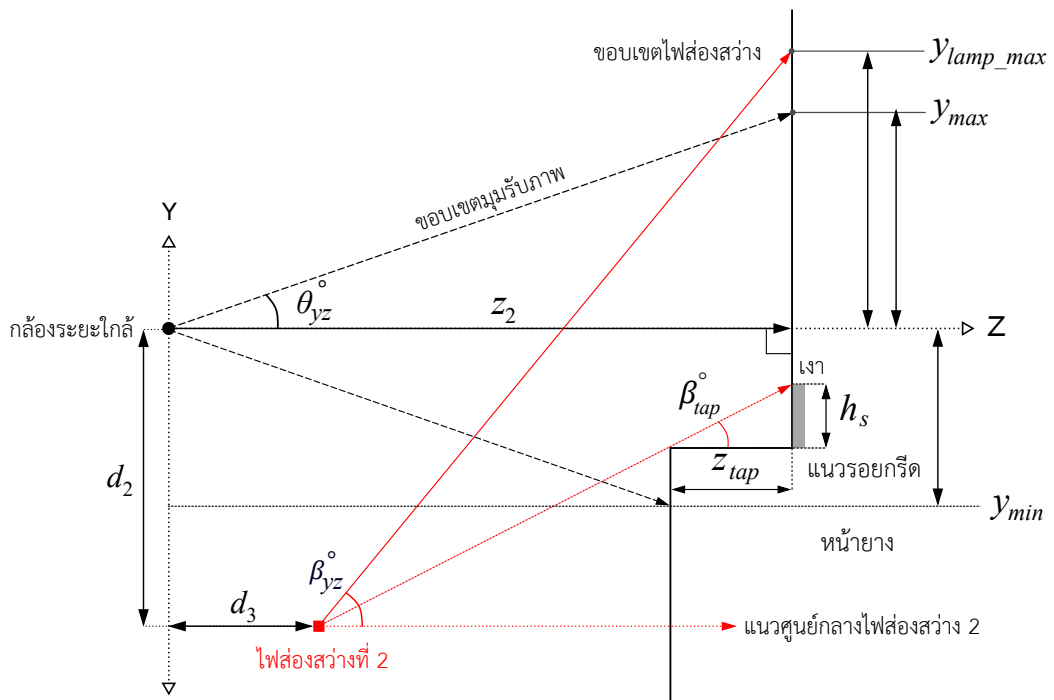
การออกแบบ กำหนดให้กล้องถ่ายภาพระยะใกล้ทำงานร่วมกับไฟส่องสว่างอีกชุดหนึ่ง ไฟส่องสว่างจะถูกยึดเข้ากับปลายแขนหุ่นยนต์ในตำแหน่งของกล้องระยะใกล้ โดยตำแหน่งไฟส่องสว่างมีระยะห่างที่คงที่กับตำแหน่งกับกล้อง และให้มุมรับภาพกับมุมฉายแสงขนานกัน (แกน Z ขนานกับแนวศูนย์กลางลำแสง) เพื่อให้เมื่อกล้องระยะใกล้หันเข้าหาแนวรอยกริดจะทำให้ไฟส่องสว่างหันลำแสงเข้าหาแนวรอยกริดในแนวเดียวกัน การกำหนดให้มุมฉายแสงของไฟส่องสว่างกับมุมรับภาพกล้องระยะใกล้ให้ขนานกัน สอดคล้องกับข้อกำหนดที่ให้กล้องระยะใกล้หันทำมุมก้มหรือมุมเงยกับหน้าียง เพื่อให้ความลึกของลำต้นส่วนบนในภาพ (z_{2top}) มีค่าประมาณเท่ากับความลึกของลำต้นส่วนล่างในภาพ (z_{2bot}) ดังแสดงในรูปที่ 2.50 จะป้องกันไม่ให้แสงสว่างจากไฟส่องสว่างทำให้เกิดความไม่สม่ำเสมอของแสงในภาพ เช่นบางมีพื้นที่ลำต้นในภาพมืดหรือสว่างมากเกินไป

การฉายไฟจากตำแหน่งเหนือกล้องเพียงเล็กน้อยเพียงตำแหน่งเดียวให้กับหน้าียงเพียงพอที่จะทำให้ความสว่างบริเวณลำต้นในภาพลำต้นมีความสว่างเฉลี่ยเท่ากัน แต่เมื่อทดลองให้ฉายไฟจากตำแหน่งต่ำกว่ากล้องให้กับหน้าียง จะทำให้เกิดเงาบริเวณพื้นที่รอยกริดเก่าตามแนวรอยกริด เนื่องจากความต่างของความลึกของเปลือกของลำต้นยาวพาราที่ยังไม่ผ่านการกริดกับพื้นที่รอยกริดเก่าซึ่งคือแห่งของแนวรอยกริด ลักษณะการให้แสงดังกล่าวสามารถใช้ถ่ายภาพและมีประโยชน์ต่อการพัฒนาวิธีตรวจหาแนวรอยกริดได้ ดังนั้นจึงออกแบบให้ระบบไฟส่องสว่างมีไฟจำนวน 2 ตำแหน่งตามรูปที่ 2.53 ประกอบด้วยไฟส่องสว่างที่ 1 มีตำแหน่งเหนือกล้องตามแนวแกน Y เป็นระยะ d_1 และไฟส่องสว่างที่ 2 มีตำแหน่งต่ำกว่ากล้องตามแนวแกน Y เป็นระยะ d_2 และอยู่ใกล้หน้าียงมากกว่ากล้องถ่ายภาพเป็นระยะ d_3 ตามแนวแกน Z ไฟแต่ละตำแหน่งมีข้อกำหนดดังนี้

- **ไฟส่องสว่างที่ 1:** เพื่อให้แสงกับหน้าียงอย่างทั่วถึง แนวศูนย์กลางลำแสงของไฟส่องสว่างที่ 1 ขนานกับแกน Z จึงทำให้มุมฉายแสงของไฟส่องสว่างที่ 1 จะต้องครอบคลุมมุมรับภาพของกล้องเพื่อให้หน้าียงภายในมุมรับภาพของกล้องได้รับแสงอย่างทั่วถึง หรือกล่าวคือไฟส่องสว่างที่ 1 ต้องมีมุมฉายแสงมากกว่ามุมรับภาพของกล้องในระนาบ YZ ($\beta_{yz} > \theta_{yz}$) ตามรูปที่ 2.53 และในระนาบ XZ ($\beta_{xz} > \theta_{xz}$) ดังแสดงในรูปที่ 2.51
- **ไฟส่องสว่างที่ 2:** เพื่อสร้างเงาตามแนวรอยกริด โดยเงาตามแนวรอยกริดจะเกิดเมื่อเมื่อฉายไฟจากไฟส่องสว่างที่ 2 เพียงตำแหน่งเดียวอธิบายตามรูปที่ 2.54 ขนาดของเงาบนหน้าียงจะสัมพันธ์กับระยะความลึก z_2 และมุมที่แสงตัดผ่านเปลือกยางที่ต่างระดับ β_{tap} ซึ่งถูกกำหนดโดยระยะ d_2 และ d_3 และกำหนดให้มุมฉายแสงของไฟส่องสว่างที่ 2 จะต้องครอบคลุมมุมรับภาพของกล้องเช่นเดียวกับไฟส่องสว่างที่ 1 เนื่องจากไฟส่องสว่างจะอยู่ต่ำกว่ากล้อง ดังนั้นการเลือกระยะ d_2 ยังคงต้องทำให้บริเวณส่วนบนของภาพอยู่ภายในเขตของไฟส่องสว่าง ($y_{lamp_max} > y_{max}$)

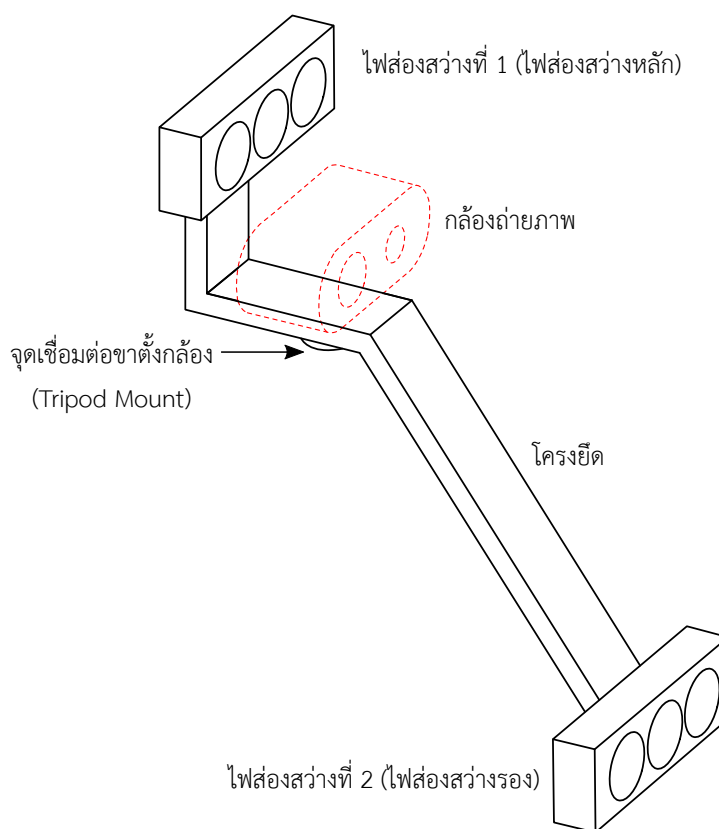


รูปที่ 2.53: การให้แสงในระนาบ YZ ของไฟส่องสว่างที่ 1, ไฟส่องสว่างที่ 2, และมุมรับภาพของกล้องระยะใกล้



รูปที่ 2.54: การให้แสงของไฟส่องสว่างที่ 2 กับหน้ายางในระนาบ YZ เพื่อสร้างเงาตามแนวรอยกรีด

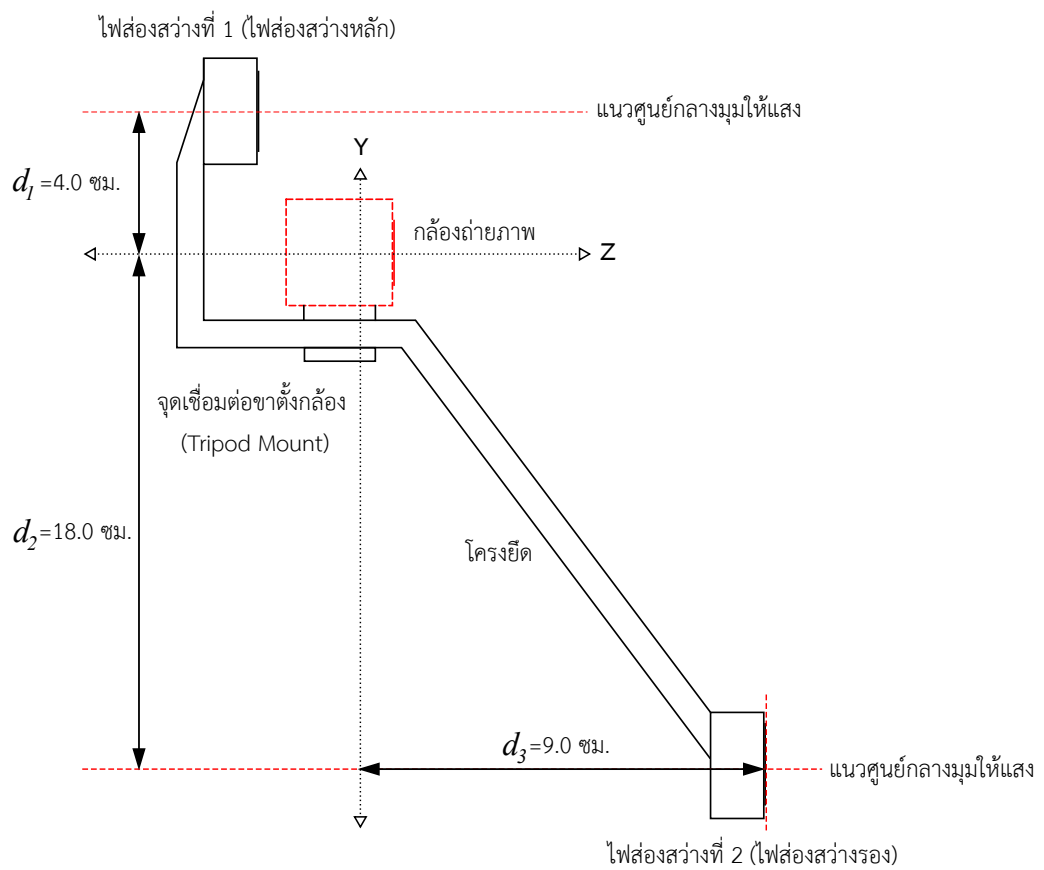
การพัฒนาอุปกรณ์ อุปกรณ์ระบบไฟส่องสว่างตามการออกแบบแสดงในรูปที่ 2.55 ประกอบด้วยไฟส่องสว่าง 2 ตำแหน่ง (ไฟส่องสว่างที่ 1 และไฟส่องสว่างที่ 2) และโครงยึด มีจุดเชื่อมต่อกล้องโมเดล D435 เพื่อใช้งานกับขาตั้งกล้อง (Tripod) โดยได้เลือกใช้หลอดไฟ LED ชนิดเดียวกันให้กับไฟส่องสว่างทั้งสองตำแหน่ง จึงทำให้ไฟส่องสว่างที่ 1 และไฟส่องสว่างที่ 2 มีมุมฉายแสงเท่ากันและครอบคลุมมุมรับภาพของกล้อง ซึ่งสำหรับอุปกรณ์ที่พัฒนาขึ้นมีมุมให้แสง $\beta_{xz} \approx 89^\circ$, $\beta_{yz} \approx 80^\circ$ และวงจรหรี่ไฟต้องเป็นวงจรที่ไม่ใช้หลักการกระพริบ (Non-PWM Dimmer) หลีกเลี่ยงปัญหาการเกิดแถบสีดำในภาพ (Flickering) ภาพแสดงอุปกรณ์ถ่ายภาพเมื่อจัดวางบนขาตั้งกล้อง (Tripod) แสดงในรูปที่ 2.57



รูปที่ 2.55: ภาพวาดอุปกรณ์ระบบไฟส่องสว่างสำหรับถ่ายภาพแนวรอยกิริตระยะใกล้

อุปกรณ์มีระยะห่างระหว่างตำแหน่งกล้องกับตำแหน่งไฟส่องสว่างที่ 1 ตามแกน Y (d_1) มีค่าประมาณ 4.0 เซนติเมตร เนื่องจากต้องเว้นระยะให้กับขนาดของตัวกล้องโมเดล D435 เพื่อไม่ให้ขวางลำแสง, ระยะห่างระหว่างตำแหน่งกล้องจนถึงตำแหน่งไฟส่องสว่างที่ 2 ตามแกน Y (d_2) มีค่าประมาณ 18.0 เซนติเมตร, และระยะห่างตามแกน Z ระหว่างตำแหน่งกล้องถึงตำแหน่งไฟส่องสว่างที่ 2 (d_3) มีค่าประมาณ 9.0 เซนติเมตร ดังแสดงในรูปที่ 2.56

ภาพอุปกรณ์ระบบส่องสว่างพร้อมติดตั้งกล้องระยะใกล้โมเดล D435 จัดวางบนขาตั้งกล้อง แสดงในรูปที่ 2.57 และภาพการจัดวางอุปกรณ์ขณะกำลังถ่ายภาพรอยกิริตแสดงในรูปที่ 2.58 (ก) โดยเมื่อเปิดใช้งานไฟส่องสว่างที่ 1 แสดงในรูปที่ 2.58 (ข) และเมื่อเปิดใช้งานไฟส่องสว่างที่ 2 แสดงในรูปที่ 2.58 (ค)



รูปที่ 2.56: ภาพวาดในระนาบ YZ ของอุปกรณ์ให้แสงเพื่อการถ่ายภาพแนวรอยกริดระยะใกล้ โดยแสดงระยะ d_1 , d_2 , และ d_3



รูปที่ 2.57: อุปกรณ์ถ่ายภาพแนวรอยกรีตระยะใกล้



(ก)



(ข)



(ค)

รูปที่ 2.58: อุปกรณ์ถ่ายภาพแนวรอยกรีตระยะใกล้ขณะถ่ายภาพต้นยางพาราในสภาพแสงน้อย

2.4.1.3 ภาพถ่ายแนวรอยกรีตระยะใกล้

ข้อมูลภาพถ่ายแนวรอยกรีตระยะใกล้ 1 ชุดภาพ เมื่อวางกล้องถ่ายภาพในตำแหน่งในสามมิติหนึ่งๆ ประกอบไปด้วยภาพถ่ายสี RGB จำนวน 2 ภาพจากการให้แสงที่แตกต่างกัน และภาพถ่ายความลึก 1 ภาพ และเป็นเฉพาะภาพถ่ายแนวรอยกรีตที่สมบูรณ์

การถ่ายภาพรอยกรีตจากต้นยางพาราแต่ละต้นจะวางกล้องถ่ายภาพใน 6 ตำแหน่งไม่ซ้ำกัน เพื่อความหลากหลายของมุมมองภาพ โดยกล้องระยะใกล้ได้ถ่ายภาพทั้งหมดจำนวน 1,405 ชุดภาพ จาก 280 หน้ายางในตำแหน่งการวางกล้องระยะใกล้ที่ไม่ซ้ำกัน ภาพถ่าย 1 ชุดภาพ มีรายละเอียดดังต่อไปนี้

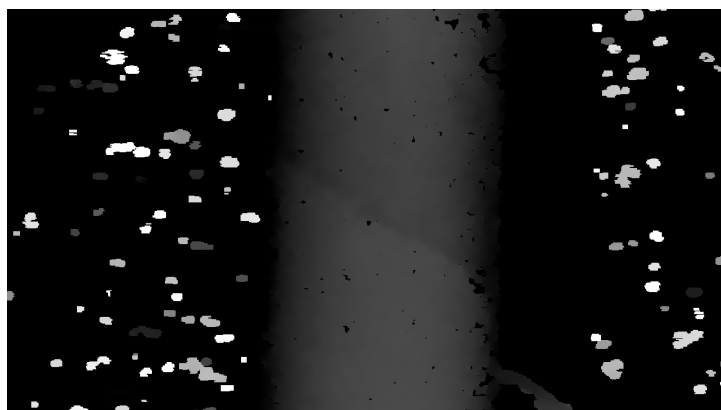
- **ภาพถ่ายสีที่ 1** (ภาพจะถูกแปลงเป็นภาพโทนเทา I_{21}), กล้องถ่ายภาพสีระยะใกล้จะถ่ายภาพถ่ายสีที่ 1 เมื่อฉายไฟบนหน้ายางด้วยไฟส่องสว่างที่ 1 เท่านั้น โดยปรับความสว่างของไฟส่องสว่างที่ 1 ให้อยู่ในช่วง 50-100 ลักซ์ ณ ตำแหน่ง $(x = 0, y = 0, z = z_2)$ ซึ่งเป็นพื้นผิวของหน้ายางบนลำต้น บันทึกภาพเป็นภาพสี RGB ขนาดกว้าง 1280 สูง 720 พิกเซล ขนาดบิตพิกเซล 8 บิต ต่อช่องสัญญาณ ตัวอย่างภาพถ่ายสีที่ 1 แสดงในรูปที่ 2.59
- **ภาพถ่ายสีที่ 2** (ภาพจะถูกแปลงเป็นภาพโทนเทา I_{22}), กล้องถ่ายภาพสีระยะใกล้จะถ่ายภาพถ่ายสีที่ 2 เมื่อฉายไฟบนหน้ายางด้วยไฟส่องสว่างที่ 2 เท่านั้น โดยปรับความสว่างของไฟส่องสว่างที่ 2 ให้อยู่ในช่วง 50-100 ลักซ์ ณ ตำแหน่ง $(x = 0, y = 0, z = z_2)$ บันทึกเป็นภาพสี RGB ขนาดกว้าง 1280 สูง 720 พิกเซล ขนาดบิตพิกเซล 8 บิต ต่อช่องสัญญาณ ตัวอย่างภาพถ่ายสีที่ 2 แสดงในรูปที่ 2.60
- **ภาพถ่ายความลึก** (ภาพโทนเทา I_{D2}), ขนาดความกว้าง 1280 สูง 720 พิกเซล ขนาดบิตพิกเซล 16 บิต ซึ่งค่าพิกเซล (p_{D2}) จะอยู่ในช่วงจำนวนเต็มตั้งแต่ 0 ถึง 65535 โดยบิตพิกเซลในภาพแปลค่าความลึกตั้งแต่ $0-65536 \times 0.1$ มิลลิเมตร จากระนาบ $Z=0$ ของกล้องระยะใกล้ ตัวอย่างภาพถ่ายความลึกแสดงในรูปที่ 2.61 โดยแสดงผลในรูปแบบภาพโทนเทาขนาด 8 บิตพิกเซล



รูปที่ 2.59: ตัวอย่างภาพถ่ายแนวรอยกรีตระยะใกล้, ภาพถ่ายสีที่ 1



รูปที่ 2.60: ตัวอย่างภาพถ่ายแนวรอยกรีตระยะใกล้, ภาพถ่ายสีที่ 2



รูปที่ 2.61: ตัวอย่างภาพถ่ายความลึกแนวรอยกรีตระยะใกล้, (แสดงในรูปแบบ 8 บิต, I_{D2_8})

2.4.2 การกำหนดแนวรอยกรีดอ้างอิงภายในภาพถ่ายระยะใกล้ (Reference Tapping-Line Annotation)

การกำหนดแนวรอยกรีดอ้างอิงในภาพถ่ายระยะใกล้ จะกำหนดตำแหน่งในระดับพิกเซล (Pixel-wise Annotation) สืบเนื่องจากการกำหนดตำแหน่งของรอยกรีดอ้างอิงในภาพจากกล้องถ่ายภาพมุมกว้างด้วยกรอบสี่เหลี่ยมล้อมรอบ และประมาณแนวรอยกรีดด้วยเส้นทแยงมุมของกรอบ ดังนั้นการกำหนดแนวรอยกรีดอ้างอิงในระดับพิกเซลในภาพถ่ายจากกล้องระยะใกล้ จะเพิ่มความละเอียดของเส้นทแยงมุมหลักของกรอบสี่เหลี่ยมล้อมรอบรอยกรีดให้มีความละเอียดของตำแหน่งที่มากขึ้น โดยอาศัยการสร้างรูปหลายเหลี่ยม (Polygon) ด้วยเส้นทแยงมุมและสามเหลี่ยมมุมบนของกรอบสี่เหลี่ยมของรอยกรีด เพื่อใช้เส้นขอบล่าง (“Lower Boundary”) ของพื้นที่รูปหลายเหลี่ยมดังกล่าวแทนแนวรอยกรีดอ้างอิงระดับพิกเซล

เส้นขอบล่างของพื้นที่ย่อยภายในภาพขาวดำใดๆ (R) คือ เซตของพิกเซลที่มีตำแหน่งตามแกนตั้งมากที่สุดในแต่ละคอลัมน์ ซึ่งสามารถเขียนในรูปความสัมพันธ์ระหว่างตำแหน่งแถว (v) และตำแหน่งคอลัมน์ (u) แทนฟังก์ชัน $g(u)$ ในสมการ 2.13 เมื่อกำหนดให้เซต $C_2 : \{(u, v) \in U_2 \times V_2\}$ แทนพิกัดตำแหน่งภายในภาพถ่ายจากกล้องระยะใกล้ (พิกัดภาพ, พิกัดจุด, Pixel Coordinate) ซึ่งมีขนาดกว้าง $w_2 = 1280$, สูง $h_2 = 720$ พิกเซล โดย $U_2 = \{1, 2, 3, \dots, w_2\}$ คือตำแหน่งตามแกนนอนของภาพ, $V_2 = \{1, 2, 3, \dots, h_2\}$ คือตำแหน่งในแกนตั้งของภาพ

$$g(u) := \max_v \{(u, v) \in R\} \quad (2.13)$$

2.4.2.1 การกำหนดแนวรอยกรีดอ้างอิงในภาพถ่ายด้วยมือ (Manual Reference Tapping-Line Annotation)

ผู้กำหนดแนวรอยกรีดอ้างอิงในภาพ (Annotator) สามารถสังเกตลักษณะของพื้นผิวหรือขอบตามแนวรอยกรีดในภาพถ่ายสีได้ดีกว่าภาพถ่ายความลึก หรือได้ดีกว่าภาพถ่ายสีที่ 2 ที่เน้นสร้างเงาบริเวณรอยกรีด ซึ่งอาจเกิดปัญหาของเงาจากการถ่ายภาพ อาทิ ความไม่สมบูรณ์ของเงาหรือความไม่สมบูรณ์ของความต่างระดับผิวเปลือกของแนวรอยกรีด เป็นต้น จึงกำหนดให้ผู้กำหนดแนวรอยกรีดกำหนดตำแหน่งของแนวรอยกรีดในภาพถ่ายสีที่ 1 เท่านั้น ซึ่งตำแหน่งแนวรอยกรีดอ้างอิงที่ได้จะส่งผลกระทบต่อรูปภาพทั้งหมดของชุดภาพนั้น เนื่องจากภาพภายในชุดภาพเดียวกันจะถูกจัดเรียงพิกเซลให้ตรงกัน (Pixel-wise Aligned) ในระบบพิกัดจุดของภาพถ่ายจากกล้องระยะใกล้ (C_2)

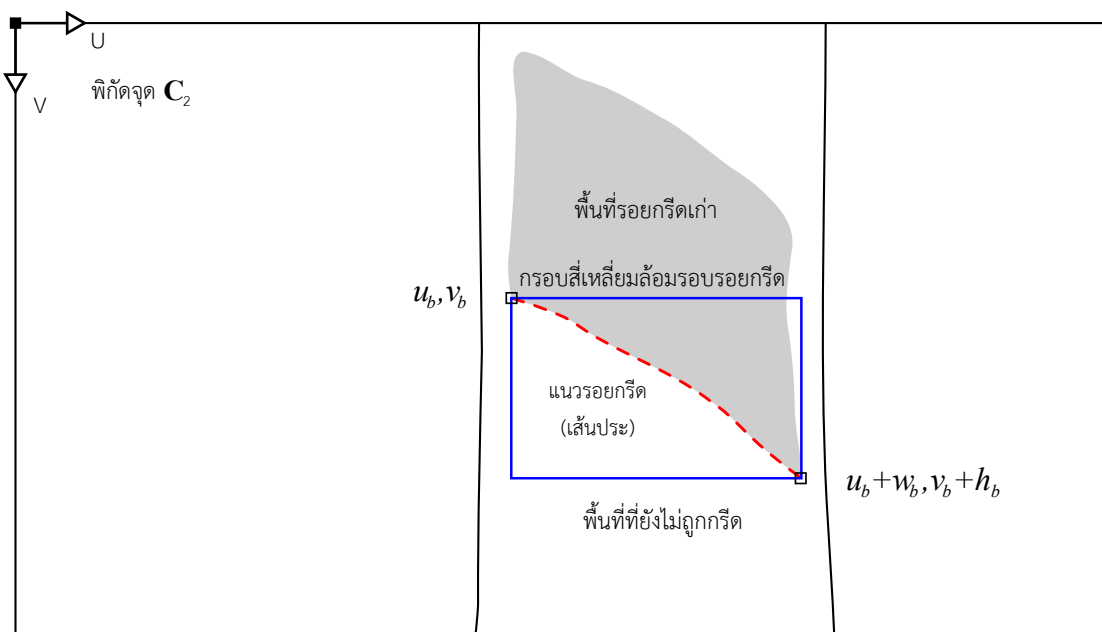
การกำหนดตำแหน่งแนวรอยกรีดอ้างอิงโดยผู้กำหนดแบ่งออกเป็น 3 ขั้นตอน ดังนี้

- 1 สร้างกรอบสี่เหลี่ยมล้อมรอบแนวรอยกรีดในภาพ ($\{u_b, v_b, w_b, h_b\}$) โดยให้มุมซ้ายบน (u_b, v_b) อยู่ในตำแหน่งซ้ายสุดของแนวรอยกรีดที่ปรากฏในภาพ และมุมขวาล่างของกรอบ ($u_b + w_b, v_b + h_b$) อยู่ในตำแหน่งขวาสุดของรอยกรีดที่ปรากฏในภาพ ตัวอย่างดังแสดงในรูปที่ 2.62
- 2 กำหนดพิกเซลเพิ่มในจำนวนที่ผู้กำหนดต้องการ ระหว่างพิกเซล (u_b, v_b) และ ($u_b + w_b, v_b + h_b$) ตามแนวเส้นทแยงมุมหลักที่ประมาณแนวรอยกรีด เพื่อสร้างรูปหลายเหลี่ยม (Polygon) ปิดล้อมพื้นที่สามเหลี่ยมบนของกรอบสี่เหลี่ยม ตามข้อ 1 โดยใช้จุดมุมซ้ายบน, มุมขวาบน, มุมขวาล่าง, และจุดที่ผู้

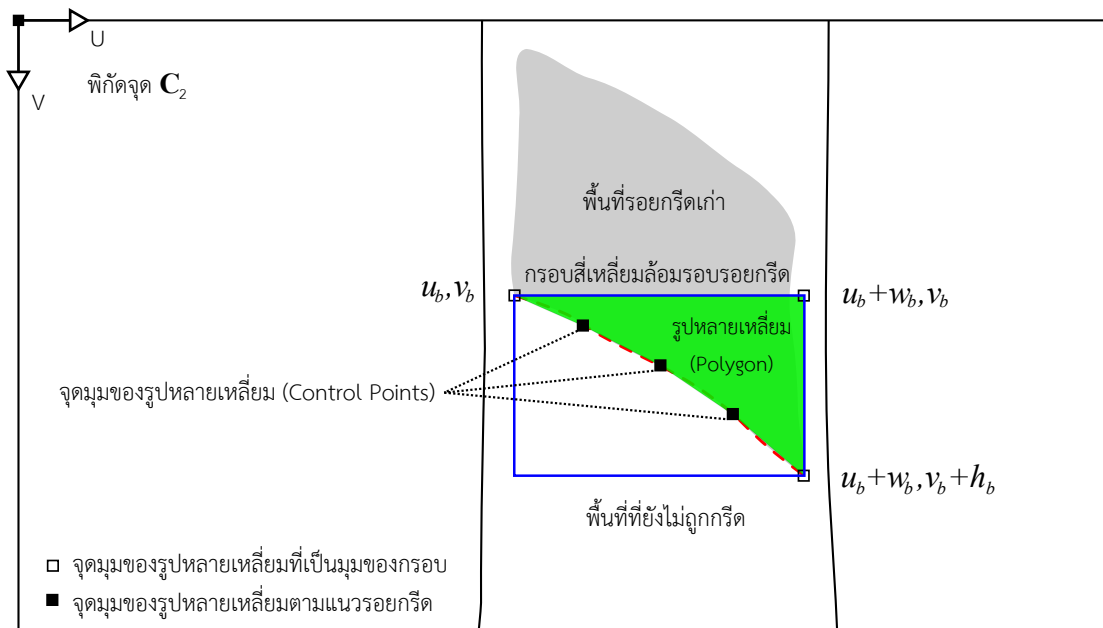
กำหนดตำแหน่งได้กำหนดเพิ่มเติม เป็นจุดมุมของรูปหลายเหลี่ยม (Polygon's Vertices) ดังแสดงในรูปที่ 2.63

- 3 คำนวณหาเส้นขอบล่างของรูปหลายเหลี่ยมในข้อ 2, ซึ่งสามารถนิยามฟังก์ชันของแนวรอยกริดได้ด้วยฟังก์ชัน $g_{\text{manual}}(u)$ ตามสมการ 2.13 ตัวอย่างแสดงในรูปที่ 2.64

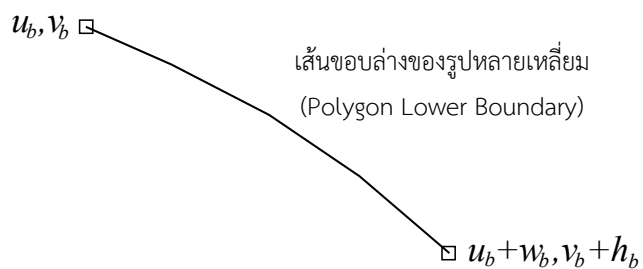
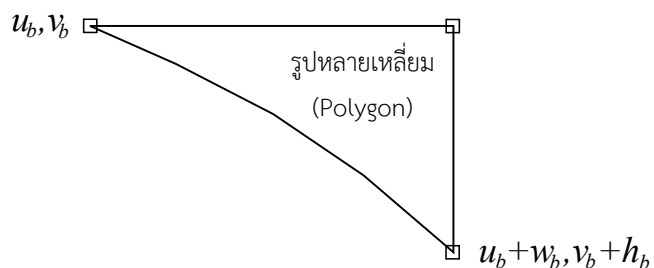
ผลลัพธ์จากวิธีการข้างต้นทำให้แนวรอยกริดอ้างอิง คือเซตของฟังก์ชันภายในภาพ T_{manual} : $\{(u, g_{\text{manual}}(u))\}$ ซึ่งถูกจำกัดอยู่ภายในกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริด, $u_b \leq u \leq u_b + w_b$ และ $v_b \leq g_{\text{manual}}(u) \leq v_b + h_b$, ซึ่งเป็นเส้นขอบล่างของรูปหลายเหลี่ยม



รูปที่ 2.62: ภาพวาดแสดงกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริดจริง (แสดงด้วยเส้นประ) ในภาพถ่ายสีที่ 1 ซึ่งกำหนดให้มุมซ้ายบนของกรอบ (u_b, v_b) อยู่ที่ตำแหน่งแนวรอยกริดซ้ายสุดที่ปรากฏในภาพ และมุมขวาล่างของกรอบ $(u_b + w_b, v_b + h_b)$ อยู่ที่ตำแหน่งขวาสุดของแนวรอยกริดที่ปรากฏในภาพ



รูปที่ 2.63: รูปหลายเหลี่ยมภายในกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริด ซึ่งผู้กำหนดได้กำหนดจุดมุมเพิ่มเติม (Control Points) ให้กับรูปหลายเหลี่ยมตามแนวรอยกริดในระหว่างพิกเซลมุมซ้ายบนและมุมขวาล่างของกรอบสี่เหลี่ยม



รูปที่ 2.64: “เส้นขอบล่าง” (“Lower boundary”) ของรูปหลายเหลี่ยมในรูปที่ 2.63

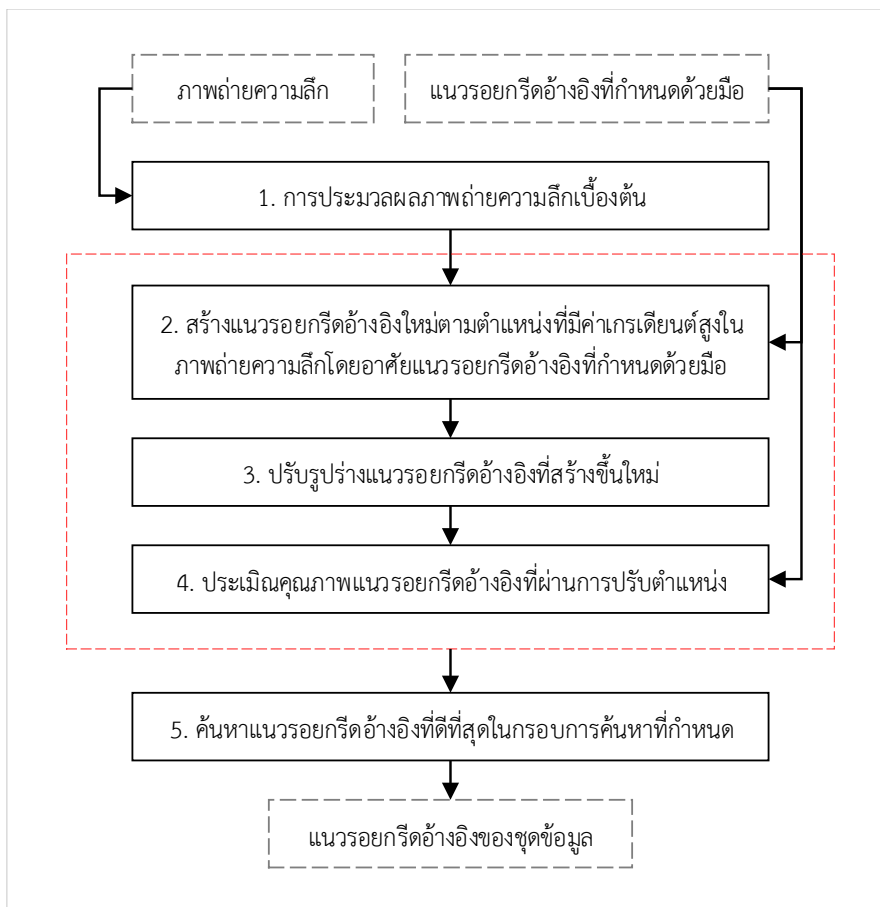
2.4.2.2 วิธีปรับตำแหน่งแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือ (Reference Tapping-Line Annotation Refinement Algorithm)

กำหนดให้ T_{manual} แทนแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือ ซึ่งเป็นลำดับที่มีสมาชิกเป็นตำแหน่งพิกเซลเส้นขอบล่างของรูปหลายเหลี่ยม โดยสมาชิกในลำดับเรียงตามแนวแกนอน ตั้งแต่ u_b ถึง $u_b + w_b$ ซึ่งสมาชิกที่อยู่ติดกันจะมีระยะห่างระหว่างพิกเซลตามแนวแกนอนเท่ากับ 1

$$T_{\text{manual}} : \{(u, g_{\text{manual}}(u))\}; u_b \leq u \leq u_b + w_b \quad (2.14)$$

ความแม่นยำของวิธีการกำหนดแนวรอยกรีดอ้างอิงด้วยมือ จะขึ้นอยู่กับผู้กำหนดตำแหน่งที่จะสามารถสังเกตเห็นขอบภาพตามแนวรอยกรีดได้แม่นยำเพียงใด ส่งผลต่อความแม่นยำของการวางตำแหน่งพิกเซลที่เป็นจุดมุมของรูปหลายเหลี่ยม ทำให้แนวรอยกรีดอ้างอิงที่กำหนดด้วยมือมีความไม่แน่นอน โดยหากเปรียบเทียบกับวิธีการกำหนดตำแหน่งขอบของวัตถุทั่วไปจะถือว่าตำแหน่งกำหนดด้วยมือมีความแม่นยำมากเพียงพอ ที่จะใช้เป็นตำแหน่งอ้างอิงเพื่อวัดผลอัลกอริทึมตรวจหาตำแหน่ง [31] อย่างไรก็ตาม ข้อกำหนดข้างต้นยังเปิดโอกาสให้มีข้อสันนิษฐานเกิดขึ้นได้ สำหรับการกำหนดตำแหน่งแนวรอยกรีดอ้างอิงในภาพถ่ายว่าแนวรอยกรีดนั้นตรงตำแหน่งทั้งในภาพถ่ายสีและภาพถ่ายความลึกได้ดีเพียงพอหรือไม่ ดังนั้นวิทยานิพนธ์นี้จึงนำเสนอขั้นตอนวิธีเพื่อที่จะประเมินคุณภาพของแนวรอยกรีดอ้างอิง และปรับตำแหน่งแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือเดิมให้ดีขึ้น เพื่อที่จะนำไปใช้เป็นแนวรอยกรีดอ้างอิง (Ground Truth Tapping-line หรือ Reference Tapping-line)

การปรับตำแหน่งแนวรอยกรีดจะอาศัยความต่อเนื่อง (Discontinuity) ที่เกิดขึ้นในบริเวณแนวรอยกรีดในภาพถ่ายความลึก เนื่องจากการวางกล่องระยะใกล้ ได้กำหนดมุมรับภาพที่ทำให้ความลึกบริเวณส่วนบนของภาพ และส่วนล่างของภาพมีความลึกที่ใกล้เคียงกัน ดังรูปที่ 2.50 ทำให้ภาพถ่ายความลึกสามารถบันทึกความแตกต่างของระดับความลึกตามแนวรอยกรีดได้อย่างชัดเจน (z_{tap} ในรูปที่ 2.54) ซึ่งคือแนวรอยต่อของพื้นที่บริเวณรอยกรีดเท่ากับพื้นที่ที่ยังไม่ถูกกรีด ดังนั้นตำแหน่งพิกเซลของเส้นขอบล่าง ที่ถูกระบุให้เป็นแนวรอยกรีดในภาพถ่ายสีที่ 1 จะต้องเป็นตำแหน่งที่มีความแตกต่างของระดับความลึก ที่มีค่าสูงกว่าตำแหน่งพิกเซลใกล้เคียง และอาศัยลักษณะความเอียงของแนวรอยกรีดของระบบการกรีดลงแบบเวียนขวาที่แนวรอยกรีดจะเอียงทำมุม 30 ถึง 45 องศา กับแนวระดับพื้น หรือแกนอนของภาพ เนื่องจากการวางตำแหน่งของกล่องที่กำหนดให้ แกน X หรือแกนอนของภาพขนานกับระดับพื้นดิน ทำให้เส้นขอบล่างที่ถูกระบุเป็นแนวรอยกรีดในภาพถ่ายสีทำมุมเอียงในช่วงดังกล่าวเช่นกัน วิธีปรับตำแหน่งแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือแบ่งออกเป็น 5 ขั้นตอนย่อย แสดงตามรูปที่ 2.65 ได้แก่



รูปที่ 2.65: แผนผังขั้นตอนวิธีการปรับตำแหน่งรอยกรีดที่กำหนดด้วยมือและการเลือกแนวรอยกรีดอ้างอิงสำหรับชุดข้อมูล

1 การประมวลผลภาพถ่ายความลึกเบื้องต้น ภาพถ่ายความลึกจากกล้องระยะใกล้เป็นภาพโทนาเทาขนาด 16 บิตพิกเซล (I_{D2}) ซึ่งค่าของพิกเซลจะมีค่าตั้งแต่ 0 ถึง 65535 ไม่เหมาะต่อการประมวลผลหาความแตกต่างตามแนวรอยกรีด (เกรเดียนต์) เนื่องจากช่วงค่าที่กว้างทำให้ความแตกต่างของระดับความลึกตามแนวรอยกรีดไม่ชัดเจน จึงเลือกแปลงเป็นภาพโทนาเทาที่มีขนาดบิตพิกเซล 8 บิต (แทนด้วย I_{D2_8}) โดยค่าพิกเซลในช่วง 0 ถึง 255 จะแปลระยะความลึกในช่วง 20-40 เซนติเมตร ทำให้ค่าความแตกต่างหรือความไม่ต่อเนื่องของความลึกในบริเวณแนวรอยกรีดถูกขยายให้เด่นชัดมากขึ้น

$$I_{D2_8} : \{(u, v, p_{D2_8}(u, v)) \mid (u, v) \in C_2, p_{D2_8} \in \{0, 1, 2, \dots, 255\}\} \quad (2.15)$$

การประมวลผลภาพเพื่อหาความไม่ต่อเนื่องของความลึกบริเวณแนวรอยกรีดในภาพถ่ายความลึก 8 บิตพิกเซล I_{D2_8} จะใช้ฟิลเตอร์โซเบลในแนวแกนตั้ง (Vertical Sobel Filter Kernel) ตามสมการ 2.16 กรองความถี่สูงในภาพถ่ายความลึก $I_{D2_8_hp}$ (High-Pass Spatial Frequency) ตามสมการที่ 2.18 ทั้งนี้กำหนดให้เลือกใช้ฟิลเตอร์กรองความถี่สูงในเฉพาะแนวตั้งของภาพ เนื่องจากส่วนโค้งงอของลำต้นจะส่งผลต่อความความไม่ต่อเนื่องตามแนวแกนนอนของภาพ และส่งผลต่อความไม่ต่อเนื่องบริเวณแนวรอยกรีดเช่นกัน

พื้นที่รอยกริดเก่าเหนือแนวรอยกริดจะมีความลึกมากกว่าพื้นที่ที่ยังไม่ถูกริดซึ่งอยู่ใต้แนวรอยกริด ผลจากฟิลเตอร์โซเบลในแนวแกนตั้งจะทำให้ค่าของเกรเดียนต์ (Gradient Magnitude) ณ ตำแหน่งดังกล่าวมีค่าเป็นบวกในพิกัดภาพ C_2 ทำให้สามารถประมาณตำแหน่งแนวรอยกริดในภาพ คือ ตำแหน่งที่มีเฉพาะค่าเกรเดียนต์เป็นบวกในภาพความถี่สูงของภาพถ่ายความลึก 8 บิต ($I_{D2_8_hp}$ ตามสมการที่ 2.18) โดยเลือกค่า (Thesholding) บิตพิกเซลที่มีค่าเป็นบวกของภาพความถี่สูงของภาพถ่ายความลึก ด้วยสมการที่ 2.19 ผลลัพธ์จะได้ภาพโทนเทาที่แสดงค่าเกรเดียนต์ตามแนวตั้งที่เป็นบวกของภาพถ่ายความลึก 8 บิต ($I_{D2_8_hp}$)

$$S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (2.16)$$

$$I_{D2_8_hp} : \{(u, v, p_{D2_8_hp})\} \quad (2.17)$$

$$p_{D2_8_hp} := p_{D2_8} * S_y = \sum_s \sum_t p_{D2_8}(u-s, v-t) S_y(u, v) \quad (2.18)$$

$$p_{D2_8_hp}(u, v) := \begin{cases} p_{D2_8_hp}(u, v) & ; p_{D2_8_hp}(u, v) \geq 0 \\ 0 & ; p_{D2_8_hp}(u, v) < 0 \end{cases} \quad (2.19)$$

2 สร้างแนวรอยกริดอ้างอิงใหม่ตามตำแหน่งที่มีค่าเกรเดียนต์สูงในภาพถ่ายความลึกโดยอาศัยแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ สร้างแนวรอยกริดขึ้นใหม่ (T_{high}) ด้วยการย้ายตำแหน่งพิกเซลในแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ (T_{manual}) ไปยังตำแหน่งที่มีค่าเกรเดียนต์สูงในภาพของขนาดเกรเดียนต์ตามแนวแกนตั้งของภาพถ่ายความลึก ($I_{D2_8_hp}$) โดยการวนลูปค้นหาตำแหน่งที่ให้ค่าเกรเดียนต์ที่สูงที่สุดในละแวกคอลัมน์ใกล้เคียง (Local Maximum) โดยกำหนดกรอบค้นหา R_{sr} ซึ่งการค้นหาคงทำงานร่วมกับฟิลเตอร์เบลอภาพแบบเกาส์ (Gaussian Smoothing Filter) เพื่อกำจัดจุดยอดอื่นใกล้เคียง (Local Maxima) ในแต่ละรอบของอัลกอริทึมการวนลูป

กำหนดให้ $R_{sr}(T_{\text{manual}}) : \{(u, v, 1)\}$ คือพื้นที่ย่อย (Region) ในภาพขาวดำในระบบพิกัดภาพของกล้องระยะใกล้ หรือเมื่อ $(u, v) \in C_2$ โดยพื้นที่ย่อยดังกล่าวเป็นพื้นที่ที่เกิดจากการขยายตัวจากตำแหน่งพิกเซลใน T_{manual} เป็นระยะ $2sr$ ในแนวคอลัมน์ (แนวตั้ง) โดย sr เป็นจำนวนเต็มบวก

$$R_{sr}(T_{\text{manual}}) := \{(u, v, 1) \mid \forall u \in T_{\text{manual}}, s_{\text{manual}}(u) - sr \leq v \leq s_{\text{manual}}(u) + sr\} \quad (2.20)$$

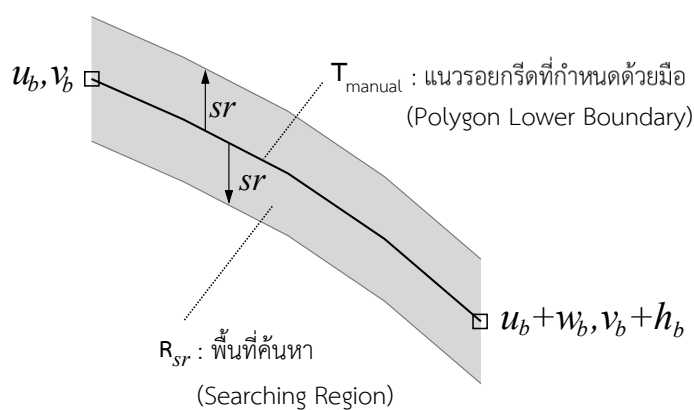
ฟังก์ชัน $T = \text{lm}(I, R_{sr})$ คือ ตัวดำเนินการสำหรับค้นหาตำแหน่งพิกเซลภายในภาพโทนาเทาใดๆ $I : \{(u, v, p(u, v))\}$ ที่เป็นตำแหน่งจุดยอดตามแนวตั้ง (Vertical Local Maxima หรือ Vertical Concave Point) ภายในพื้นที่ R_{sr} ในแต่ละคอลัมน์ และกำหนดให้ G_σ แทนฟิลเตอร์เบลอภาพแบบเกาส์ (2D Square Size Gaussian Kernel) ที่มีค่าการกระจายของน้ำหนัก (Standard Variation) เท่ากับ σ ในทั้งสองแกน ($\sigma = \sigma_u = \sigma_v$)

$$\text{lm}(I, R_{sr}) = \{(u, v) \mid p(u, v - 1) < p(u, v) \text{ และ } p(u, v) > p(u, v + 1)\} \quad (2.21)$$

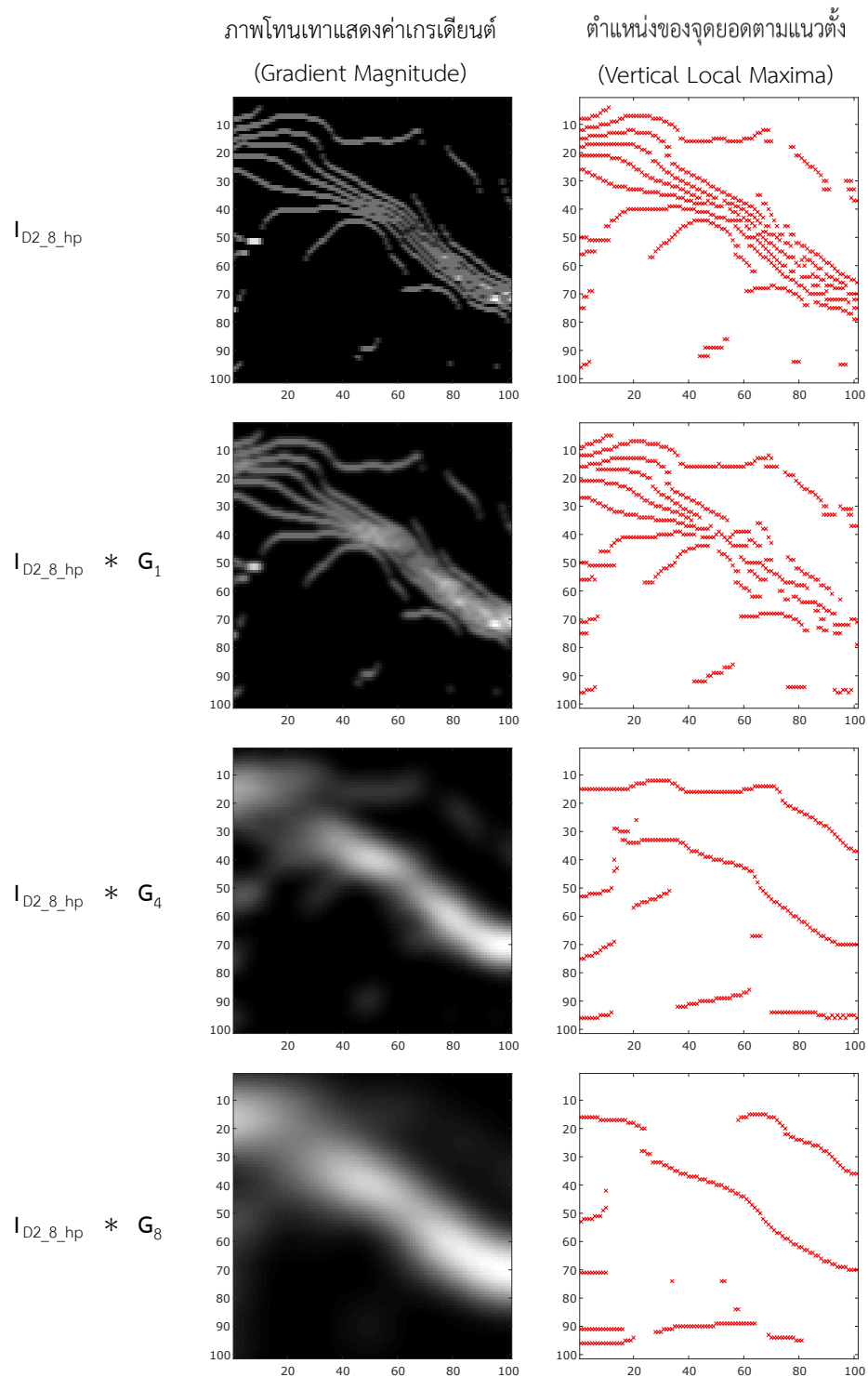
อัลกอริทึม 2.4.2 จะย้ายตำแหน่งแต่ละพิกเซลของ T_{manual} ไปตามแนวคอลัมน์ในขอบเขต $[v_i - sr, v_i + sr]$ จากตำแหน่งเดิม (u_i, v_i) ไปยังตำแหน่งที่มีค่าพิกเซลที่สูงและเป็นจุดยอด (Concave Point) ภายในภาพ $I_{D2_8_hp}$ (ภาพขนาดเกรเดียนต์ของภาพถ่ายความลึก) โดยจะวนลูปเบลอภาพ $I_{D2_8_hp}$ ด้วยฟิลเตอร์เบลอภาพแบบเกาส์ที่ค่า σ เพิ่มขึ้นจากเดิม $\sqrt{2}$ เท่าในแต่ละรอบ จนกระทั่งในขอบเขต $[v_i - sr, v_i + sr]$ เหลือจุดยอดเพียงจุดเดียว แต่เนื่องจากการเพิ่มค่า σ ในแต่ละรอบ อาจทำให้ภายในขอบเขต $[v_i - sr, v_i + sr]$ ของคอลัมน์หนึ่งๆ (u_i) ไม่มีจุดยอด (Non-Concave) หรือเป็นสัญญาณเรียบ ในกรณีนี้จึงกำหนดให้ใช้ตำแหน่งพิกเซลเดิมจากแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ (T_{manual}) ของคอลัมน์นั้นๆ

อัลกอริทึม 2.4.2: การย้ายตำแหน่งพิกเซลภายใน T_{manual} ไปยังตำแหน่งใกล้เคียงที่มีค่าเกรเดียนต์สูงของภาพถ่ายความลึก (Gradient Magnitude) ภายในกรอบการค้นหาที่กำหนด

Input : $sr, l_{D2_8_hp}, T_{\text{manual}}$
Output : T_{high}
Initialize: $R_{sr}(T_{\text{manual}}), \sigma \leftarrow \sigma_0, i \leftarrow \infty$
while $i > 1$ **do**
 $l_{\text{smooth}} \leftarrow \{(u, v, p_{D2_8_hp} * G_\sigma)\}$;
 $T \leftarrow \text{lm}(l_{\text{smooth}}, R_{sr})$;
 A : เซตของจำนวนรวมพิกเซลตามแนวตั้ง ในแต่ละคอลัมน์ (u_i) ของ T ;
 $i \leftarrow \max(A)$;
 $\sigma \leftarrow \sqrt{2}\sigma$
end
 $T_{\text{high}} \leftarrow T$
กำหนดค่าในตำแหน่งคอลัมน์ที่ว่างภายใน T_{high} โดยใช้ค่าเดิมจาก T_{manual}



รูปที่ 2.66: ภาพแสดงพื้นที่กรอบค้นหา R_{sr} ในระนาบพิกัด C_2 ซึ่งเกิดจากการขยายแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ (T_{manual}) ไปตามแกนตั้งเป็นระยะ sr ในทั้งสองทิศทาง



รูปที่ 2.67: ตัวอย่างภาพโทนเทาแสดงค่าเกรเดียนต์ที่เป็นบวกรวมของภาพถ่ายความลึก ($I_{D2_8_hp}$) เมื่อผ่านการทำคอนโวลูชันกับฟิลเตอร์เกาส์ที่มีค่าการกระจายของน้ำหนักต่างกัน โดยแสดงขนาดขยาย 100×100 พิกเซล (Cropped) และตำแหน่งที่เป็นจุดยอดตามแนวตั้ง (Vertical Local Maxima)

3 ปรับรูปร่างแนวรอยกริดอ้างอิงที่สร้างขึ้นใหม่ ปรับรูปร่างของแนวรอยกริดจากข้อ 2 ให้เรียบขึ้น โดยให้มีค่าความชันเฉลี่ยอยู่ในช่วง 30 ถึง 45 องศา กับแนวแกนนอนของภาพ [1] และจำกัดการแกว่งของค่าอนุพันธ์ (Derivatives) ให้อยู่ในช่วงที่แคบ ซึ่งค่าอนุพันธ์สามารถคำนวณได้จากการลบกันของค่าที่อยู่ติดกัน (Finite Difference) จากค่าก่อนหน้า (Backward Difference) การเกลี่ยรูปร่างแนวรอยกริดอ้างอิงที่สร้างขึ้นใหม่ให้เรียบขึ้นอาศัยวิธีการเกลี่ยเส้นโค้ง (Smoothing Spline) และนิยามแนวรอยกริดในภาพจากพิกเซลของเส้นขอบล่างด้วยเส้นโค้งหลายมุม (Polygonal Curve) แบบฟังก์ชันพหุนามดีกรี 3 (Piecewise Cubic Polynomial) [68]

แนวรอยกริดอ้างอิงที่สร้างขึ้นใหม่จากอัลกอริทึม 2.4.2, $T_{\text{high}} : \{(u, g_{\text{high}}(u))\}$ จะถูกปรับรูปร่างด้วยอัลกอริทึม 2.4.3 ให้มีความชันเฉลี่ยอยู่ในช่วง 30 ถึง 45 องศา กับแนวแกนนอนของภาพ แต่เนื่องจากทั้งค่าของ u, g_{high} เป็นค่าของตำแหน่งพิกเซลในโดเมน C_2 ซึ่งเป็นค่าที่ไม่ต่อเนื่อง (Discrete Values) จึงประมาณค่าระหว่างพิกเซลที่มีคอลัมน์ติดกันจำนวน $n_T = |T_{\text{high}}| - 1$ ช่วง ด้วยพหุนามดีกรี 3 (H) ในโดเมนจำนวนจริง

กำหนดให้ $H(x)$ แทนเส้นโค้งหลายมุม (Polygonal Curve) แบบฟังก์ชันพหุนามดีกรี 3 ที่นิยามบน n_T ช่วงของค่า u ใน T_{high} เมื่อ $K = \{1, 2, 3, \dots, n_T\}$ และ $n_T = |T_{\text{high}}| - 1$ ดังสมการ 2.22

$$H(x) : \{h_k(x) = a_k x^3 + b_k x^2 + c_k x + d_k \mid a \neq 0, x \in [u_k, u_{k+1}], \forall k \in K\} \quad (2.22)$$

การเกลี่ยเส้นโค้ง $h(x)$ ใดๆ ในสองมิติให้เรียบ (Smoothing) ด้วยวิธี Smoothing Spline [19] คือการค่าประมาณของฟังก์ชัน $f(x)$ เพื่อทำให้ความแตกต่างระหว่าง $h(x)$ กับ $f(x)$ มีค่าน้อยที่สุด โดยขณะที่ยังต้องทำให้ค่าของ $f(x)$ มีการเปลี่ยนแปลงน้อยที่สุดเช่นกัน หรือกล่าวคือทำให้ $f''(x)$ มีค่าน้อยที่สุด จึงกำหนดให้ฟังก์ชัน $A(f, \lambda)$ เป็นฟังก์ชันของโค้ง $f(x)$ และค่า λ ซึ่ง λ เป็นตัวแปรสำหรับควบคุมความเรียบ (Smoothing Parameter) ของโค้ง $f(x)$ โดย λ จะมีค่าอยู่ในช่วง $[0, \infty]$ โดยค่า $\lambda = 0$ จะทำให้โค้ง $f(x)$ เป็นไปตามโค้งเป้าหมาย $h(x)$ ซึ่งทำให้ $f(x)$ หยาบ (Roughness หรือ No-Smoothing) และหาก $\lambda \rightarrow \infty$ จะทำให้ $f(x)$ เรียบจนเข้าใกล้เส้นตรง ซึ่งการเกลี่ยเส้นโค้งคือการทำให้ฟังก์ชัน A มีค่าน้อยที่สุด (Minimize A) [37]

$$A(f, \lambda) = \left(\sum_i [h(x_i) - f(x_i)]^2 \right) - \lambda \int (f''(x))^2 dx \quad (2.23)$$

อัลกอริทึม 2.4.3 จะเกลี่ยเส้นโค้งฟังก์ชันพหุนามดีกรี 3 ทุกช่วงใน H จากสมการ 2.22 ให้มีค่าความชันแต่ละช่วง $[u_k, u_{k+1}]$ อยู่ในช่วง 30 ถึง 45 องศา กับแนวระดับพื้นดิน เนื่องจากลักษณะของแนวรอยกริดของระบบกริดลงเวียนขวา [1, 23] เอียงเป็นเส้นค่อนข้างตรงโดยมีความชันเฉลี่ย 30 ถึง 45 องศา กับแนวระดับพื้นดิน จึงสามารถอนุมานได้ว่าความชันในแต่ละจุดบนแนวรอยกริดอ้างอิงควรมีค่าอยู่ในช่วง 30 ถึง 45 องศา กับแนวแกนของภาพเช่นกัน

อัลกอริทึม 2.4.3: ปรับรูปร่างแนวรอยกรีด

Input : $T_{\text{high}}, H(x)$
Output : T_{refined}
Initialize: $\lambda \leftarrow \lambda_0, F(x) \leftarrow H(x)$: interpolating spline
while $\forall \{F'(x)\} \notin [\tan 30^\circ, \tan 45^\circ]$ **do**
 $F(x) \leftarrow$ Smoothing Spline : Minimize $A(H, \lambda)$;
 เพิ่มค่า λ ;
end
 $T_{\text{refined}} \leftarrow \{(u, F(u)) \mid \forall u \in T_{\text{high}}\}$

4 ประเมินคุณภาพแนวรอยกรีดอ้างอิงที่ผ่านการปรับตำแหน่ง การประเมินคุณภาพจะเปรียบเทียบคุณภาพของแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือ (T_{manual}) และแนวรอยกรีดอ้างอิงใหม่ที่ปรับแต่งจากวิธีการปรับตำแหน่งแนวรอยกรีด (T_{refined}) ทั้งนี้เพื่อนำแนวรอยกรีดอ้างอิงที่มีคุณภาพมากกว่า ใช้เป็นแนวรอยกรีดอ้างอิงสำหรับชุดข้อมูล (Ground Truth) ซึ่งสามารถแบ่งการพิจารณาคุณภาพของแนวรอยกรีดได้ เป็น 2 ประเด็น ได้แก่ ความเรียบ (Smoothness) และความตรงตำแหน่งที่มีความต่างของระดับของความลึกในภาพถ่ายความลึก (Spatial Localization)

แนวรอยกรีดที่ปรากฏในภาพ คือเซตของพิกเซลที่เรียงต่อกันซึ่งอยู่ในรูปเซตของคู่ลำดับ $T : \{(u, g(u))\}$ ที่มีจำนวนสมาชิกเท่ากับ $n = |T|$ สามารถประมาณเส้นทางของแนวรอยกรีดได้ โดยสร้างเส้นตรงเชื่อมระหว่างจุดที่อยู่ติดกันตามตำแหน่งในแกนนอนในลำดับ ซึ่งคือเส้นโค้งหลายเหลี่ยมแบบเชิงเส้น (Piecewise Linear Curve) จึงทำให้ความเรียบของแนวรอยกรีด (Smoothness) อาจกำหนดได้โดยความตรง (Straightness) ของเส้นโค้งหลายเหลี่ยมแบบเชิงเส้น โดยจะถือว่าแนวรอยกรีดอ้างอิงที่มีความตรงมากกว่ามีคุณภาพมากกว่า ทำให้สามารถใช้ดัชนีชี้วัดความตรง (Straightness Index) [5] เป็นค่าคะแนนชี้วัดความตรงของแนวรอยกรีดอ้างอิง (Straightness Score, SS) ดังสมการ 2.24 ซึ่งแนวรอยกรีดที่มีความตรงมากกว่าจะมีค่าดัชนีชี้วัดความตรงน้อยกว่า

$$SS(T) = \frac{[(u_1 - u_n)^2 + (g(u_1) - g(u_n))^2]^{\frac{1}{2}}}{\sum_{i=1}^{n-1} [(u_i - u_{i+1})^2 + (g(u_i) - g(u_{i+1}))^2]^{\frac{1}{2}}} \quad (2.24)$$

ค่าคะแนนความตรงตำแหน่งของแนวรอยกรีด (Spatial Location Score, LS) คำนวณได้จากผลรวมของค่าพิกเซลภายในภาพ $I_{D2_8_hp}$ หรือคือผลรวมของค่าเกรเดียนต์ของภาพถ่ายความลึก (Gradient Magnitude) ในตำแหน่งที่แนวรอยกรีดพาดผ่าน ดังสมการ 2.25 โดยแนวรอยกรีดที่มีคุณภาพดีกว่าจะคำนวณได้ LS สูง

$$LS(T) = \sum p_{D2_8_hp}(u_i, g(u_i)); \forall (u_i, g(u_i)) \in T \quad (2.25)$$

คะแนนคุณภาพแนวรอยกรีดอ้างอิง (Quality Score, QS) คืออัตราส่วนระหว่างค่าคะแนนความเรียบ (Smoothness Score, SS) และค่าคะแนนความตรงตำแหน่ง (Spatial Location Score, LS) สำหรับแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือ (T_{manual}) และแนวรอยกรีดอ้างอิงที่ถูกปรับตำแหน่ง (T_{refined}) ตามสมการ 2.26 และ 2.27 ตามลำดับ เนื่องจากค่าคะแนนความเรียบจะแปรผกผันกับคุณภาพของแนวรอยกรีด ในขณะที่ค่าคะแนนความตรงตำแหน่งจะแปรผันโดยตรงกับคุณภาพของแนวรอยกรีดอ้างอิง

$$QS(T_{\text{manual}}) = \frac{LS(T_{\text{manual}})}{SS(T_{\text{manual}})} \quad (2.26)$$

$$QS(T_{\text{refined}}) = \frac{LS(T_{\text{refined}})}{SS(T_{\text{refined}})} \quad (2.27)$$

การเปรียบเทียบคะแนนคุณภาพของแนวรอยกรีดอ้างอิง (T_{manual} กับ T_{refined}) ในแต่ละประเด็นคุณสมบัติ จะคำนวณในเชิงอัตราส่วน (Relative Score) โดยกำหนดคะแนนความเรียบสัมพัทธ์ (Relative Smoothness Score, RSS), คะแนนตำแหน่งสัมพัทธ์ (Relative Spatial Location Score, RLS)

$$RSS = \frac{SS(T_{\text{refined}})}{SS(T_{\text{manual}})} \quad (2.28)$$

$$RLS = \frac{LS(T_{\text{refined}})}{LS(T_{\text{manual}})} \quad (2.29)$$

$$RQS = \frac{QS(T_{\text{refined}})}{QS(T_{\text{manual}})} \quad (2.30)$$

5 ค้นหาแนวรอยกรีดอ้างอิงที่ดีที่สุดในรอบการค้นหาที่กำหนด แนวรอยกรีดอ้างอิงที่สร้างขึ้นใหม่จากอัลกอริทึม 2.4.2 และ 2.4.3 ต่างถูกควบคุมด้วยค่า sr ซึ่งเป็นพารามิเตอร์สำหรับกำหนดขนาดของพื้นที่ย่อย (R_{sr}) ในอัลกอริทึม 2.4.2 ในการย้ายตำแหน่งแนวรอยกรีดเดิม ภาพเกรเดียนต์ตามแกนตั้งของภาพถ่ายความลึกจะถูกเบลอด้วยฟิลเตอร์เกาส์ (Gaussian Filter) G_{σ} ที่มีค่า σ เพิ่มมากขึ้นเรื่อยๆ เพื่อทำให้สัญญาณภายใน R_{sr} เป็นสัญญาณที่มีจุดยอดเดียวในแต่ละคอลัมน์ ซึ่งทำให้ค่า σ จะถูกควบคุมด้วยค่า sr ในทางอ้อม เป็นผลให้คุณภาพของแนวรอยกรีดที่ได้ขึ้นอยู่กับค่า sr

กำหนดให้อัลกอริทึม 2.4.4 ทำการค้นหาแนวรอยกรีดที่มีคุณภาพมากที่สุดในช่วงของ sr โดยใช้การประเมินคุณภาพของแนวรอยกรีดอ้างอิงด้วยค่า RSS และ RLS ข้างต้นจากสมการที่ 2.28, 2.29 ตามลำดับ

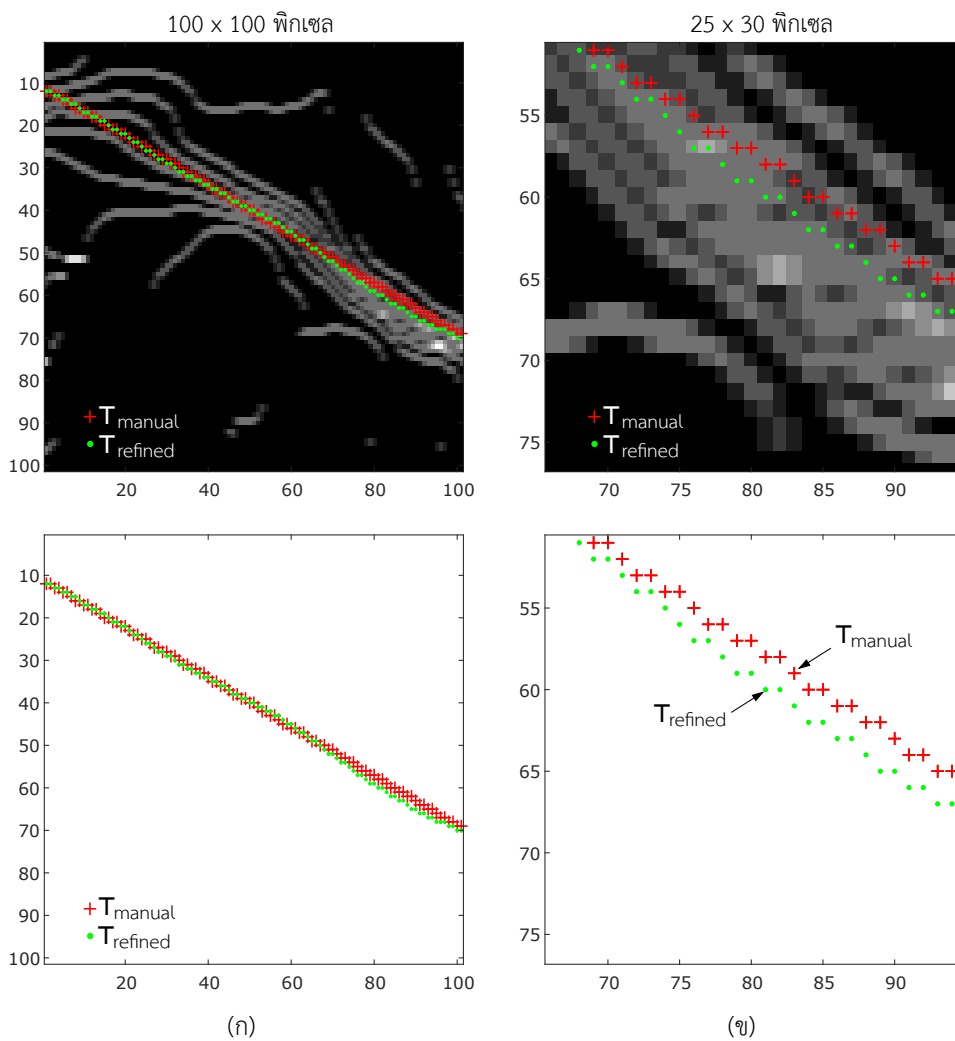
อัลกอริทึม 2.4.4: ค้นหาแนวรอยกรีดที่ดีที่สุดในรอบการค้นหาที่กำหนด

```

Input   :  $sr_{max} \geq 1, T_{manual}$ 
Output  :  $T_{GT}$ 
Initialize:  $RQS_{temp} \leftarrow 1$ 
for  $i \in \{1, 2, 3, \dots, sr_{max}\}$  do
    /* ปรับตำแหน่งแนวรอยกรีด */
     $T_{refined} \leftarrow$  แนวรอยกรีดใหม่ของ  $T_{manual}$  จากอัลกอริทึม 2.4.2 และ 2.4.3
    โดยใช้กรอบการค้นหา  $R_{sr=i}$ 
    /* ประเมินคุณภาพ */
    คำนวณค่า  $RLS_{crnt}, RSS_{crnt}, RQS_{crnt}$  ของแนวรอยกรีด  $T_{manual}, T_{refined}$ 
    /* เลือกแนวรอยกรีด */
    if  $RLS_{crnt} > 1$  and  $RSS_{crnt} \leq 1$  and  $RQS_{crnt} > RQS_{temp}$ 
    then
         $T_{GT} \leftarrow T_{refined}$ 
         $RQS_{temp} \leftarrow RQS_{crnt}$ 
    else
         $T_{GT} \leftarrow T_{manual}$ 
    end
end
  
```

การเลือกแนวรอยกรีดที่กำหนดด้วยมือหรือแนวรอยกรีดอ้างอิงที่ผ่านการปรับตำแหน่งคู่ใดๆ เพื่อใช้เป็นแนวรอยกรีดอ้างอิงสำหรับชุดข้อมูล (T_{GT}) จะอิงตามค่า RSS และ RLS โดยกำหนดเงื่อนไขเมื่อ $RSS < 1.00$ และ $RLS \geq 1$ หรือกล่าวคือหากแนวรอยกรีดอ้างอิงที่สร้างขึ้นใหม่มีตำแหน่งที่ดีกว่าแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือเดิม ($RLS \geq 1$) ในขณะที่แนวรอยกรีดอ้างอิงที่สร้างขึ้นใหม่นั้นมีความตรงมากกว่าแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือ ($RSS < 1.00$) จะเลือกแนวรอยกรีดใหม่ ($T_{refined}$) เป็นแนวรอยกรีดอ้างอิงสำหรับชุดข้อมูล แต่หากคะแนนของแนวรอยกรีดอ้างอิงที่สร้างขึ้นใหม่ไม่ตรงตามเงื่อนไขข้างต้น จะเลือกแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือเดิม (T_{manual}) เป็นแนวรอยกรีดอ้างอิงสำหรับชุดข้อมูล ซึ่งอัลกอริทึม 2.4.4 จะค้นหาแนวรอยกรีดที่ดีที่สุด (RQS สูงที่สุด) ในช่วง $sr = \{1, 2, 3, \dots, sr_{max}\}$

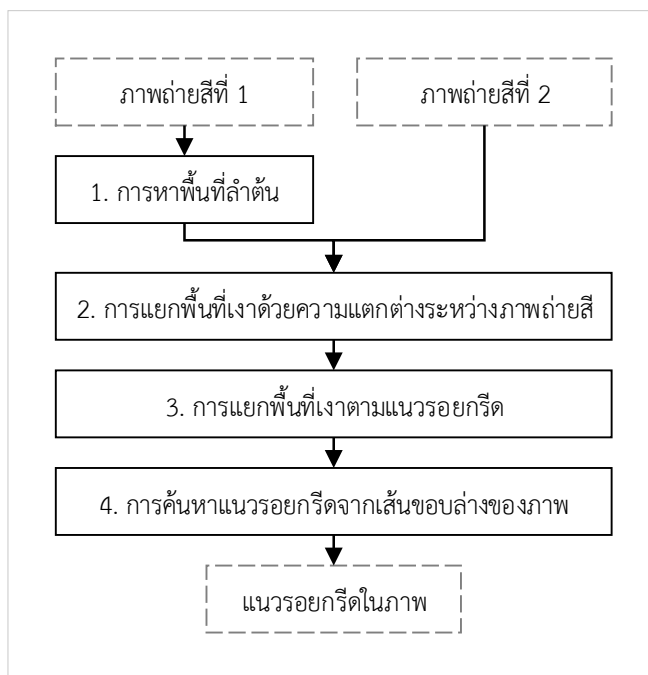
ตัวอย่างภาพแนวรอยกรีดอ้างอิงเปรียบเทียบระหว่างแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือ (T_{manual}) กับแนวรอยกรีดอ้างอิงใหม่จากการปรับตำแหน่ง ($T_{refined}$) โดยภาพโทนาเทาซึ่งแสดงค่าเกรเดียนต์ของภาพถ่ายความลึก ($I_{D2_8_hp}$) ขนาด 100×100 พิกเซล (รูปที่ 2.68 ก) และภาพขยายขนาด 25×30 พิกเซล (รูปที่ 2.68 ข)



รูปที่ 2.68: ภาพตัวอย่างเปรียบเทียบระหว่างแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ (T_{manual}) กับแนวรอยกริดอ้างอิงใหม่จากวิธีปรับตำแหน่ง (T_{refined})

2.4.3 วิธีตรวจหาแนวรอยกริดในภาพถ่ายระยะใกล้

ภาพถ่ายแนวรอยกริดระยะใกล้ภาพสีที่ 2 (I_{22}) มีเงาเกิดขึ้นตามแนวรอยกริด ในขณะที่ภาพสีที่ 1 (I_{21}) ที่ไม่มีเงาตามแนวรอยกริด ภาพสีทั้งสองจะถูกแปลงเป็นภาพโทนเทาเพื่อคำนวณหาความแตกต่างระหว่างภาพ โดยวิธีตรวจหาแนวรอยกริดในภาพถ่ายระยะใกล้ แบ่งออกเป็น 4 ขั้นตอนตามแสดงในรูปที่ 2.69 ดังนี้



รูปที่ 2.69: แผนผังขั้นตอนวิธีตรวจหารอยกริดในภาพถ่ายจากกล้องระยะใกล้

1 การหาพื้นที่ลำต้น (Trunk Segmentation) วิธีการหาพื้นที่ลำต้นของภาพถ่ายกล้องระยะใกล้ จะแตกต่างจากการหาพื้นที่ของลำต้นในภาพด้วยภาพถ่ายความลึกที่ถ่ายจากกล้องถ่ายภาพมุมกว้าง เนื่องจากภาพถ่ายความลึกจากกล้องระยะใกล้โมเดล D435 ได้สร้างสัญญาณรบกวนในบริเวณพื้นหลังมากกว่าโมเดล D415 [27] ส่งผลให้ไม่สามารถใช้ภาพถ่ายความลึกจำกัดพื้นที่ลำต้นโดยตรงได้ ดังนั้นสำหรับภาพถ่ายจากกล้องระยะใกล้ จึงเลือกจำกัดพื้นที่ลำต้นด้วยภาพถ่าย โดยลักษณะของภาพถ่ายในสภาพแสงน้อยเวลากลางคืน จะมีพื้นหลังที่ปรากฏในภาพถ่ายสีระยะใกล้เป็นพื้นที่สีดำชัดเจน เมื่อเทียบกับบริเวณลำต้นที่สามารถสะท้อนแสงจากไฟส่องสว่างได้ จึงสามารถใช้วิธีเลือกค่า (Thresholding) แยกระหว่างพื้นที่ลำต้นกับพื้นหลังได้

กำหนดให้ I_{2_trunk} คือภาพขาวดำของพื้นที่ลำต้นที่เกิดจากการเลือกค่าพิคเซลในภาพโทนเทาของภาพถ่ายสีที่ 1 (I_{21}) ด้วยวิธี Otsu's [59] และ C_{2_trunk} แทนเซตของตำแหน่งพิคเซลที่มีค่าพิคเซลเท่ากับ 1 (Valid Pixel) ในภาพ I_{2_trunk} ดังสมการ 2.31

$$C_{2_trunk} = \{(u, v) \mid \forall (u, v, 1) \in I_{2_trunk}\} \quad (2.31)$$

2 การแยกพื้นที่เงาด้วยความแตกต่างระหว่างภาพถ่ายสี การหาความแตกต่างระหว่างภาพถ่ายสีที่ 1 กับภาพถ่ายสีที่ 2 คำนวณได้จากการลบกันโดยตรงของค่าพิกเซลในตำแหน่งที่ตรงกัน (Pixel-wise Subtraction) โดยหากพิจารณาความสว่าง (Brightness) ในภาพถ่ายสีที่ 2 การกระจายแสงของไฟส่องสว่างตำแหน่งรอง (ไฟส่องสว่างที่ 2) ทำให้บริเวณส่วนล่างของภาพมีความสว่างมากกว่าบริเวณส่วนบนของภาพ เพราะไฟส่องสว่างรองมีตำแหน่งอยู่ใกล้ผิวเปลือกยางมากกว่าไฟส่องสว่างหลัก ในขณะที่ภาพถ่ายสีที่ 1 หน้ายางมีความสว่างสม่ำเสมอ และระดับความสว่างเฉลี่ย (Average Brightness) ในพื้นที่ลำต้นของภาพสีที่ 1 กับภาพสีที่ 2 จะมีค่าต่างกัน จากเหตุผลตัวอย่างข้างต้นเป็นผลให้ การลบกันของค่าพิกเซลโดยตรงไม่สามารถกำจัดพื้นที่ที่ไม่ใช่เงาออกได้หมด ดังนั้นจึงออกแบบให้อัลกอริทึม 2.4.5 กำจัดพื้นที่พื้นหลัง (พื้นที่ที่ไม่ใช่เงา) ออกจากภาพ โดยทำการปรับค่าความสว่างเฉลี่ยภายในพื้นที่ลำต้นของภาพโทนเทาของภาพสีที่ 1 และค่าความสว่างเฉลี่ยภายในพื้นที่ลำต้นของภาพโทนเทาของภาพสีที่ 2 ให้มีระดับความสว่างเฉลี่ยเท่ากัน เพื่อหาความแตกต่างของค่าพิกเซลจากทั้งสองภาพ โดยการลบกันของค่าพิกเซลโดยตรงในบริเวณพื้นที่ลำต้น (C_{2_trunk})

จากนั้นจัดกลุ่มค่าพิกเซลจากผลการลบกันของภาพโทนเทาของภาพถ่ายสีที่ 1 และภาพโทนเทาของภาพถ่ายสีที่ 2 ด้วยวิธีจัดกลุ่มแบบค่ากลาง K กลุ่ม ภายในพื้นที่ลำต้น เพื่อแยกกลุ่มพิกเซลที่มีค่าความแตกต่างของความสว่างในระดับต่างๆ ทั้งนี้หากสังเกตในภาพโทนเทาของภาพถ่ายสีที่ 2 พื้นที่เงาตามแนวรอยกรีด จะมีพื้นที่ที่สว่างมากกว่าห้อมล้อมโดยรอบซึ่งเป็นลักษณะเด่นต่างจากบริเวณอื่นในภาพ โดยภาพผลลัพธ์จากอัลกอริทึม 2.4.5, เป็นภาพขาวดำ (Binary Image, $I_{2_seg_1}$) ซึ่งพื้นที่ที่ไม่ใช่พื้นหลังที่เหลือจะถูกเปลี่ยนค่าให้มีค่าพิกเซลเท่ากับ 1 และค่าพิกเซลของพื้นหลังเท่ากับ 0

อย่างไรก็ตามในภาพถ่ายสีที่ 2 อาจมีพื้นที่เงาอื่นที่ไม่ใช่เงาตามแนวรอยกรีด เนื่องจากอาจมีพื้นที่ที่มีความต่างระดับความลึกอื่นๆ ในหน้ายาง ยกตัวอย่างเช่น ผิวขรุขระของพื้นที่รอยกรีดเก่า หรือสิ่งแปลกปลอมอื่นที่อาจเกาะอยู่บนผิวลำต้น เช่น แมลง, เปลือกยางที่เสื่อมสภาพ, รอยน้ำยางแห้ง เป็นต้น ดังนั้นเงาตามแนวรอยกรีดจะถูกแยกหาด้วยอัลกอริทึม 2.4.6 ในลำดับต่อไป

ตัวอย่างภาพผลลัพธ์ในแต่ละขั้นตอนของอัลกอริทึมแสดงในรูปที่ 2.70 และรูปที่ 2.71 โดยรูปที่ 2.70 (ก) และ 2.70 (ข) แสดงภาพโทนเทาของภาพถ่ายสีที่ 1 และภาพโทนเทาของภาพถ่ายสีที่ 2 เมื่อถูกปรับค่าความสว่างตามลำดับ, ในรูปที่ 2.71 (ก) แสดงผลการลบระหว่างภาพที่ผ่านการปรับค่าความสว่าง, และรูปที่ 2.71 (ข) คือภาพโทนเทาผลลัพธ์ที่ผ่านการเลือกค่าในข้อ 3 ของอัลกอริทึม 2.4.5

อัลกอริทึม 2.4.5: การแยกพื้นที่เงาในภาพถ่ายสีที่ 1 และภาพถ่ายสีที่ 2 จากกล้องระยะใกล้ (กล้อง 2)

Input : ภาพโทนเทาของภาพถ่ายสีที่ 1 และภาพถ่ายสีที่ 2: I_{21}, I_{22}

Output : ภาพขาวดำ: $I_{2_seg_1}$

Initialize: ภาพขาวดำของพื้นที่ลำต้น: C_{2_trunk} ตามสมการ 2.31

- 1 คำนวณความสว่างเฉลี่ยในพื้นที่ลำต้นของแต่ละภาพ, val_1, val_2

$$val_1 = \frac{\sum_{(u,v) \in C_{2_trunk}} p_{21}(u,v)}{|C_{2_trunk}|}$$

$$val_2 = \frac{\sum_{(u,v) \in C_{2_trunk}} p_{22}(u,v)}{|C_{2_trunk}|}$$

- 2 ปรับระดับความสว่างในแต่ละภาพ,

$$p_{21_n} : \{(u, v, p_{21_n})\}, p_{22_n} : \{(u, v, p_{22_n})\}, (u, v) \in C_2$$

$$p_{21_n}(u, v) = p_{21}(u, v) - val_1; \forall (u, v) \in C_2$$

$$p_{22_n}(u, v) = p_{22}(u, v) - val_2; \forall (u, v) \in C_2$$

- 3 คำนวณหาภาพที่เกิดจากความแตกต่างของค่าพิกเซล พร้อมทั้งเลือกขจัดค่าลบให้เท่ากับศูนย์ (Soft Thresholding),

$$I_{2_diff} : \{(u, v, p_{2_diff}(u, v))\}, (u, v) \in C_2 \text{ โดย}$$

$$p_{2_diff}(u, v) = \begin{cases} 0 & ; p_{21_n}(u, v) - p_{22_n}(u, v) < 0 \\ p_{21_n}(u, v) - p_{22_n}(u, v) & ; \text{อื่นๆ} \end{cases}$$

- 4 แบ่งกลุ่มค่าพิกเซล เป็น 3 กลุ่มด้วยวิธีการจัดกลุ่มแบบค่ากลาง K กลุ่ม (K-means) โดยกำหนดค่ากลางเริ่มต้นในแต่ละกลุ่ม (Initial Centroids) ดังนี้

กลุ่มที่ 1: 0

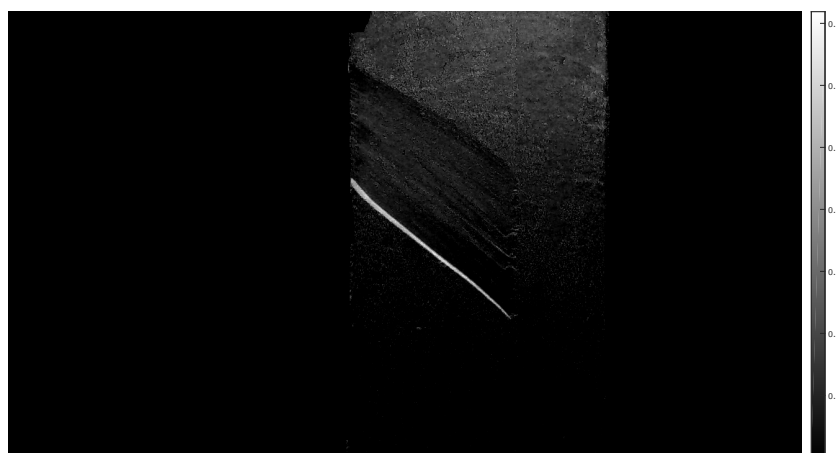
กลุ่มที่ 2: ค่าเฉลี่ยของค่าพิกเซลในพื้นที่ลำต้น

กลุ่มที่ 3: $\max(p_{2_diff}(u, v)), \forall (u, v) \in C_2$

- 5 $I_{2_seg_1}$ เป็นภาพขาวดำที่มีพิกเซลที่มีค่าเท่ากับ 1 (Valid Pixel) เป็นสมาชิกของกลุ่มที่ 3 ของผลจากวิธีการจัดกลุ่มด้วยค่ากลาง K กลุ่ม ตามข้อ 4.
-

(ก) I_{21_n} (ข) I_{22_n}

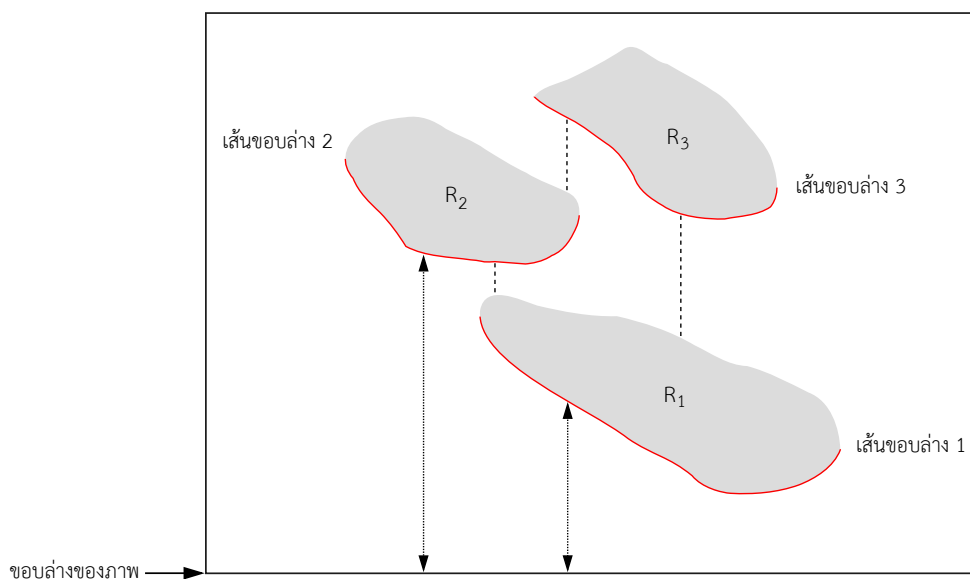
รูปที่ 2.70: ตัวอย่างภาพโทนเทาที่ผ่านการปรับค่าความสว่างให้มีค่าความสว่างเฉลี่ยภายในพื้นที่ลำต้นเท่ากัน ระหว่างภาพโทนเทาของภาพถ่ายสีที่ 1 (I_{21_n}) ในรูป (ก) และภาพโทนเทาของภาพถ่ายสีที่ 2 (I_{22_n}) ในรูป (ข)

(ก) $I_{21_n} - I_{22_n}$ (ข) I_{2_diff}

รูปที่ 2.71: ตัวอย่างภาพผลลัพธ์การลบระหว่างภาพโทนเทาที่ผ่านการปรับความสว่างของภาพโทนเทาของภาพถ่ายสีที่ 1 และภาพโทนเทาของภาพถ่ายสีที่ 2 แสดงในรูป (ก) และภาพผลลัพธ์ของการเลือกค่าตามข้อ 3 ในอัลกอริทึม 2.4.5 (I_{2_diff}) ในรูป (ข)

3 การแยกพื้นที่เงาตามแนวรอยกริด ภาพขาวดำผลลัพธ์จากการแยกพื้นที่เงาด้วยอัลกอริทึม 2.4.5 ($I_{2_seg_1}$) จะถูกประมวลผลเพื่อแยกหาพื้นที่ย่อย (Regions) ด้วยวิธีแยกพื้นที่ตามการเชื่อมต่อกับพื้นที่ใกล้เคียง (Connected Component Labeling [26]) หากสังเกตขนาดและการกระจายตัวของพื้นที่ย่อย ในภาพขาวดำ $I_{2_seg_1}$ พบว่า ลักษณะพื้นที่เงาของแนวรอยกริดจะทอดยาวตามแนวนอนของภาพ ในขณะที่พื้นที่ย่อยที่อยู่ต่ำกว่าพื้นที่เงาของแนวรอยกริด จะมีความหนาแน่น (ขนาดและปริมาณ) น้อยกว่าพื้นที่ที่อยู่เหนือพื้นที่เงาของแนวรอยกริด เนื่องจากลักษณะของภาพถ่ายสีที่ 2 ที่บริเวณส่วนครึ่งล่างของภาพ มีความสว่างโดยเฉลี่ยมากกว่าส่วนครึ่งบน ทำให้เมื่อความแตกต่างของภาพถ่ายสีที่ 1 และ ภาพถ่ายสีที่ 2 ในอัลกอริทึม 2.4.5 ทำให้พื้นที่ครึ่งล่างถูกขจัดออกได้มากกว่าพื้นที่ครึ่งบนของภาพ

อัลกอริทึม 2.4.6 จะอาศัยฟังก์ชันการทำงานย่อย (Subroutine) ทำหน้าที่ประเมินพื้นที่ย่อยทั้งหมด ตามการกีดขวางหรือการบดบังโดยพื้นที่อื่นตามแนวตั้งระหว่างเส้นขอบล่างในทุกพิกเซลของพื้นที่ย่อยนั้น ไปยังขอบล่างของภาพหรือแถวสุดท้ายในภาพ ($v = h_2$) โดยหากพื้นที่ย่อยใดที่มีเส้นขอบล่างถูกบดบังด้วยพื้นที่ย่อยอื่นอย่างสมบูรณ์จะถูกนำออกจากภาพ จากนั้นหมุนภาพตามลำดับแล้วใช้งานการทำงานย่อยเดิม เพื่อนำพื้นที่ที่ถูกบดบังในด้านอื่นๆ ออกจากภาพ โดยอัลกอริทึมสามารถรองรับการแยกพื้นที่เงาของแนวรอยกริดที่เป็นพื้นที่ชนิดต่อเนื่อง (Filled Regions) และชนิดหลายพื้นที่ย่อย (Multiple Regions)



รูปที่ 2.72: ตัวอย่างการประเมินพื้นที่ย่อย (Regions), จากภาพ เส้นขอบล่าง 3 ของพื้นที่ย่อย R_3 ถูกกีดขวางจากขอบล่างของภาพอย่างสมบูรณ์โดยพื้นที่ย่อยอื่นในภาพ ในขณะที่เส้นขอบล่าง 2 ของพื้นที่ย่อย R_2 ถูกกีดขวางจากขอบล่างของภาพเพียงบางส่วนด้วยพื้นที่ย่อย R_1

อัลกอริทึม 2.4.6: วิธีหาพื้นที่ย่อยตามแนวรอยกริด (Tapping Line Shadow Region Segmentation)

```

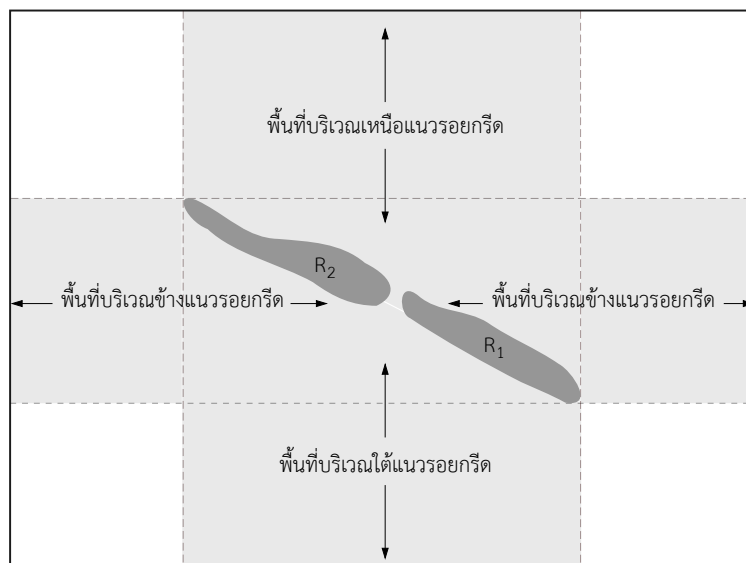
Input: ภาพขาวดำ  $I_{2\_seg\_1}$ 
Output: ภาพขาวดำ  $I_{2\_seg\_2}$ 

/* pass 1 : นำออกพื้นที่ย่อยที่ถูกกีดขวางส่วนบนรูปภาพ */
1  $I_{pass\_1} \leftarrow \text{rmRegion}(I_{2\_seg\_1})$ 
/* pass 2 : นำออกพื้นที่ย่อยที่ถูกกีดขวางส่วนล่างรูปภาพ */
2  $I_{step\_1} \leftarrow \text{imRotate}(I_{pass\_1}, 180)$ 
3  $I_{step\_2} \leftarrow \text{rmRegion}(I_{step\_1})$ 
4  $I_{pass\_2} \leftarrow \text{imRotate}(I_{step\_2}, -180)$ 
/* pass 3 : นำออกพื้นที่ย่อยที่ถูกกีดขวางส่วนข้างลำต้นซ้าย */
5  $I_{step\_3} \leftarrow \text{imRotate}(I_{pass\_2}, 90)$ 
6  $I_{step\_3} \leftarrow \text{rmRegion}(I_{step\_3})$ 
7  $I_{pass\_3} \leftarrow \text{imRotate}(I_{step\_3}, -90)$ 
/* pass 4 : นำออกพื้นที่ย่อยที่ถูกกีดขวางส่วนข้างลำต้นขวา */
8  $I_{step\_4} \leftarrow \text{imRotate}(I_{pass\_2}, -90)$ 
9  $I_{step\_4} \leftarrow \text{rmRegion}(I_{step\_4})$ 
10  $I_{pass\_4} \leftarrow \text{imRotate}(I_{step\_4}, 90)$ 
/* ผลรวมจากแต่ละขั้นตอน (Intersection, Pixelwise Multiplication) */
 $I_{2\_seg\_2} = I_{pass\_2} \cap I_{pass\_3} \cap I_{pass\_4}$ 
/* นิยามการทำงานย่อย */

Function imRotate ( $I, deg$ )
┌    $I_{rot}$  เกิดจากการหมุนภาพ  $I$  เป็นมุม  $|deg|$  องศา เมื่อ
│    $deg > 0$  : กำหนดให้หมุนในทิศตามเข็มนาฬิกา
│    $deg < 0$  : กำหนดให้หมุนในทิศทวนเข็มนาฬิกา ;
└   return  $I_{rot}$ 

Function rmRegion ( $I$ )
┌    $R$  = เซตของพื้นที่ย่อยในภาพ  $I$  คำนวณจาก Connected Component ;
│   ตรวจสอบในแต่ละพิกเซลเส้นขอบล่างของแต่ละพื้นที่ใน  $R$  ว่าถูกกีดขวางจาก
│   ขอบล่างของภาพหรือไม่;
│    $I_{rm}$  = นำพื้นที่ย่อยใน  $R$  ที่มีเส้นขอบล่างทุกพิกเซลถูกกีดขวางออกจากภาพ  $I$ ;
└   return  $I_{rm}$ 

```

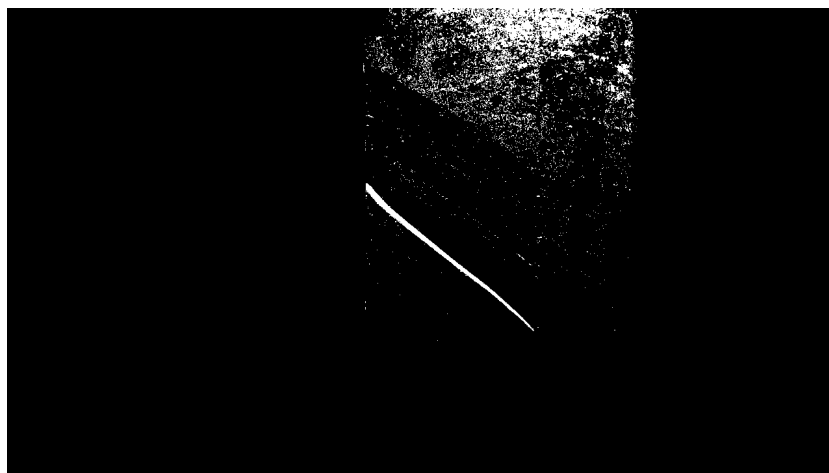
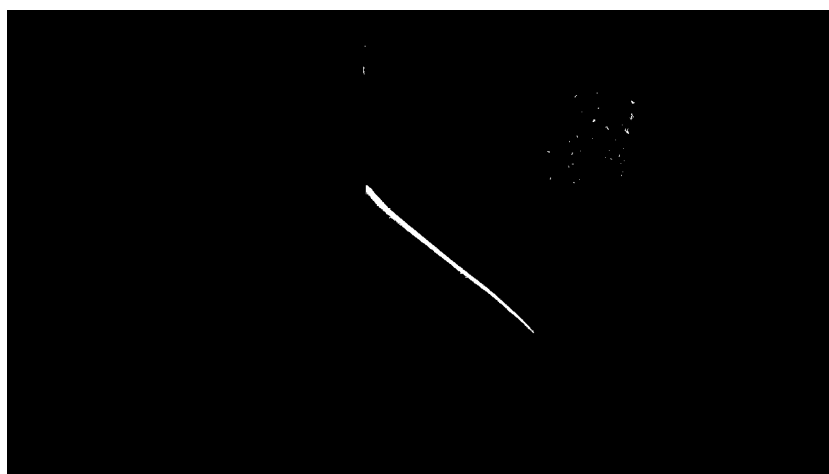


รูปที่ 2.73: ตัวอย่างภาพวาดแสดงลักษณะพื้นที่ย่อยที่เป็นเงาแนวรอยกรีดและพื้นที่บริเวณข้างเคียง

การทำงานของอัลกอริทึม 2.4.6 อธิบายประกอบตามรูปที่ 2.73 แสดงตัวอย่างลักษณะพื้นที่เงาตามแนวรอยกรีด สมมติให้พื้นที่ในกรอบสี่เหลี่ยมใหญ่ที่บ คือพื้นที่ลำดับ ประกอบด้วยพื้นที่ย่อย R_1 และ R_2 เป็นพื้นที่เงาแนวรอยกรีด โดยพื้นที่ย่อยที่อยู่ในเฉพาะพื้นที่ย่อยโดยรอบของเงาแนวรอยกรีด (ภายในพื้นที่แรเงาสั้นประ) ซึ่งล้อมรอบในทิศหลักทั้ง 4 (4-Neighbourhood) จะถูกจัดออกจากภาพตามลำดับได้แก่ พื้นที่ย่อยที่อยู่เหนือพื้นที่ย่อยของเงาแนวรอยกรีด, พื้นที่ย่อยใต้พื้นที่ย่อยของเงาแนวรอยกรีด, และพื้นที่บริเวณด้านข้างทั้งสอง

อัลกอริทึม 2.4.6 สามารถแยกหาพื้นที่เงาของแนวรอยกรีดในภาพ $I_{2_seg_1}$ ได้ โดยกลุ่มของพื้นที่ย่อยที่มีความหนาแน่นสูง (ขนาดและปริมาณ) เหนือพื้นที่เงาของแนวรอยกรีด จะถูกนำออกจากภาพเป็นลำดับแรก ตามด้วยกลุ่มของพื้นที่ย่อยใต้พื้นที่เงาของรอยแนวกรีดที่มีความหนาแน่นต่ำ (ขนาดเล็กและมีระยะห่างจากพื้นที่ย่อยใกล้เคียง) จะถูกนำออกจากภาพ และพื้นที่ย่อยในด้านข้าง ที่มีตำแหน่งแถวในภาพซ้อนทับกับพื้นที่เงาตามแนวรอยกรีดจะถูกนำออกจากภาพในขั้นตอนสุดท้าย ทั้งนี้เป็นเพราะผลของอัลกอริทึมการแยกพื้นที่เงา (อัลกอริทึม 2.4.5) ทำให้พื้นที่เงาตามแนวรอยกรีดมีลักษณะพาดยาวไปตามแกนนอนของภาพ ทำมุมเอียงกับพื้นที่ในแนวตั้งของภาพภายในพื้นที่ลำดับ

ตัวอย่างผลลัพธ์การทำงานของอัลกอริทึม 2.4.6 แสดงในรูปที่ 2.74 เมื่อให้ภาพอินพุตคือภาพขาวดำจากขั้นตอนที่ 2 ($I_{2_seg_1}$) แสดงในรูปที่ 2.74 (ก) และอัลกอริทึมให้ภาพขาวดำเอาท์พุต ($I_{2_seg_2}$) แสดงในรูปที่ 2.74 (ข)

(ก) $I_{2_seg_1}$ (ข) $I_{2_seg_2}$

รูปที่ 2.74: ตัวอย่างภาพขาวดำผลลัพธ์จากอัลกอริทึม 2.4.5 ในรูป (ก) และผลจากอัลกอริทึม 2.4.6 ในรูป (ข)

4 การค้นหาแนวรอยกริดจากเส้นขอบล่างของภาพ ภาพขาวดำผลลัพธ์จากขั้นตอนการแยกพื้นที่เงาของแนวรอยกริด ($I_{2_seg_2}$) พื้นที่ย่อยภายในภาพจะไม่ถูกลบทิ้งจากขอบล่างของภาพ (พื้นที่ย่อยในภาพทั้งหมดมีเส้นขอบล่างที่ไม่ซ้อนทับกัน หรือ เหลื่อมกันเพียงบางส่วนเท่านั้น) สามารถนิยามการค้นหาแนวรอยกริด คือ การค้นหาเส้นขอบล่างของพื้นที่ย่อยที่มาจากพื้นที่เงาของแนวรอยกริด จากเส้นขอบล่างจากพื้นที่ย่อยทั้งหมดภายในพื้นที่ลำดับ

เส้นขอบล่างของภาพ (T_2) คำนวณได้จากวิธีการสแกนคอลัมน์ของรูปภาพในพื้นที่ลำดับ (C_{2_trunk}) โดยเลือกพิกเซลที่มีค่าพิกเซลเท่ากับ 1 (Valid Pixel) ที่มีตำแหน่งแถวสูงสุดในแต่ละคอลัมน์ในภาพ และตำแหน่งพิกเซลของเส้นขอบล่าง สามารถเขียนอยู่ในรูปความสัมพันธ์เชิงฟังก์ชันระหว่างแถว v กับคอลัมน์ u ตามสมการ 2.33

$$T_2 : \{(u, g_2(u))\} \quad (2.32)$$

$$g_2(u) = \max_v \{(u, v)\}; \forall (u, v) \in C_{2_trunk} \quad (2.33)$$

การประมวลผลภาพขาวดำ $I_{2_seg_2}$ เพื่อหาพิกเซลของเส้นขอบล่างข้างต้น อาจทำให้เส้นขอบล่างมีความไม่ต่อเนื่องตามแกนนอนของภาพ ดังในกรณีเช่น มีบางพื้นที่ย่อยอื่นเหลืออยู่ในภาพ ในบริเวณด้านข้างเยื้องบนหรือล่างกับพื้นที่เงาของแนวรอยกริด (นอกพื้นที่แรกในรูปที่ 2.73) หรือพื้นที่เงาของแนวรอยกริดไม่ได้เป็นพื้นที่เดียวกัน (Multiple Regions) ตัวอย่างในรูปที่ 2.73 กล่าวคือ มีบางค่า u ในช่วงปิด $[a, b]$ ที่ทำให้ $g_2(u)$ ไม่มีค่า โดยที่ $b - a > 1$, $g_2(a) \geq 1$ และ $g_2(b) \geq 1$ ดังนั้นจึงประมาณค่าช่วงที่ไม่ต่อเนื่องด้วยเส้นตรงที่เชื่อมระหว่างพิกเซลของเส้นขอบล่างใกล้สุดทางซ้ายของช่วง $(a, g_2(a))$ และพิกเซลของเส้นขอบล่างใกล้สุดทางขวาของช่วง $(b, g_2(b))$ ดังตัวอย่างตามรูปที่ 2.75 จนกระทั่งไม่มีช่วงปิดเหลือใน $\{u\}$ ที่ทำให้ $g_2(u)$ ไม่มีค่า ภายในเส้นขอบล่าง T_2

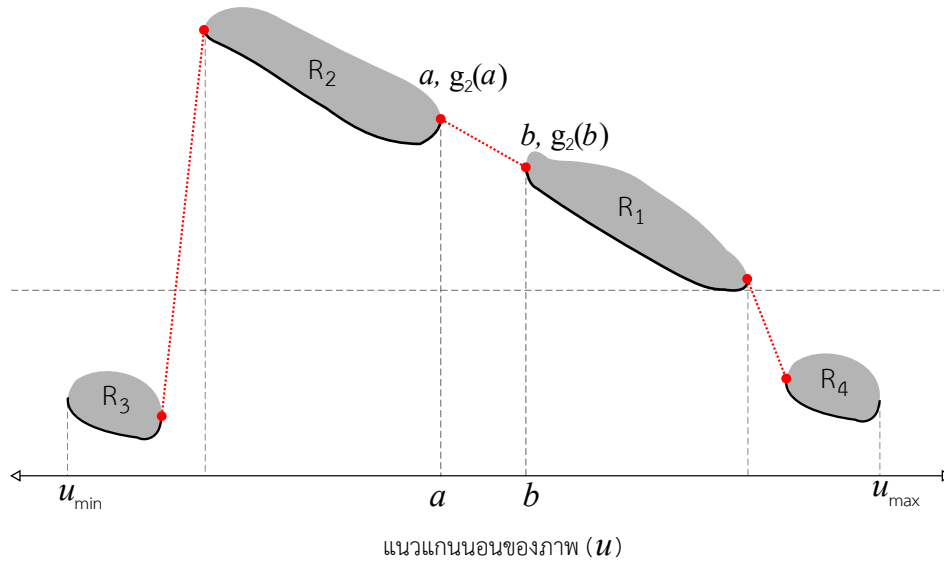
ตัวอย่างตามรูปที่ 2.75 สมมติให้ภาพขาวดำมีพื้นที่ย่อย R_1, R_2 เป็นพื้นที่เงาแนวรอยกริด และ R_3, R_4 เป็นพื้นที่ย่อยอื่นนอก 4 ทิศทางหลักล้อมรอบ (4-Neighbourhood) ของพื้นที่เงาแนวรอยกริด ทั้งนี้เส้นขอบล่างของภาพเมื่อถูกประมาณค่าในช่วงปิด (g_2) จะทำให้ลำดับ $\{u\} \in T_2$ เป็นลำดับเลขคณิต (Arithmetic Sequence, อัตราเพิ่มขึ้นของ u เท่ากับ 1 ตลอดทั้งลำดับ) และอยู่ในช่วงปิด $[u_{\min}, u_{\max}]$ เมื่อ

$$u_{\min} = \min(\{u \mid g_2(u) \geq 1\})$$

$$u_{\max} = \max(\{u \mid g_2(u) \geq 1\})$$

ดังนั้นแนวรอยกริด คือลำดับที่เป็นส่วนหนึ่งของเซต $T_2 : \{(u, g_2(u))\}$ ที่แต่ละพิกเซลที่อยู่ติดกันมีระยะห่างตามแกนนอนของภาพเท่ากับ 1 ($u_{i+1} - u_i = 1$)

การค้นหาแนวรอยกริดเป้าหมาย ($T_{\text{detected}} : \{(u, g_{\text{detected}}(u))\}$) คือการค้นหาลำดับที่มีระยะห่างตามแกนนอนของภาพเท่ากับ 1 ที่มีความยาวมากที่สุดภายใน T_2 และมีรูปร่างที่เรียบ (Smooth)



รูปที่ 2.75: ตัวอย่างการประมาณเส้นตรงในช่วงว่างระหว่างพื้นที่ย่อย

และเอียงทำมุมในช่วง 30 ถึง 45 องศา กับแกนนอนของภาพ ซึ่งเป็นปัญหาแบบการหาค่าที่ดีที่สุดหลายเงื่อนไข (Multi-Objective Optimization) ซึ่งคำตอบต้องพิจารณาจาก 2 ประเด็น ได้แก่ ความยาวของลำดับต้องมากที่สุด ($\max |T_{\text{detected}}|$) และความเรียบซึ่งอาจนิยามด้วยผลรวมของอนุพันธ์อันดับสองกำลังสองของโค้งใดๆ ต้องมีค่าน้อยที่สุด ($\min \sum (g''_{\text{detected}}(u))^2$) อ้างอิงตามวิธีการเกลี่ยเส้นโค้ง (Smoothing Spline) [19] อีกทั้งยังต้องทำให้แนวรอยกริดอยู่ภายใต้เงื่อนไขรูปร่างของแนวรอยกริด ซึ่งทำมุมเอียงเฉลี่ย 30 ถึง 45 องศา กับแนวแกนนอนของภาพ ซึ่งเป็นปัญหาที่อาจมีได้หลายคำตอบ (Pareto Front)

เพื่อลดความซับซ้อนของปัญหาจึงนิยามความเรียบของแนวรอยกริดด้วยการจำกัดการแกว่งของแนวรอยกริด ซึ่งกำหนดให้แนวรอยกริดเป้าหมาย T_{detected} ต้องมีค่าอนุพันธ์อันดับหนึ่งแต่ละตำแหน่งอยู่ในช่วง $[T_{\text{up}}, T_{\text{down}}]$ ทำให้การค้นหาแนวรอยกริดจากเส้นขอบล่าง คือ ค้นหาลำดับที่อยู่ติดกัน และมีความยาวมากที่สุด ใน $g'_2(u)$ และทุกตำแหน่งในลำดับมีค่าในช่วง $[T_{\text{up}}, T_{\text{down}}]$ และค่าเฉลี่ยต้องมีค่าที่จะส่งผลให้ลำดับดังกล่าวทำมุมเอียงเฉลี่ยระหว่าง 30 ถึง 45 องศา กับแกนนอนของภาพ

อนุพันธ์อันดับหนึ่งของ $g_2(u)$ เทียบกับ u คำนวณได้จากการลบกันของค่าที่อยู่ติดกัน (Backward Difference) เมื่อ n_{t2} คือความยาวของ T_2 , $i = 1, 2, 3, \dots, n_{t2}$ ตามสมการ 2.34

$$g'_2(u_i) = g_2(u_i) - g_2(u_{i-1}); \quad g(u_0) := g(u_1) \quad (2.34)$$

อัลกอริทึม 2.4.7: ค้นหาลำดับย่อยติดกันที่ยาวมากที่สุดโดยกำหนดเงื่อนไข

Input : $T_2, \{g'_2(u) | u_{\min} \leq u \leq u_{\max}\}, n_{t2}, T_{up}, T_{down}$
Output : Sequence: T_{detected}
Initialize: Starting Index: $m \leftarrow 1$,
 Target Sequence Length: $n \leftarrow 1$,
 Source Sequence Length: $k \leftarrow n_{t2}$,
 Target Sequence: $q \leftarrow \{g'_2(u) | u_{\min} \leq u \leq u_{\max}\}$

```

for  $i \in \{1, 2, 3, \dots, k - 1\}$  do
  for  $j \in \{j > i\}$  do
     $q_{temp} \leftarrow q(i : j)$ 
     $n_{temp} \leftarrow j - i + 1$ 
     $avg \leftarrow \sum q_{temp} / n_{temp}$ 
    /* ตรวจสอบเงื่อนไข */
    if  $n_{temp} \geq n$  and  $\tan 30^\circ \leq avg \leq \tan 45^\circ$  and
        $T_{up} \leq \forall q \in q_{temp} \leq T_{down}$  then
      |  $n \leftarrow n_{temp}$ 
      |  $m \leftarrow i$ 
    end
  end
end
 $T_{\text{detected}} \leftarrow T_2(m : m + n)$ 

```

อัลกอริทึม 2.4.7 จะค้นหาลำดับย่อยภายใน $\{g'_2(u) | u_{\min} \leq u \leq u_{\max}\}$ ที่มีสมาชิกติดกันที่ยาวมากที่สุด โดยกำหนดเงื่อนไขการค้นหา คือ

1. ลำดับเป้าหมายมีค่าเฉลี่ยในช่วง $[\tan 30^\circ, \tan 45^\circ]$
2. ทุกๆ สมาชิกภายในลำดับเป้าหมายมีค่าอยู่ในช่วง $[T_{up}, T_{down}]$

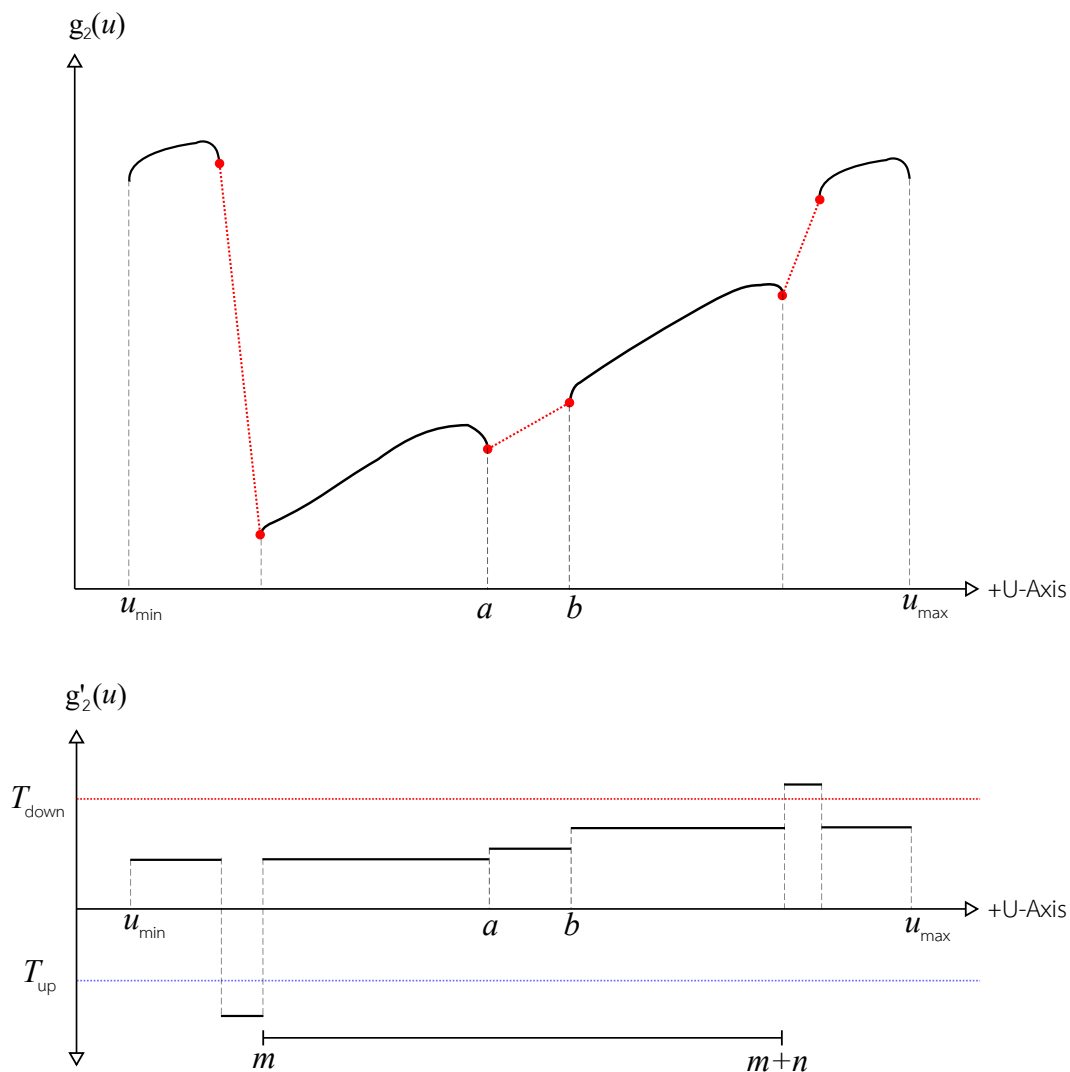
ทำให้การค้นหาแนวรอยกริดจะถูกควบคุมด้วย 2 พารามิเตอร์ได้แก่ T_{up}, T_{down} ทั้งนี้การเลือกค่าของสองตัวแปรเป็นจำนวนเต็มที่ไม่กำหนดตายตัว ทำให้ T_{up}, T_{down} เป็นตัวแปรต้นสำหรับการทดลอง

T_{up} ควบคุมการแกว่งของแนวรอยกริดในทิศทางขึ้น (Maximum-Allowed Upward Direction Derivative)

T_{down} ควบคุมการแกว่งของแนวรอยกริดในทิศทางลง (Maximum-Allowed Downward Direction Derivative)

ตัวอย่างในรูปที่ 2.76 กราฟแสดงเส้นขอบล่างจากรูปที่ 2.75 และกราฟอนุพันธ์อันดับหนึ่งของเส้นขอบล่าง (โดยประมาณอย่างง่าย) ทั้งนี้แสดงตัวอย่างการเลือกค่า T_{up}, T_{down} ที่ทำให้อัลกอริทึม 2.4.7 สามารถหาเส้นขอบล่างของพื้นที่เงาแนวรอยกริดได้ครบ ซึ่งคือช่วง $m \leq u \leq m + n$

ภาพตัวอย่างผลลัพธ์ของวิธีการตรวจหารอยกริดทั้งหมดแสดงในรูปที่ 2.77 ซึ่งเป็นภาพถ่ายของภาพถ่ายสีที่ 1 ที่มีแนวรอยกริดเป็นแนวพิกเซลภายใน



แนวแกนของภาพ (u)

รูปที่ 2.76: กราฟแสดงเส้นขอบล่างของรูปที่ 2.75 และอนุพันธ์อันดับหนึ่งของ $g(u)$ เทียบกับค่า u แทนด้วย $g'(u)$ และแสดงการเลือกค่า T_{up}, T_{down} ที่สามารถแยกเส้นขอบล่างที่เป็นแนวรอยกรีด ออกจากเส้นขอบล่างทั้งหมดได้ถูกต้อง



รูปที่ 2.77: ตัวอย่างผลการตรวจหาแนวรอยกรีดในภาพถ่ายระยะใกล้ เมื่อแสดงแนวรอยกรีด (แนวจุดพิกเซลสีขาว) แบบซ้อนทับกับภาพขยายของภาพถ่ายสีที่ 1

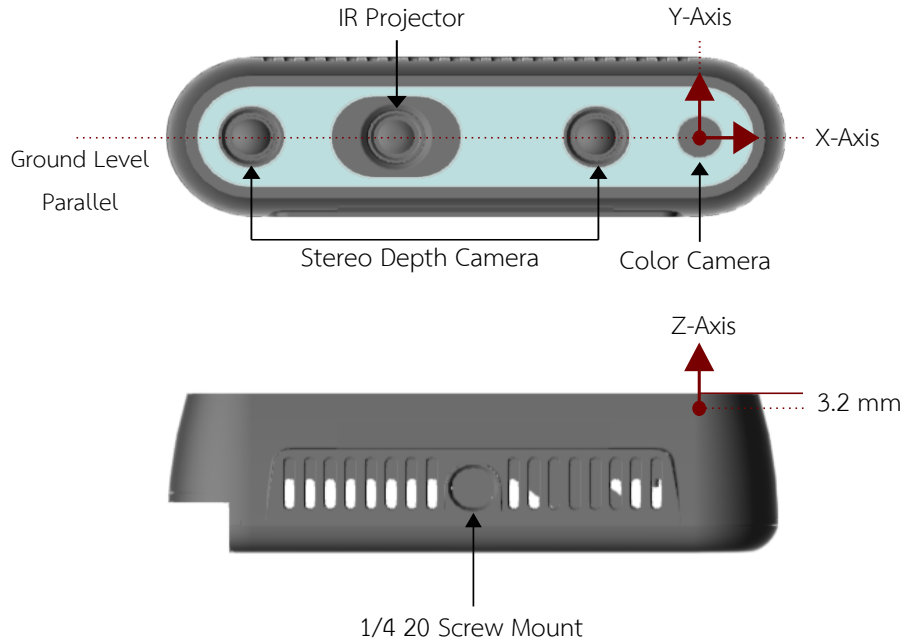
2.4.4 ตำแหน่งแนวรอยกรีดในพิกัดสามมิติอ้างอิงกล้องระยะใกล้ (กล้อง 2)

พิกัดภาพของกล้องระยะใกล้อยู่ในระบบพิกัดจุด C_2 สามารถแปลงไปเป็นตำแหน่งในระบบพิกัดฉากสามมิติในกรอบอ้างอิงของกล้องระยะใกล้ได้ โดยอาศัยค่าพารามิเตอร์สำหรับแปลงพิกัด (Intrinsic Parameters) ซึ่งกำหนดให้ระบบพิกัดภาพของภาพถ่ายระยะใกล้ทั้งหมดจะอ้างอิงกับโมดูลกล้องถ่ายภาพสีของกล้องโมเดล D435

อ้างอิงจากเอกสารกำกับกล้อง [10] ได้กำหนดจุดกำเนิดของกรอบอ้างอิงกล้องถ่ายภาพสีในตำแหน่งกึ่งกลางหน่วยรับภาพของกล้องถ่ายภาพสี และมีระนาบความลึกศูนย์ (Depth Start Point, $z = 0$) ณ ระนาบถัดจากกระจกหน้าเลนส์ (Cover Glass) เข้าไปในตัวกล้องเป็นระยะ 3.2 มิลลิเมตร (รูปที่ 2.78) สำหรับภาพถ่ายสี RGB ที่ความละเอียด 1280×720 พิกเซล มีค่าพารามิเตอร์สำหรับแปลงพิกัด (Intrinsic Parameters) ได้แก่ ทางยาวโฟกัสในแนวแกน X (f_x), ทางยาวโฟกัสในแนวแกน Y (f_y), และตำแหน่งจุดสำหรับฉาย (Projection Points, pp_x, pp_y) ดังนี้

$$f_x = 923.531, f_y = 923.956$$

$$pp_x = 642.823, pp_y = 352.496$$



รูปที่ 2.78: กรอบอ้างอิงของกล้องโมเดล D435 (กล้องระยะใกล้) มีจุดกำเนิดที่ตำแหน่งกึ่งกลางของกล้องถ่ายภาพสี (Color Camera), โมเดลสามมิติของกล้องจาก [11]

ภาพถ่ายความลึกจากกล้องระยะใกล้ (I_{D2}) เป็นภาพโทนเทา 16 บิต ค่าของพิกเซลถูกแมปกับระยะความลึกโดยตรงในหน่วย $\times 0.1$ มิลลิเมตร ดังนั้นตำแหน่งพิกเซลแนวรอยกริด $(u, v) \in T_{\text{detected}}$ ในระบบพิกัดภาพที่มีค่าความลึกเท่ากับ $p_{D2}(u, v)$ สามารถแปลงไปยังตำแหน่งในพิกัดสามมิติ (x, y, z) ในหน่วยมิลลิเมตร ตามสมการ 2.35

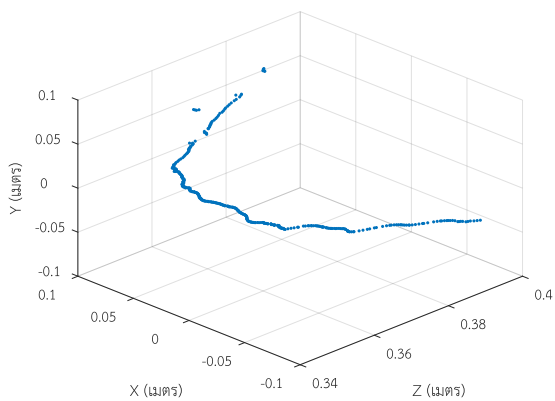
แนวรอยกริดจากนิยามเส้นขอบล่าง จะมีเพียงพิกเซลเดียวในแต่ละคอลัมน์ในภาพ ทำให้เส้นทางของแนวรอยกริดในภาพ จะมีลำดับเรียงตามค่าตำแหน่งตามแกนของภาพ ทำให้ตำแหน่งพิกเซลของแนวรอยกริดถูกแปลงไปยังพิกัดสามมิติ เส้นทางของแนวรอยกริดในระบบพิกัดสามมิติจะเรียงลำดับเหมือนใน T_{detected} เช่นกัน

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{f_x} & 0 & 0 \\ 0 & \frac{1}{f_y} & 0 \\ 0 & 0 & 0.1 \end{bmatrix} \begin{bmatrix} u \\ v \\ p_{D2}(u, v) \end{bmatrix} + \begin{bmatrix} -\frac{pp_x}{f_x} \\ -\frac{pp_y}{f_y} \\ 0 \end{bmatrix} \quad (2.35)$$

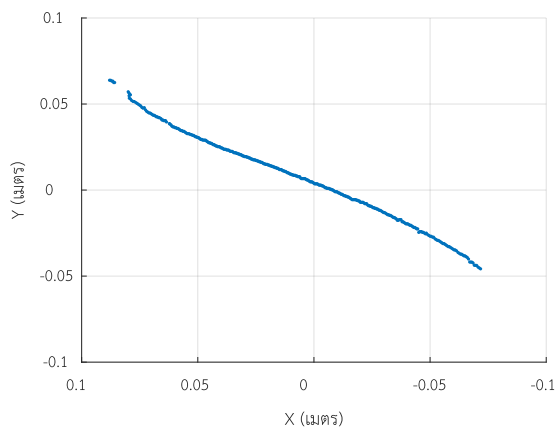
ตัวอย่างแนวรอยกริดในภาพ (รูปที่ 2.79 ก) เมื่อแปลงเป็นตำแหน่งในพิกัดสามมิติของกล้องระยะใกล้แสดงในรูปที่ 2.79 ข) และแสดงผลในระนาบสองมิติในรูปที่ 2.79 ค) , 2.79 จ) , และ 2.79 ง)



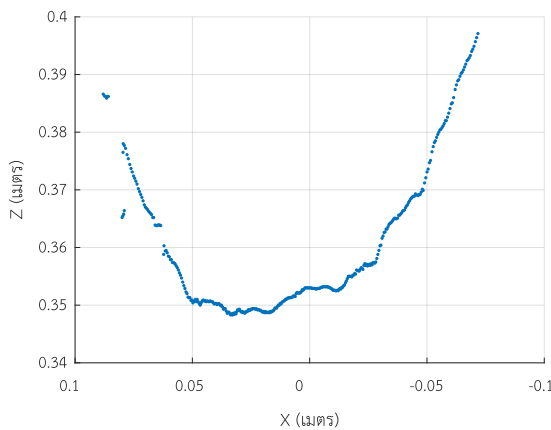
(ก) แนวรอยกรีดในภาพจากวิธีตรวจหา



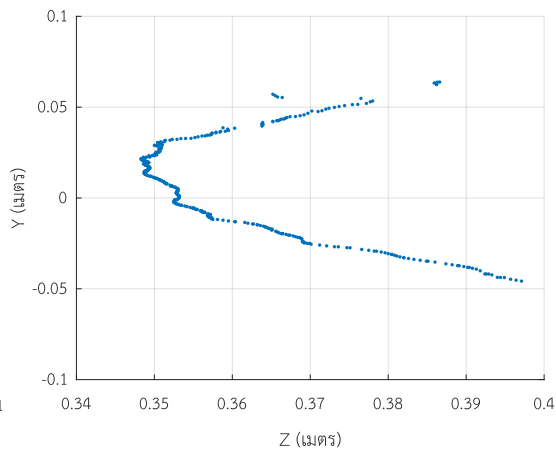
(ข) แนวรอยกรีดในปริภูมิสามมิติ



(ค) แนวรอยกรีดในระนาบ XY



(ง) แนวรอยกรีดในระนาบ XZ



(จ) แนวรอยกรีดในระนาบ YZ

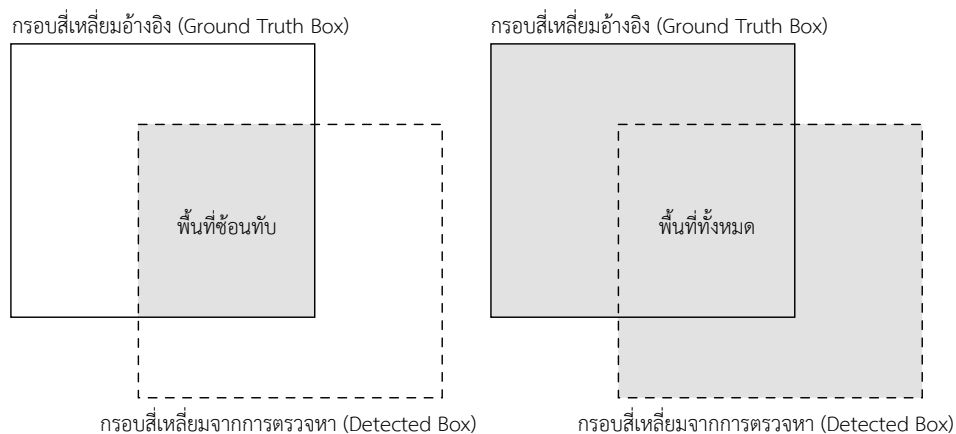
รูปที่ 2.79: แผนภาพแสดงตำแหน่งแนวรอยกรีดในพิกัดสามมิติของกล้องระยะใกล้ โดยค่า XYZ ในหน่วยเมตร

2.5 การวัดผลการตรวจหารอยกริด, ถ้วยรอน้ำยาง, และแนวรอยกริดในภาพ

2.5.1 การตรวจหารอยกริดและถ้วยรอน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้าง

2.5.1.1 อัตราส่วนการซ้อนทับระหว่างกรอบสี่เหลี่ยม (Intersection-Over-Union, IoU)

ตำแหน่งรอยกริดและถ้วยรอน้ำยางในภาพถ่ายถูกกำหนดด้วยกรอบสี่เหลี่ยมล้อมรอบ (Bounding Box) ทำให้สามารถวัดผลการตรวจหาด้วยอัตราส่วนการซ้อนทับระหว่างกรอบสี่เหลี่ยมอ้างอิง (Ground Truth Bounding Box) กับกรอบสี่เหลี่ยมล้อมรอบจากการตรวจหา (Detected Bounding Box) ด้วยค่าอัตราส่วนการซ้อนทับที่มากกว่าหรือเท่ากับ 0.5 เพื่อใช้บ่งชี้ว่ากรอบสี่เหลี่ยมล้อมรอบจากการตรวจหาตรงตำแหน่งกับกรอบสี่เหลี่ยมอ้างอิง หรือ “ตรวจพบ” [17] แต่ทั้งนี้ค่าอัตราส่วนการซ้อนทับสำหรับบ่งชี้ผลการตรวจหาไม่ได้มีค่าที่กำหนดแน่นอน โดยหากกำหนดให้ค่าเข้าใกล้ 1 มากขึ้น จะบ่งชี้ว่าตำแหน่งของกรอบสี่เหลี่ยมล้อมรอบจากการตรวจหา นั้นใกล้เคียงกับกรอบสี่เหลี่ยมอ้างอิงได้มากขึ้น (Well-localized) [33]



รูปที่ 2.80: การซ้อนทับของกรอบสี่เหลี่ยมอ้างอิงกับกรอบสี่เหลี่ยมจากการตรวจหาวัตถุ

2.5.1.2 ความแม่นยำเฉลี่ย

ค่าความแม่นยำเฉลี่ย (Average Precision, AP) เป็นค่าความแม่นยำต่อการตรวจหา 1 ชนิดวัตถุ ในชุดข้อมูลหนึ่งๆ โดยสำหรับชุดข้อมูลภาพถ่ายจากกล้องถ่ายภาพมุมกว้างแต่ละภาพ จะมีรอยกริดและถ้วยรอน้ำยางอยู่เสมอ ทำให้ค่าความแม่นยำเฉลี่ยเป็นไปตามสมการ 2.36

$$AP = \frac{\text{จำนวนการตรวจพบของวัตถุชนิดหนึ่งๆ}}{\text{จำนวนภาพในชุดข้อมูล}} \quad (2.36)$$

2.5.2 การวัดตำแหน่งของแนวรอยกริดในภาพถ่ายจากกล้องระยะใกล้

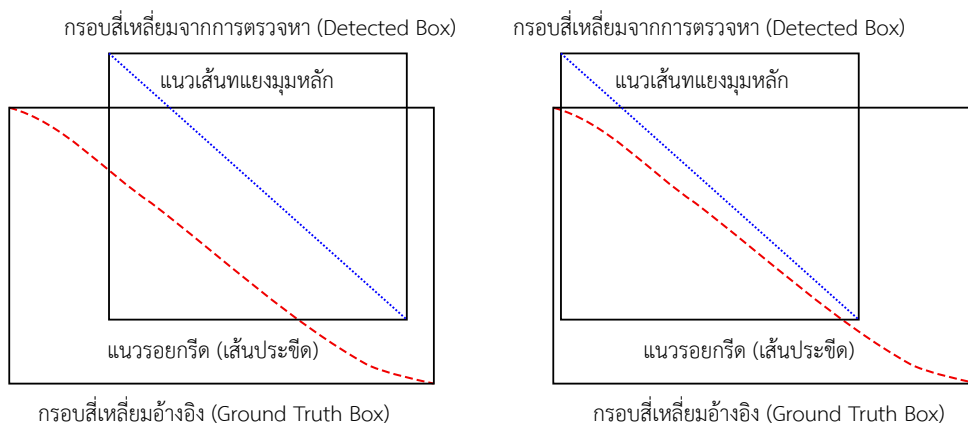
ปัญหาการตรวจหาแนวรอยกริดในภาพจากกล้องระยะใกล้ แบ่งออกเป็น 2 ปัญหาย่อย ได้แก่

1. การรู้จำแนวรอยกริดซึ่งเป็นการตอบคำถามตรวจพบหรือตรวจไม่พบแนวรอยกริด
2. ระยะคลาดเคลื่อนของตำแหน่งแนวรอยกริดที่ตรวจพบ

ซึ่งการกำหนดตำแหน่งแนวรอยกริดด้วยกรอบสี่เหลี่ยมล้อมรอบและกลุ่มของพิกเซลเส้นขอบล่าง ตามข้อ 2.4.2 รองรับการตอบปัญหาทั้ง 2 ประเด็นดังกล่าว โดย

2.5.2.1 การรู้จำแนวรอยกริด

การรู้จำแนวรอยกริดสามารถวัดผลได้ด้วยค่าความแม่นยำด้วยกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริด โดยอาศัยอัตราส่วนการซ้อนทับระหว่างกรอบ ได้เช่นเดียวกับการวัดผลการตรวจหาภาพจากกล้องถ่ายภาพมุมกว้าง ตามสมการ 2.36 เพื่อบ่งชี้ผลของการตรวจพบแนวรอยกริดได้เท่านั้น เพราะกรอบสี่เหลี่ยมไม่ได้ให้ข้อมูลตำแหน่งของพิกเซลตามแนวรอยกริดที่แน่นอน แม้ว่าจะประมาณแนวรอยกริดด้วยแนวเส้นทแยงมุมหลักของกรอบ ดังกรณีในรูปที่ 2.81 กรอบสี่เหลี่ยมของการตรวจหา (Detected Box) และกรอบสี่เหลี่ยมอ้างอิง (Ground Truth Bounding Box) ทั้งสองกรณีมีค่าอัตราส่วนการซ้อนทับที่เท่ากัน แต่ระยะห่างระหว่างแนวรอยกริดกับแนวเส้นทแยงมุมหลักของกรอบสี่เหลี่ยมของการตรวจหาไม่เท่ากัน



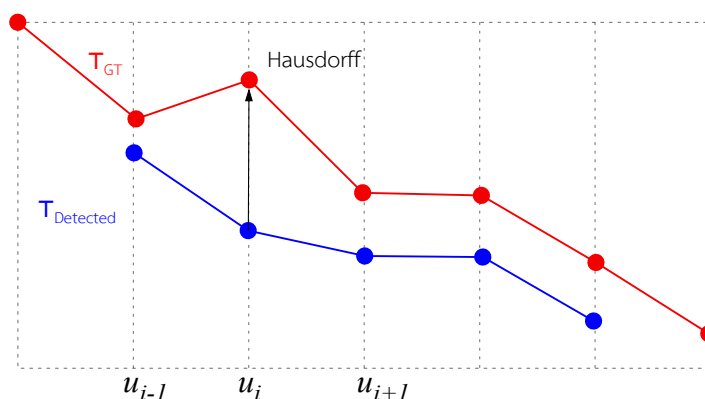
รูปที่ 2.81: ตัวอย่างผลการตรวจหากรอบสี่เหลี่ยมแนวรอยกริดในภาพถ่ายจากกล้องระยะใกล้ (กล้อง 2) ซึ่งทั้งสองกรณีมีค่าอัตราส่วนการซ้อนทับเท่ากัน

2.5.2.2 ระยะคลาดเคลื่อนของแนวรอยกริด

หากพิจารณาแนวรอยกริด คือเส้นโค้งหลายมุม (Polygonal Curve) ซึ่งประกอบด้วยจุดมุม (Vertices) ซึ่งคือตำแหน่งของพิกเซล และเส้นตรงระหว่างจุด (Edges) ดังนั้นระยะคลาดเคลื่อนของแนวรอยกริด คือระยะห่าง

ระหว่างเส้นโค้งหลายมุมของแนวรอยกริดที่ตรวจหาได้ (T_{detected}) กับเส้นโค้งหลายมุมแนวของรอยกริดอ้างอิง (T_{GT}) ซึ่งโดยทั่วไปการวัดผลอาจใช้ผลรวมของระยะห่างกำลังสองระหว่างจุดต่อจุด (พิกเซลต่อพิกเซล) ณ ลำดับหรือตำแหน่งพิกเซลที่ใกล้กันที่มากที่สุด (Sum of Square Distance, Sum of Square Error, ตัวย่อ SSE) ซึ่งให้ค่าเป็นระยะห่างสะสมรวมของทุกจุดระหว่างแนวรอยกริดทั้งสอง แต่การวัดผลด้วย SSE อาจเกิดปัญหาในบางกรณี เช่น เส้นโค้งทั้งสองมีจำนวนจุด (พิกเซล) ไม่เท่ากัน ซึ่งอาจเป็นปัญหาในการเลือกจุดเปรียบเทียบ และค่า SSE ที่ได้มีความแปรปรวนสูง ซึ่งไม่ได้บ่งชี้ความใกล้เคียงกันของเส้นโค้งหลายมุมของแนวรอยกริด จึงเลือกใช้การวัดความคล้ายของเส้นโค้งหลายมุมด้วยระยะทางแบบเฮาส์ดอร์ฟฟ์

การวัดระยะทางแบบเฮาส์ดอร์ฟฟ์ (Hausdorff Distance) [2, 34] ซึ่งจะวัดระยะห่างระหว่างโค้ง T_{detected} กับ T_{GT} โดยจะหาระยะห่างที่มากที่สุดระหว่างจุดที่อยู่ใกล้กันมากที่สุด ระหว่างเส้นโค้งทั้งสอง ดังตัวอย่างแสดงในรูป 2.82 ซึ่งสามารถคำนวณได้จากสมการ 2.37



รูปที่ 2.82: แสดงตัวอย่างเส้นโค้งหลายมุม (Polygonal Curve) ของแนวรอยกริดอ้างอิง (T_{GT}) และเส้นโค้งหลายมุมของแนวรอยกริดที่ตรวจพบ (T_{Detected}) ซึ่งมีจำนวนจุด (พิกเซล) ไม่เท่ากัน และแสดงระยะทางแบบเฮาส์ดอร์ฟฟ์ จาก T_{Detected} ไปยัง T_{GT} , $h(T_{\text{detected}}, T_{\text{GT}})$

กำหนดให้ $h(P, Q) = \max_{p \in P} \min_{q \in Q} \|p - q\|$ คือระยะทางแบบเฮาส์ดอร์ฟฟ์ทางเดียว (Directional Hausdorff Distance) จากเส้นโค้งหลายมุม P ไปยัง Q เมื่อ P, Q คือเซตของจุดในระบบพิกัด C_2 ดังนั้นระยะทางแบบเฮาส์ดอร์ฟฟ์ (δ_H) ระหว่าง T_{detected} กับ T_{GT} เป็นไปตามสมการ 2.37

$$\delta_H = \max\{h(T_{\text{detected}}, T_{\text{GT}}), h(T_{\text{GT}}, T_{\text{detected}})\} \quad (2.37)$$

ดังนั้นการวัดผลการตรวจหาแนวรอยกริดในภาพถ่ายจากกล้องระยะใกล้จะใช้ทั้งสองวิธีการวัดผลข้างต้นร่วมกัน โดยวัดความแม่นยำเฉลี่ยด้วยอัตราส่วนการซ้อนทับของกรอบสี่เหลี่ยมล้อมรอบที่มีค่ามากกว่าหรือเท่ากับ 0.5 เพื่อบ่งชี้ผลการตรวจหาที่เป็นชนิด ตรวจพบ จากนั้นคำนวณระยะห่างระหว่างแนวรอยกริดที่ตรวจพบกับแนวรอยกริดอ้างอิง ด้วยการวัดระยะทางแบบเฮาส์ดอร์ฟฟ์

บทที่ 3

ผล

3.1 การตรวจหารอยกรีดและถ้วยร่องน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้าง

การทดลองตรวจหารอยกรีดและถ้วยร่องน้ำยางในภาพแบ่งออกเป็น 2 การทดลองจาก 2 วิธีตรวจหา ได้แก่ การทดลองวิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพและความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี่ การทดลองวิธีตรวจหารอยกรีดและถ้วยร่องน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้างด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN โดยแต่ละการทดลองมีภาพถ่ายจากกล้องมุมกว้างที่ใช้ทดลองทั้งหมด 500 ภาพ

3.1.1 ผลการทดลอง: วิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี่

วิธีตรวจหารอยกรีดสามารถเลือกใช้งานโมเดลสำหรับแยกความแตกต่างของสีได้ 2 โมเดลสี ได้แก่ โมเดลสี HSV และโมเดลสี $L^*a^*b^*$ และมีตัวแปรควบคุม 3 ตัวแปร สำหรับกำหนดกรอบสี่เหลี่ยมล้อมรอบเริ่มต้นสำหรับค้นหาตำแหน่งรอยกรีด ได้แก่ ระยะเลื่อนของกรอบ (n_s), ความกว้างของกรอบ (w_b), และมุมที่เส้นทแยงมุมหลักทำกับด้านกว้างของกรอบ (θ_b) ตามข้อย่อย 2.3.3.2

ดังนั้นการทดลองจึงประกอบไปด้วยการทดลองแยกสีด้วยภาพโมเดลสี HSV กับการทดลองแยกสีด้วยภาพโมเดลสี $L^*a^*b^*$ โดยแต่ในโมเดลสี จะแบ่งกลุ่มทดลองออกเป็น 6 กลุ่มทดลอง เพื่อกำหนดค่าความละเอียดของตัวแปรกำหนดขนาดของกรอบสี่เหลี่ยม ตามตารางที่ 3.1

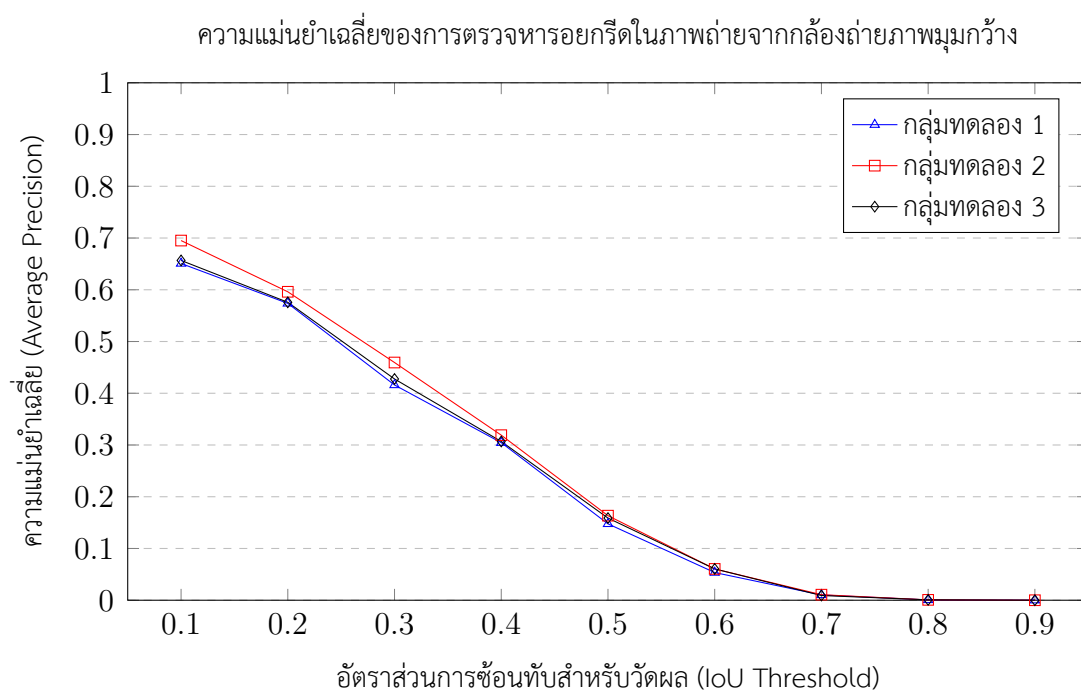
ตารางที่ 3.1: ค่าตัวแปรของแต่ละกลุ่มทดลอง วิธีตรวจหารอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี่

| กลุ่มทดลอง | n_s (พิกเซล) | w_b (ต่อความกว้างลำต้น, $\times w_t$) | θ_b (องศา, °) |
|------------|-------------------|---|----------------------------|
| 1 | 5 | 0.5, 1 | 30, 45 |
| 2 | 5 | 0.5, 0.75, 1 | 30, 37.5, 45 |
| 3 | 5 | 0.5, 0.625, 0.75, 0.875, 1 | 30, 33.75, 37.5, 41.25, 45 |
| 4 | 15 | 0.5, 1 | 30, 45 |
| 5 | 15 | 0.5, 0.75, 1 | 30, 37.5, 45 |
| 6 | 15 | 0.5, 0.625, 0.75, 0.875, 1 | 30, 33.75, 37.5, 41.25, 45 |

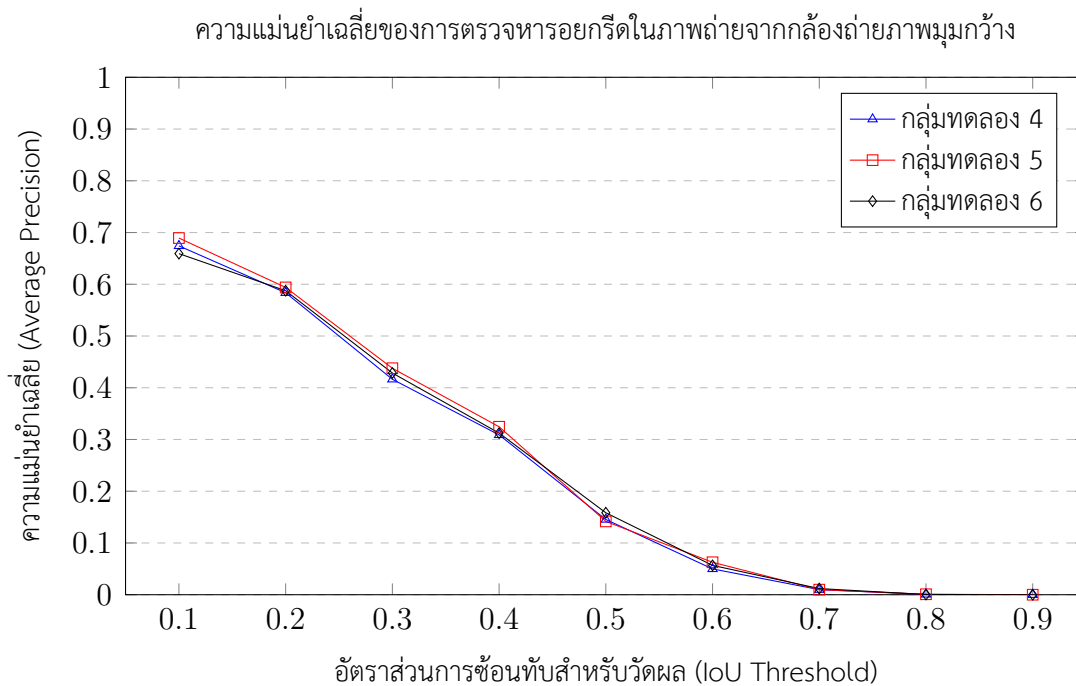
ตารางที่ 3.3: เวลาที่ใช้ประมวลผลด้วยโปรเซสเซอร์หลัก (CPU) เฉลี่ยต่อภาพในหน่วยวินาที ของวิธีตรวจหา รอยกรีดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความ เี่ียงของแนวรอยกรีดผ่านการจัดกลุ่มสี่ ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV

| กลุ่มทดลอง | กลุ่ม 1 | กลุ่ม 2 | กลุ่ม 3 | กลุ่ม 4 | กลุ่ม 5 | กลุ่ม 6 |
|---------------------------|---------|---------|---------|---------|---------|---------|
| เวลาเฉลี่ยต่อภาพ (วินาที) | 1.5572 | 3.0390 | 7.8143 | 0.7406 | 0.9687 | 1.6294 |

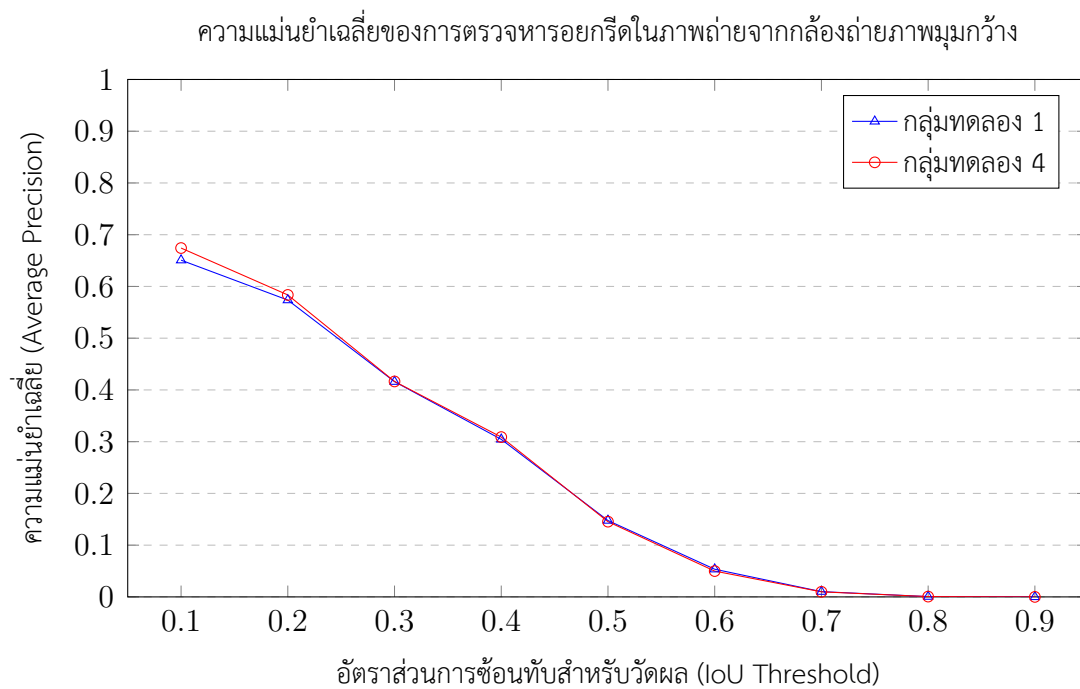
หมายเหตุ ทดลองด้วยโปรแกรม MATLAB R2019a บนคอมพิวเตอร์ Core i7-8086K, RAM 32GB



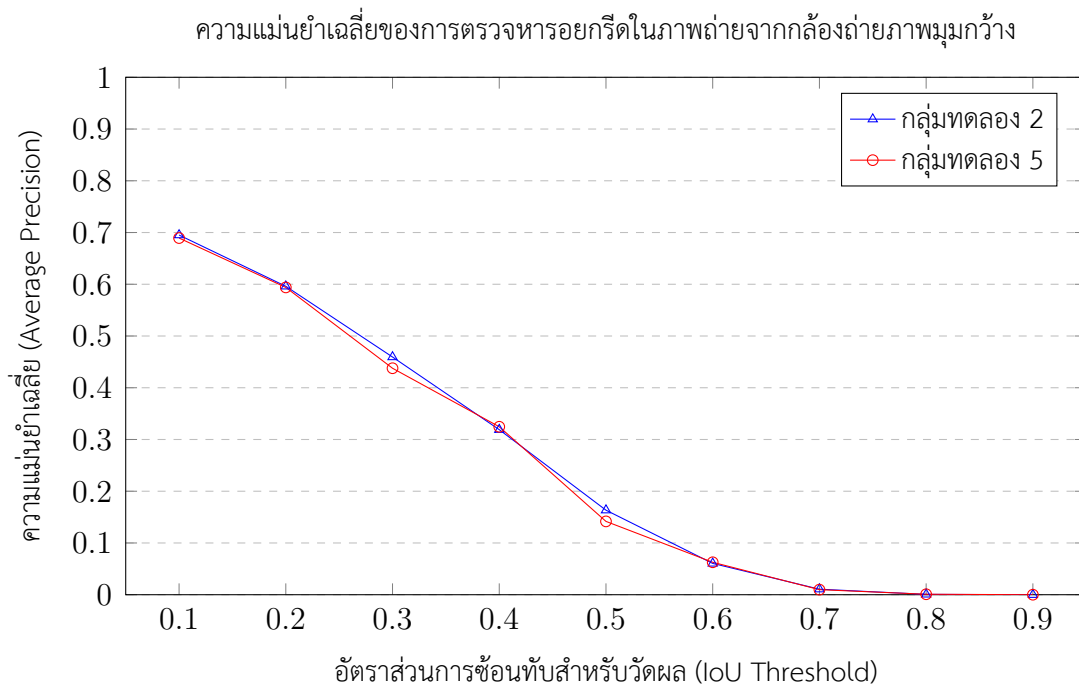
รูปที่ 3.1: ความแม่นยำเฉลี่ยในการตรวจหารอยกรีด ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV ของกลุ่มทดลอง ที่ 1, 2 และ 3 ซึ่งใช้ค่าระยะเลื่อนเท่ากับ 5 พิกเซล ($n_s = 5$)



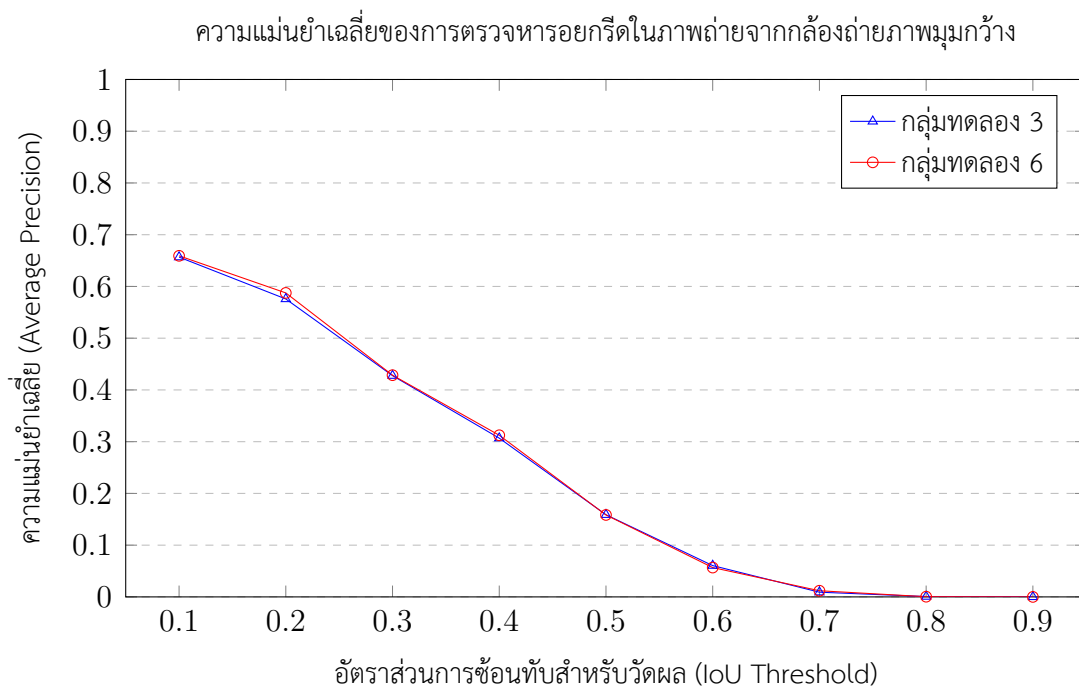
รูปที่ 3.2: ความแม่นยำเฉลี่ยในการตรวจหารอยกริด ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV ของกลุ่มทดลองที่ 4, 5 และ 6 ซึ่งใช้ค่าระยะเลื่อนเท่ากับ 15 พิกเซล ($n_s = 15$)



รูปที่ 3.3: ความแม่นยำเฉลี่ยในการตรวจหารอยกริด ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV ของกลุ่มทดลอง 1 กับกลุ่มทดลอง 4 ซึ่งใช้ค่าระยะเลื่อน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน



รูปที่ 3.4: ความแม่นยำเฉลี่ยในการตรวจหารอยกริด ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV ของกลุ่มทดลอง 2 กับกลุ่มทดลอง 5 ซึ่งใช้ค่าระยะเลื่อน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน



รูปที่ 3.5: ความแม่นยำเฉลี่ยในการตรวจหารอยกริด ในกรณีเมื่อประมวลผลภาพโมเดลสี HSV เปรียบเทียบกลุ่มทดลอง 3 กับกลุ่ม 6 ซึ่งใช้ค่าระยะเลื่อน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน

3.1.1.2 กรณีจัดกลุ่มสี่ด้วยภาพโมเดลสี่ $L^*a^*b^*$

ผลการทดลองในตารางที่ 3.4 แสดงผลการตรวจหารอยกริด โดยทดลองปรับค่าอัตราส่วนการซ้อนทับสำหรับวัตถุผล (IoU Threshold) ตั้งแต่ 0.1 ถึง 0.9 เพื่อวัดผลการตรวจพบ และเพื่อความสามารถในการวางตำแหน่งกรอบสี่เหลี่ยมล้อมรอบรอยกริดในหลายระดับ โดยแบ่งการแสดงผลการทดลองดังนี้

- กราฟผลการตรวจหารอยกริดของกลุ่มที่มีค่า w_b, θ_b ต่างกัน ในขณะที่ค่าระยะเลื่อน (Sliding Distance) เท่ากัน โดย
 - กลุ่มที่ใช้ค่าระยะเลื่อน 5 พิกเซล ($n_s = 5$) ได้แก่ กลุ่มทดลอง 1, 2, และ 3 แสดงในรูปที่ 3.6
 - กลุ่มที่ใช้ค่าระยะเลื่อน 15 พิกเซล ($n_s = 15$) ได้แก่ กลุ่มทดลอง 4, 5, และ 6 แสดงในรูปที่ 3.7
- กราฟผลการตรวจหารอยกริดของกลุ่มทดลองที่มีค่าระยะเลื่อน (n_s) ต่างกัน ในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน ได้แก่
 - กลุ่มทดลอง 1 และ 4 แสดงในรูปที่ 3.8
 - กลุ่มทดลอง 2 และ 5 แสดงในรูปที่ 3.9
 - กลุ่มทดลอง 3 และ 6 แสดงในรูปที่ 3.10

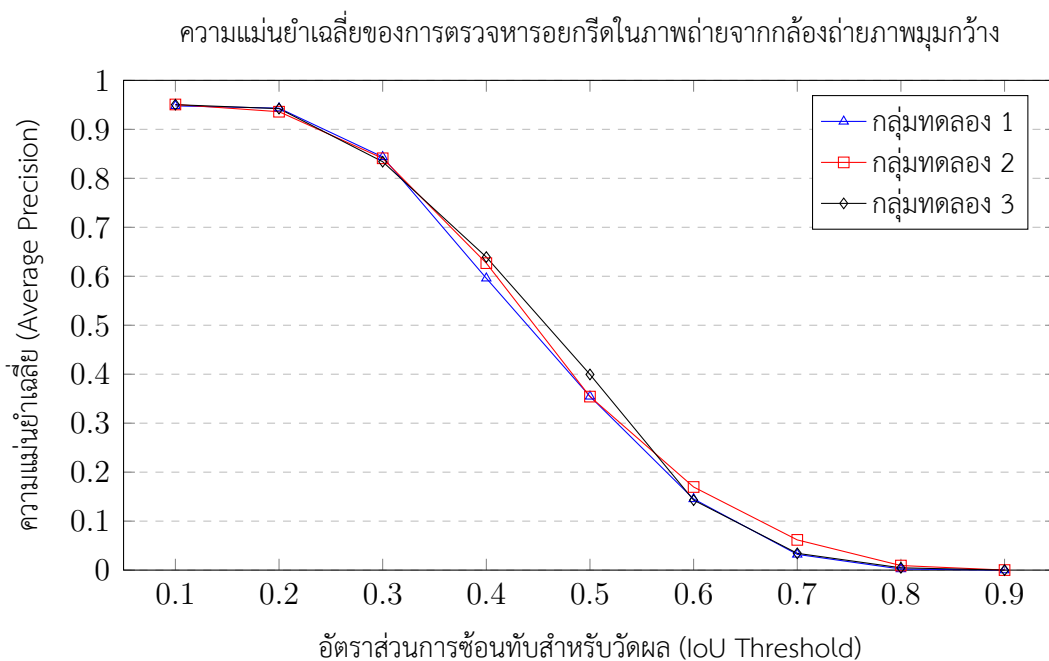
ตารางที่ 3.4: ความแม่นยำเฉลี่ย (Average Precision) ที่อัตราส่วนการซ้อนทับสำหรับวัตถุผลระดับต่างๆ ของวิธีตรวจหารอยกริดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี่ ในกรณีเมื่อประมวลผลภาพโมเดลสี่ $L^*a^*b^*$

| IoU Threshold | กลุ่มทดลอง | | | | | | เฉลี่ยทุกกลุ่ม |
|---------------|------------|---------|---------|---------|---------|---------|----------------|
| | กลุ่ม 1 | กลุ่ม 2 | กลุ่ม 3 | กลุ่ม 4 | กลุ่ม 5 | กลุ่ม 6 | |
| 0.1 | .9479 | .9508 | .9504 | .9391 | .9426 | .9542 | .9475 |
| 0.2 | .9432 | .9360 | .9423 | .9242 | .9372 | .9420 | .9375 |
| 0.3 | .8437 | .8409 | .8339 | .8022 | .8240 | .7956 | .8234 |
| 0.4 | .5959 | .6268 | .6392 | .5733 | .6512 | .6127 | .6165 |
| 0.5 | .3547 | .3543 | .3995 | .3202 | .3293 | .3729 | .3552 |
| 0.6 | .1453 | .1697 | .1429 | .1459 | .1506 | .1271 | .1469 |
| 0.7 | .0321 | .0617 | .0343 | .0371 | .0518 | .0406 | .0429 |
| 0.8 | .0022 | .0093 | .0047 | .0012 | .0075 | .0036 | .0048 |
| 0.9 | .0000 | .0001 | .0000 | .0000 | .0000 | .0000 | .0000 |

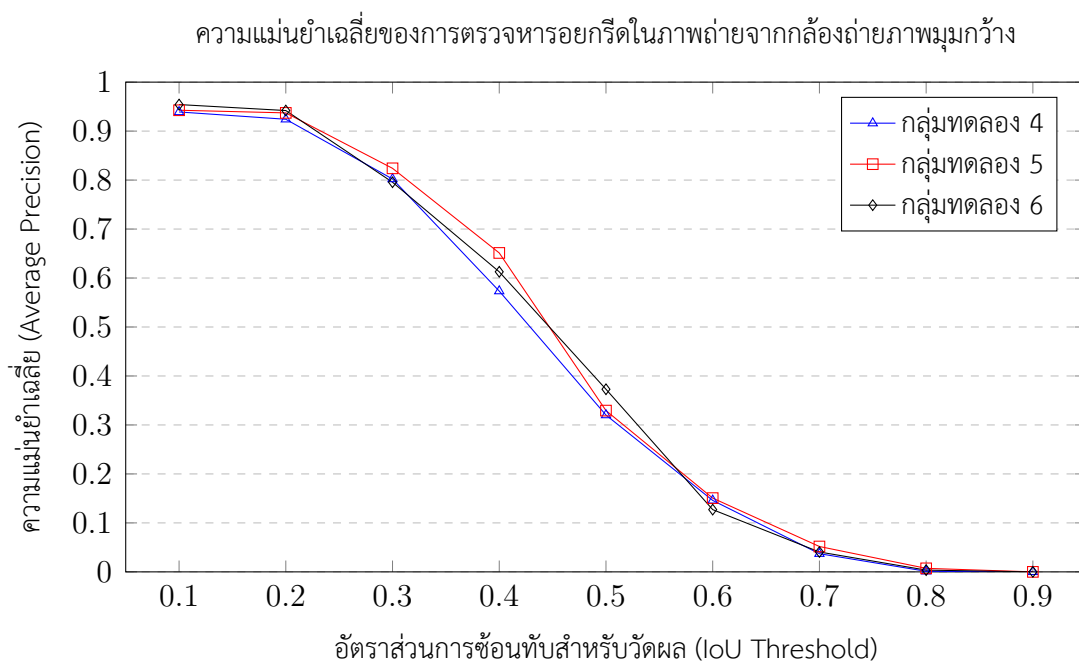
ตารางที่ 3.5: เวลาที่ใช้ประมวลผลด้วยโปรเซสเซอร์หลัก (CPU) เฉลี่ยต่อภาพในหน่วยวินาที ของวิธีตรวจหา รอยกริดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความ เี่ียงของแนวรอยกริดผ่านการจัดกลุ่มสี่ ในกรณีเมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$

| กลุ่มทดลอง | กลุ่ม 1 | กลุ่ม 2 | กลุ่ม 3 | กลุ่ม 4 | กลุ่ม 5 | กลุ่ม 6 |
|---------------|---------|---------|---------|---------|---------|---------|
| เวลา (วินาที) | 1.4275 | 2.9712 | 7.9393 | 0.5769 | 0.7953 | 1.4972 |

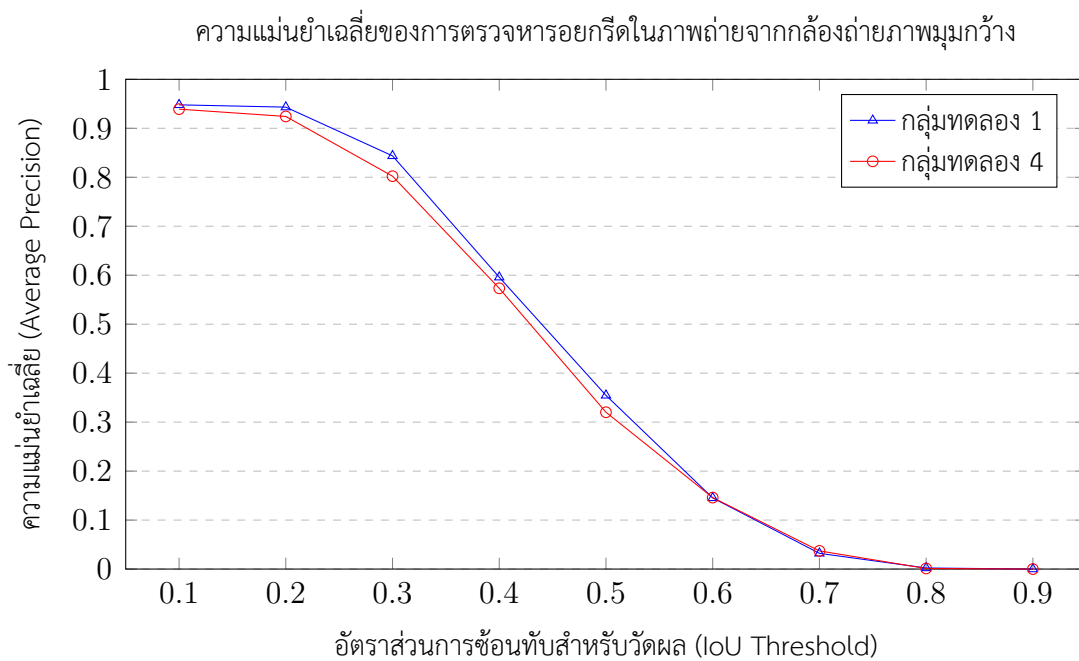
หมายเหตุ ทดลองด้วยโปรแกรม MATLAB R2019a บนคอมพิวเตอร์ Core i7-8086K, RAM 32 GB



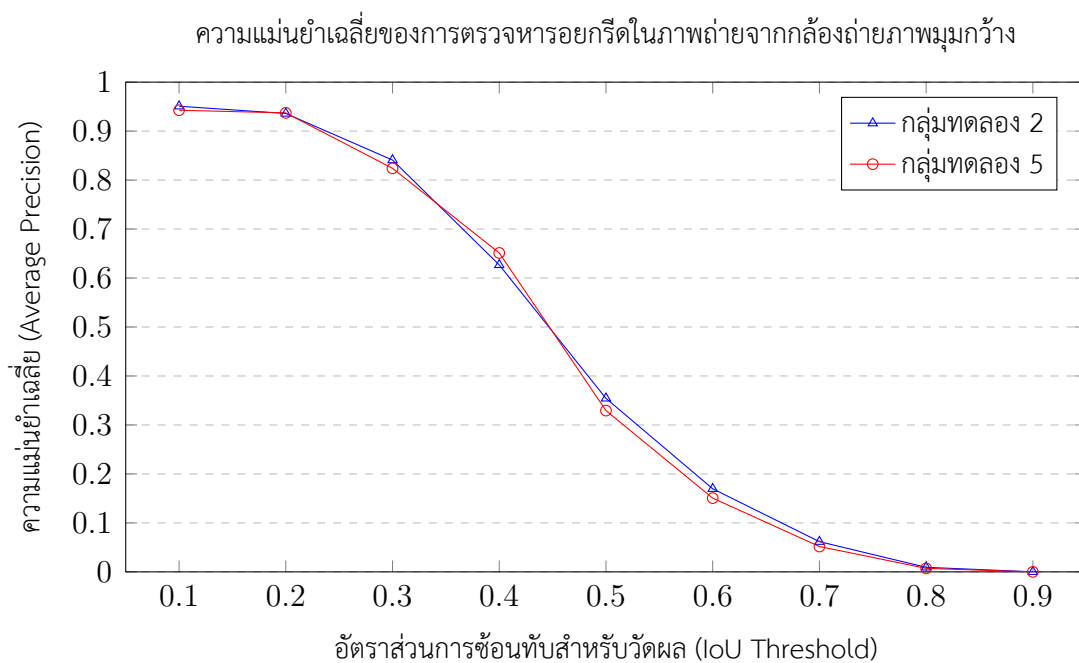
รูปที่ 3.6: ความแม่นยำเฉลี่ยในการตรวจหารอยกริด ในกรณีเมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ ของกลุ่ม ทดลอง 1, 2 และ 3 ซึ่งใช้ค่าระยะเลื่อนเท่ากับ 5 พิกเซล ($n_s = 5$)



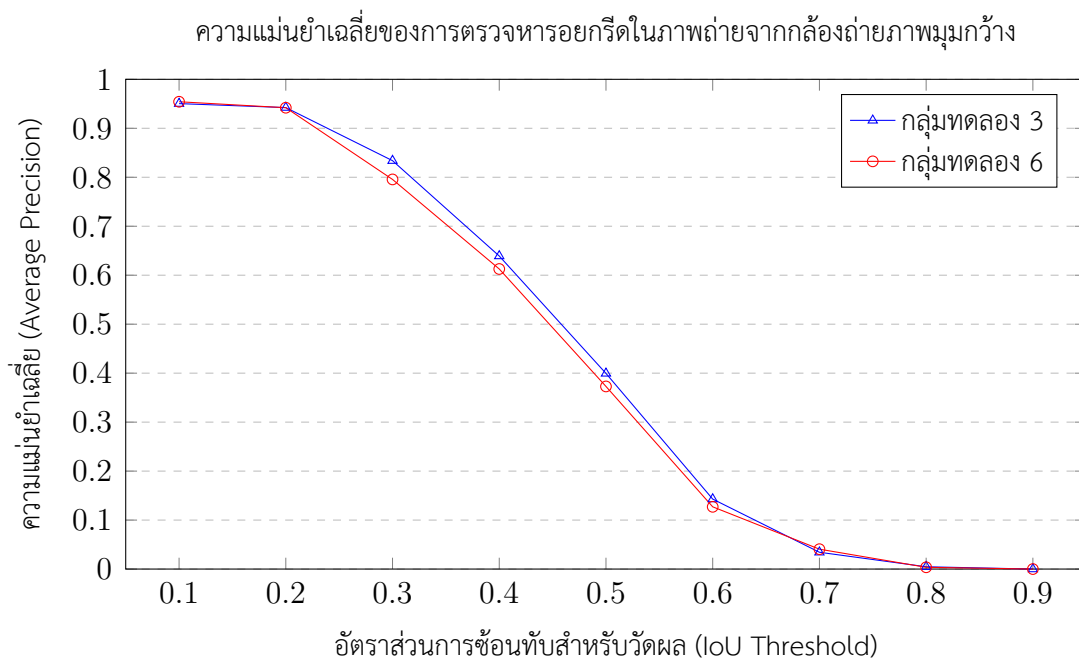
รูปที่ 3.7: ความแม่นยำเฉลี่ยในการตรวจหารอยกริด ในกรณีเมื่อประมวลผลภาพโมเดลสี่ L^*a*b^* ของกลุ่มทดลอง 4, 5 และ 6 ซึ่งใช้ค่าระยะเลื่อนเท่ากับ 15 พิกเซล ($n_s = 15$)



รูปที่ 3.8: ความแม่นยำเฉลี่ยในการตรวจหารอยกริด ในกรณีเมื่อประมวลผลภาพโมเดลสี่ L^*a*b^* ของกลุ่มทดลอง 1 กับ 4 ซึ่งใช้ค่าระยะเลื่อน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน



รูปที่ 3.9: ความแม่นยำเฉลี่ยในการตรวจหารอยกริด ในกรณีเมื่อประมวลผลภาพโมเดลสี่ $L^*a^*b^*$ ของกลุ่มทดลอง 2 กับ 5 ซึ่งใช้ค่าระยะเลื่อน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน

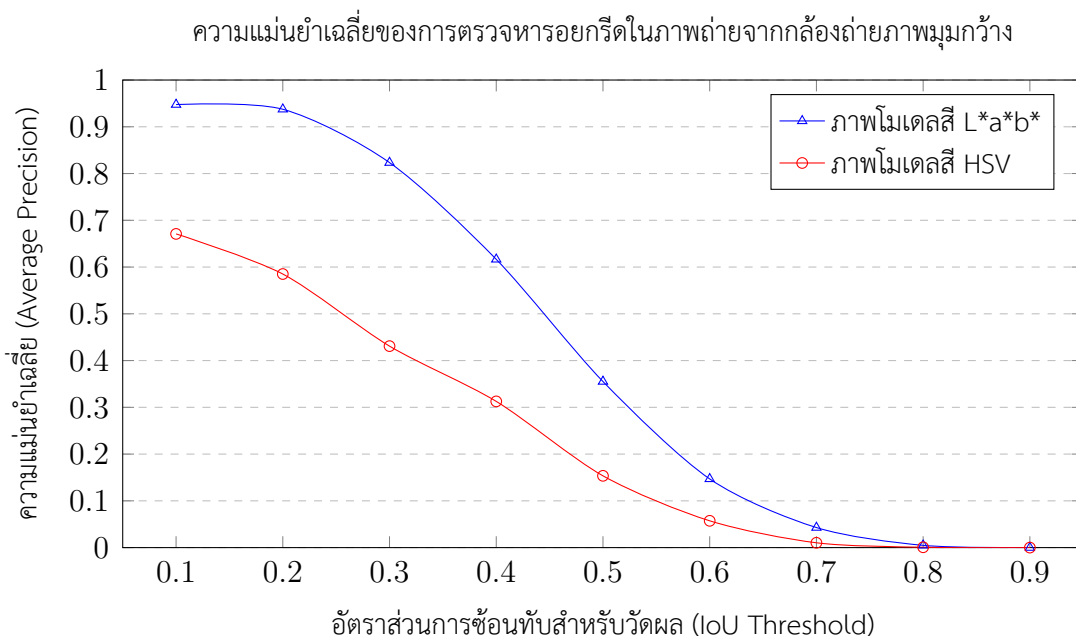


รูปที่ 3.10: ความแม่นยำเฉลี่ยในการตรวจหารอยกริด ในกรณีเมื่อประมวลผลภาพโมเดลสี่ $L^*a^*b^*$ ของกลุ่มทดลองที่ 3 กับ 6 ซึ่งใช้ค่าระยะเลื่อน (n_s) ต่างกันในขณะที่ใช้ค่า w_b, θ_b ชุดเดียวกัน

3.1.1.3 สรุปผลการทดลอง: วิธีตรวจหารอยกริดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี

การแยกกลุ่มสีในภาพโมเดลสี HSV และ $L^*a^*b^*$ การตรวจหารอยกริดด้วยภาพโมเดลสี $L^*a^*b^*$ ได้ความแม่นยำเฉลี่ยสูงกว่าการตรวจหารอยกริดด้วยภาพโมเดลสี HSV ในทุกอัตราส่วนการซ้อนทับสำหรับวัตถุ สามารถสังเกตได้อย่างชัดเจนเมื่อวัดผลการตรวจหารอยกริดด้วยภาพโมเดลสี $L^*a^*b^*$ ด้วยอัตราส่วนการซ้อนทับที่มีค่าในช่วง 0.1 ถึง 0.3 ตามรูปที่ 3.11 พบว่าความแม่นยำจะอยู่ในช่วง 0.8 ถึง 0.9 หรือกล่าวคืออัลกอริทึมสามารถตรวจหาแนวรอยกริดได้ถูกต้องถึง 90% ในทุกกลุ่มทดลอง ในขณะที่การแยกกลุ่มสีด้วยภาพโมเดลสี HSV สามารถตรวจหารอยกริดได้สูงสุดเพียง 70% ทำให้สามารถสรุปได้ว่าการจัดกลุ่มสีโทนแดงในช่องสัญญาณ a^* ของโมเดลสี $L^*a^*b^*$ ด้วยวิธีการจัดกลุ่มแบบค่ากลาง K กลุ่ม สามารถใช้เป็นฟีเจอร์เพื่อรู้จำบริเวณและระบุตำแหน่งรอยกริดได้ดีกว่าการจุดกลุ่มสีในช่องสัญญาณ Hue ของภาพโมเดลสี HSV

ดังนั้นวิธีการจัดกลุ่มสีและการตรวจหากรอบสี่เหลี่ยมด้วยค่าคะแนน *boxScore* สามารถระบุตำแหน่งของรอยกริดได้ไม่ผิดนัก เนื่องจากความแม่นยำเมื่อวัดผลด้วยอัตราส่วนการซ้อนทับสำหรับวัตถุที่ 0.5 ซึ่งเป็นค่ามาตรฐานสำหรับบ่งชี้ผลการตรวจหาวัตถุ อัลกอริทึมสามารถระบุตำแหน่งของกรอบล้อมรอบรอยกริดได้สูงสุดเพียง 36% เมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ และค่าความแม่นยำจะลดลงอย่างรวดเร็วในระดับอัตราส่วนการซ้อนทับสำหรับวัตถุที่สูงขึ้น



รูปที่ 3.11: ความแม่นยำเฉลี่ยในการตรวจหารอยกริดด้วยวิธีตรวจหารอยกริดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี เปรียบเทียบระหว่างในกรณีเมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ และโมเดลสี HSV

ความละเอียดของกรอบสี่เหลี่ยมเพื่อค้นหาตำแหน่งรอยกรีด ความละเอียดของการกำหนดกรอบสี่เหลี่ยมล้อมรอบสำหรับค้นหาตำแหน่งรอยกรีดตามข้อย่อย 2.3.3.2 ก็ส่งผลต่อความแม่นยำเฉลี่ย โดยในการทดลองเพิ่มความละเอียดและจำนวนของค่า w_b, θ_b ทำให้ค่าความแม่นยำเฉลี่ยเพิ่มสูงขึ้นโดยเฉพาะช่วงอัตราส่วนการซ้อนทับสำหรับวัดผลที่ 0.5 ถึง 0.7 ในกลุ่มทดลองที่ 1, 2, และ 3 (รูปที่ 3.6) และในกลุ่มทดลองที่ 4, 5, และ 6 (รูปที่ 3.7) ในขณะที่การเพิ่มจำนวนกรอบสี่เหลี่ยมล้อมรอบเริ่มต้นโดยลดระยะห่างระหว่างกัน (ลดระยะเลื่อน, n_s) ทำให้ค่าความแม่นยำของการตรวจหาเพิ่มขึ้นเพียงเล็กน้อย ในช่วงอัตราส่วนการซ้อนทับสำหรับวัดผลที่ค่าน้อยกว่า 0.5 สังเกตได้ในรูปที่ 3.8, 3.9 และ 3.10 เมื่อกำหนดให้กลุ่มทดลองใช้ค่า w_b, θ_b ชุดเดียวกัน แต่มีระยะห่าง (n_s) เป็น 5 และ 15 พิกเซล ทำให้สรุปได้ว่าการเพิ่มความละเอียดของกรอบเพื่อค้นหาตำแหน่งรอยกรีด สามารถเพิ่มความแม่นยำของการตรวจหารอยกรีดได้

3.1.2 ผลการทดลอง: วิธีตรวจหารอยกริดและถ้วยรองน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้างด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN

3.1.2.1 ตัวแปรต้นของการทดลอง

กลุ่มทดลอง จากการอธิบายในหัวข้อที่ 2.3.4.5 ได้จัดกลุ่มการทดลองเป็น 3 กลุ่ม ตามการจัดสรรช่องสัญญาณของภาพอินพุต ซึ่งข้อมูลในแต่ละกลุ่มทดลองจะใช้สอน (Training) และใช้ทดสอบ (Testing) เน็ตเวิร์กFaster-RCNN ด้วยวิธีการแบ่งข้อมูลทดสอบ K กลุ่ม (K-fold Cross Validation) ที่มีค่า K เท่ากับ 5 โดยมีกลุ่มทดลองดังนี้

- 1 RGB, ภาพถ่ายสี RGB ขนาดรวม 24 บิต (ขนาดบิตพิกเซล 8 บิตต่อช่องสัญญาณภาพ)
- 2 DDD, ภาพถ่ายความลึก ซึ่งเป็นภาพโทนเทาขนาดรวม 24 บิต ที่เกิดจากการขยายภาพโทนเทา 1 ช่องสัญญาณขนาด 8 บิต เป็น 3 ช่องสัญญาณ
- 3 RGBxD, ภาพถ่ายสี RGB คู่ณด้วยภาพถ่ายความลึกโทนเทาแบบพิกเซลต่อพิกเซล

การสร้าง Faster-RCNN ให้ทำงานร่วมกับ MobileNetV2 เพื่อทดลองข้อมูลในแต่ละกลุ่มทดลองข้างต้น มีประเด็นหรือรายละเอียดที่สามารถควบคุมหรือกำหนดได้ซึ่งถือเป็นตัวแปรต้นสำหรับการทดลอง มีรายละเอียดดังต่อไปนี้

- 1 ลำดับตำแหน่งเลเยอร์ใน MobileNetV2 ในการแทรกกลุ่มของเลเยอร์ส่วนปลายของ RPN
ลำดับเลเยอร์ใน MobileNetV2 จะประมวลผลเทนเซอร์ที่มีขนาดแตกต่างกันออกไป การเลือกเลเยอร์สำหรับป้อนเอาต์พุตจากเลเยอร์ส่วนปลาย RPN เพื่อทำนายพื้นที่ (Activation Map) ขนาดของเทนเซอร์ที่เลเยอร์ดังกล่าวจะส่งผลกระทบต่อความสามารถของเน็ตเวิร์กในการตรวจหาวัตถุขนาดเล็ก ดังนั้นในการทดลองจึงกำหนดให้ทดลองสร้าง RPN ใน 3 ตำแหน่งเลเยอร์ใน MobileNetV2 ที่มีขนาดเทนเซอร์ต่างกัน ได้แก่
 - ตำแหน่ง RPN บล็อก 3, เลเยอร์ “block_3_depthwise_relu” ซึ่งประมวลผลเทนเซอร์ที่มีขนาดกว้าง×สูง เท่ากับ 28×28 พิกเซล
 - ตำแหน่ง RPN บล็อก 10, เลเยอร์ “block_10_depthwise_relu” ซึ่งประมวลผลเทนเซอร์ที่มีขนาดกว้าง×สูง เท่ากับ 14×14 พิกเซล
 - ตำแหน่ง RPN บล็อก 13, เลเยอร์ “block_13_depthwise_relu” ซึ่งประมวลผลเทนเซอร์ที่มีขนาดกว้าง×สูง เท่ากับ 7×7 พิกเซล
- 2 ขนาดของภาพ การใช้ภาพขนาดใหญ่ทำการสอนแบบปรับละเอียดเน็ตเวิร์ก เน็ตเวิร์กที่เรียนรู้จะมีโอกาสตรวจหาวัตถุขนาดเล็ก ได้แม่นยำทั้งในด้านผลการทำนายและการวางตำแหน่งของกรอบสี่เหลี่ยมได้ดีกว่าการใช้ภาพขนาดเล็ก เนื่องจากผลของการคอนโวลูชันแบบต่อเนื่องภายในเน็ตเวิร์กทำให้ข้อมูลมีโอกาสดูสูญหาย (Information Loss) แต่ทั้งนี้หากกำหนดให้เน็ตเวิร์กประมวลผลภาพขนาดใหญ่ จะใช้หน่วยความจำและทรัพยากรอื่นของเครื่องมากกว่าการใช้ภาพขนาดเล็ก โดยในการทดลองกำหนดให้ใช้

ภาพขนาด $224 \times 224 \times 3$ คงที่ตลอดทั้งการทดลอง ประโยชน์เพื่อทำให้การทดลองสามารถทำได้บนคอมพิวเตอร์เดสก์ท็อป และเพื่อเปรียบ Faster-RCNN กับสถาปัตยกรรมของเน็ตเวิร์กแบบอื่นที่ใช้ภาพขนาดเท่ากัน

- 3 จำนวนพื้นที่ย่อยที่ผลิตจาก RPN ขณะตรวจหาวัตถุ (RPN Proposal for Object Inferencing) Faster-RCNN อนุญาตให้ผู้พัฒนาสามารถเลือกจำนวนได้อย่างอิสระ [66] ทั้งนี้การทดลองจะกำหนดให้จำนวนพื้นที่ย่อย (RPN Proposals) คงที่สำหรับการวัดความแม่นยำของเน็ตเวิร์กที่จำนวน 100 พื้นที่ต่อภาพ (รายละเอียดกล่าวถึงในหัวข้อที่ 3.1.2.6)

3.1.2.2 ผลการสอนและค่าลอสฟังก์ชันของการสอน (Training Loss)

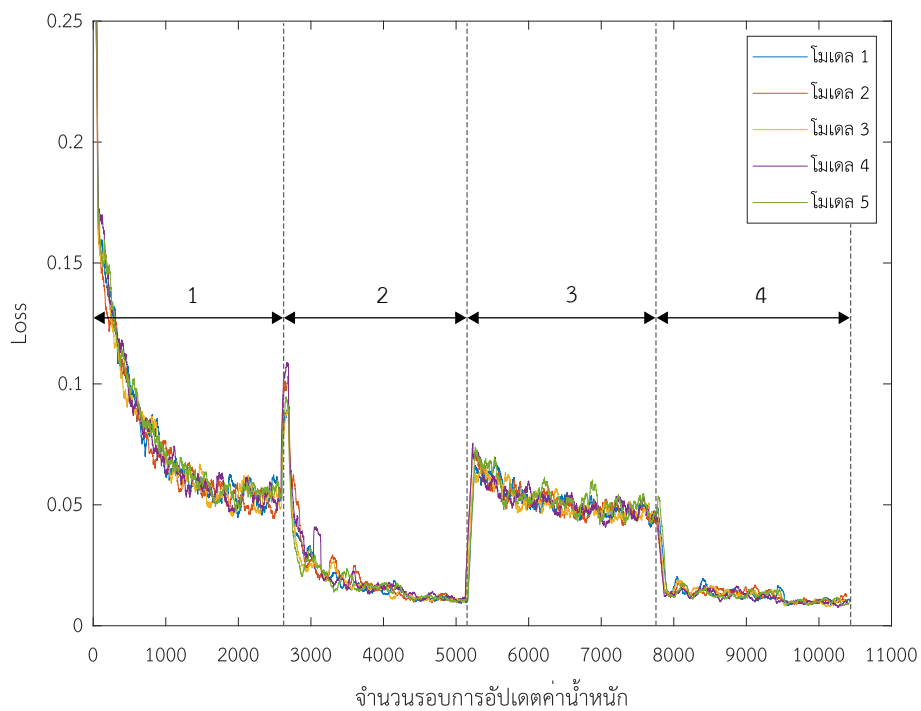
ผลการสอน Faster-RCNN ด้วยวิธีสอนแบบสลับ 4 ชั้นตอน (4-Alternating Training) ทำให้ Faster-RCNN ที่มี RPN ในตำแหน่งเลเยอร์บล็อก 13 ใน MobileNetV2 ไม่สามารถสอนได้ เนื่องจากค่าฟังก์ชัน Loss สูงมากเกินไปในทุกกลุ่มทดลอง (Loss = NaN) ซึ่งอาจเกิดจากเน็ตเวิร์ก RPN ไม่สามารถทำนายทั้งตำแหน่งของพื้นที่ที่น่าสนใจ (ROI) ได้ถูกต้อง ส่งผลต่อการสอนเน็ตเวิร์กรู้จำวัตถุ (Detection Network) ไม่มีพื้นที่อินพุตที่เป็นวัตถุเพื่อสอนการทำนายต่อได้ หรือเน็ตเวิร์ก RPN สามารถทำนายพื้นที่ที่น่าสนใจได้ แต่เน็ตเวิร์กตรวจหาวัตถุไม่มีสามารถสกัดหาพีเจอร์ที่มีประโยชน์ต่อการทำนายพื้นที่ได้

การลอสฟังก์ชันการสอนในแต่ละรอบการอัปเดตค่าน้ำหนักใน Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อกที่ 3 (RPN บล็อก 3) และเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อกที่ 10 (RPN บล็อก 3) ของแต่ละกลุ่มทดลอง โดยหมายเลข 1 ถึง 4 แสดงค่าลอสฟังก์ชันของเน็ตเวิร์กย่อย ตามขั้นตอนที่ 1 ถึง 4 ของวิธีสอนแบบสลับ 4 ชั้นตอน

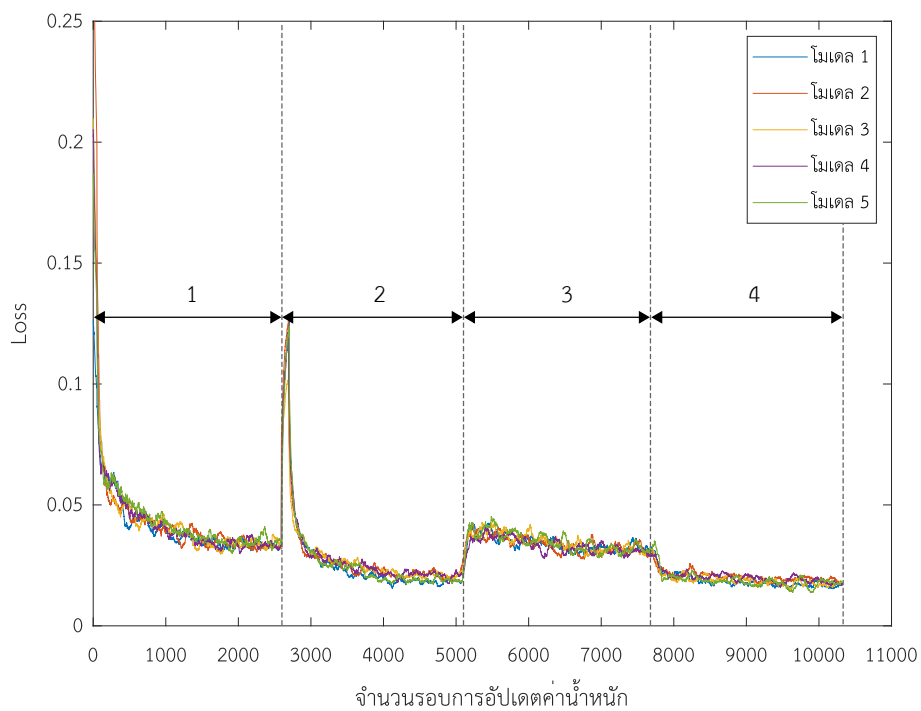
ในแต่ละกลุ่มข้อมูลทดลอง จะทำการสอน Faster-RCNN จำนวน 5 เน็ตเวิร์ก (โมเดล) ตามวิธีการแบ่งข้อมูลทดสอบ K กลุ่ม (K=5) โดยกำหนดพารามิเตอร์การสอน (Hyper-parameters) ที่ทำให้ลอสฟังก์ชันของแต่ละโมเดลมีค่าใกล้เคียงกัน ประโยชน์เพื่อทำให้ความสามารถของแต่ละโมเดลที่ได้ใกล้เคียงกัน ซึ่งจะลดความแปรปรวนในการวัดผลการตรวจหาวัตถุของเน็ตเวิร์ก

ผลการสอนเน็ตเวิร์กในแต่ละกลุ่มทดลองจำนวน 5 เน็ตเวิร์ก (โมเดล) มีค่าลอสฟังก์ชันแสดงในรูปที่ 3.12 ถึงรูปที่ 3.16 ซึ่งลอสฟังก์ชันในแต่ละโมเดลที่มีค่าใกล้เคียงกัน หมายถึงโมเดลมีความสามารถในการทำนายวัตถุใกล้เคียงกัน ทั้งนี้เพื่อเป็นการตรวจสอบอย่างคร่าว และป้องกันการสอนที่ทำให้บางเน็ตเวิร์กมีพารามิเตอร์ไม่ลู่เข้า ทำให้ค่าความแม่นยำเฉลี่ยของโมเดลภายในแต่ละกลุ่มทดลองมีความน่าเชื่อถือมากขึ้น

กลุ่มทดลองที่ 1 : ภาพสี RGB

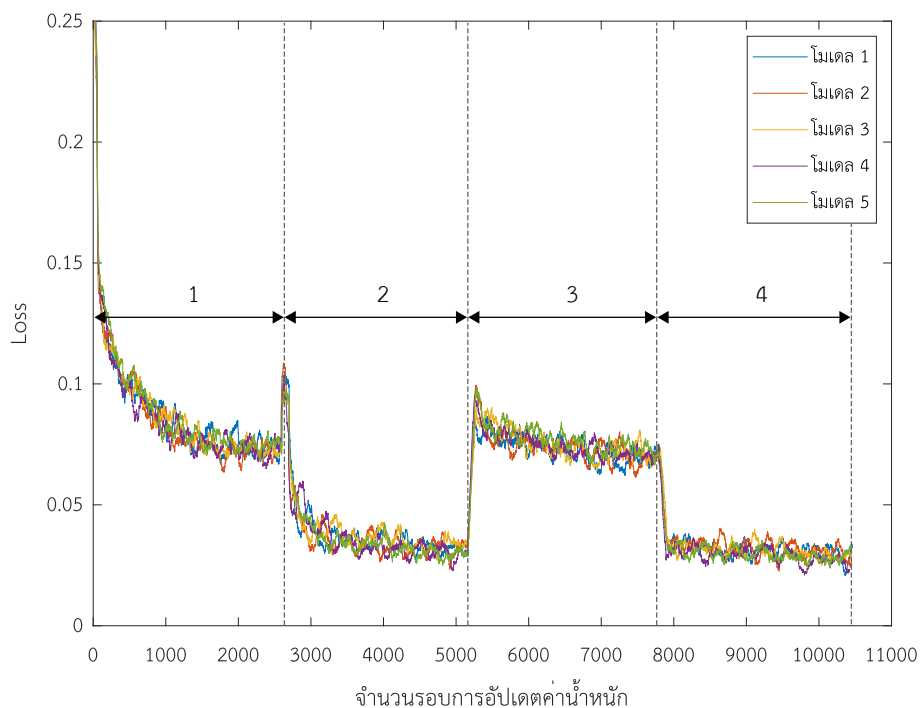


รูปที่ 3.12: ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 3

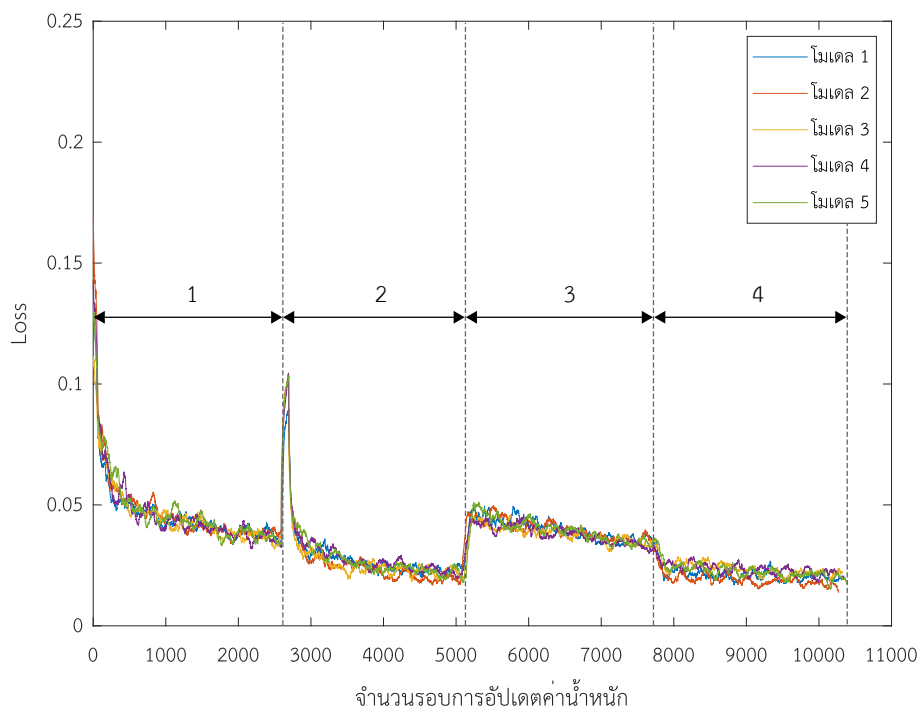


รูปที่ 3.13: ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 10

กลุ่มทดลองที่ 2 : ภาพความลึกโทนเทา

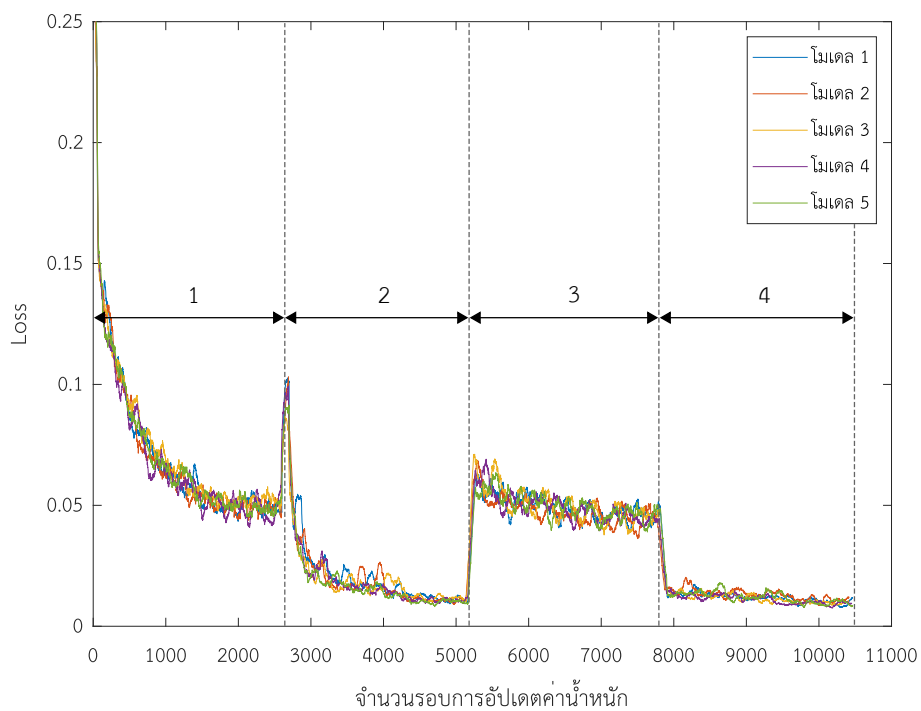


รูปที่ 3.14: ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 3

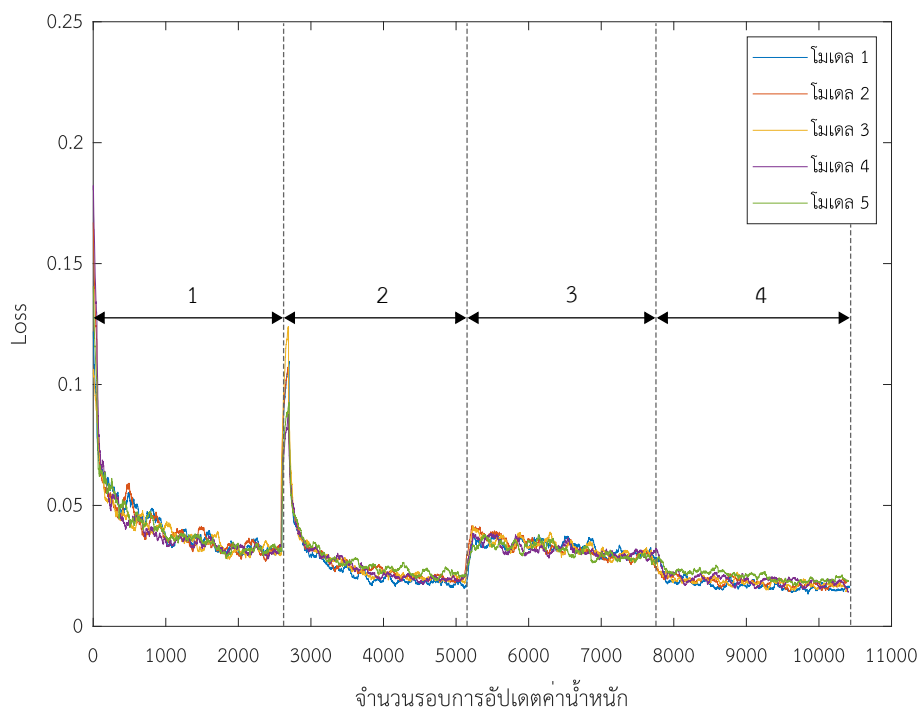


รูปที่ 3.15: ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 10

กลุ่มทดลองที่ 3 : ภาพสี RGB เมื่อคูณด้วยภาพความลึกโทนเทา



รูปที่ 3.16: ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 3



รูปที่ 3.17: ลอสฟังก์ชันการสอนเน็ตเวิร์กที่มี RPN บล็อก 10

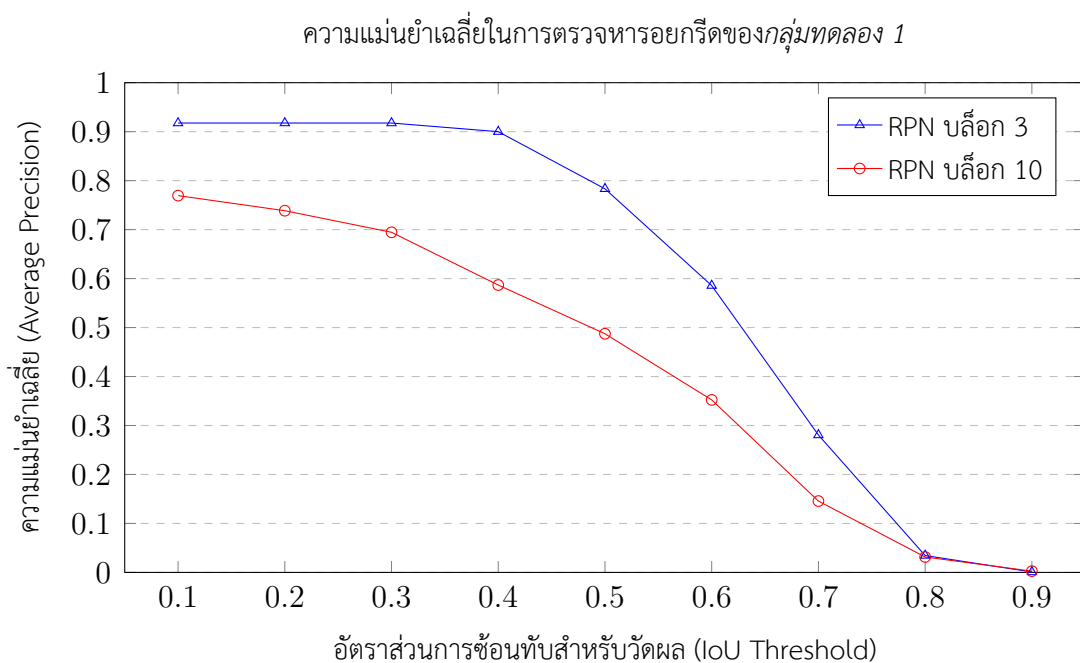
3.1.2.3 ความแม่นยำของการตรวจหารอยกริด

การตรวจหารอยกริดในภาพด้วยกรอบสี่เหลี่ยมจะวัดผลด้วยค่าความแม่นยำเฉลี่ย (Average Precision) ด้วยค่าอัตราส่วนการซ้อนทับสำหรับวัดผล (IoU Threshold) ตั้งแต่ 0.1 ถึง 0.9 เพื่อสังเกตความแม่นยำของการวางตำแหน่งกรอบสี่เหลี่ยมของ Faster-RCNN และเพื่อเปรียบเทียบกับวิธีตรวจหารอยกริดก่อนหน้านี้ (วิธีตรวจหารอยกริดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี่)

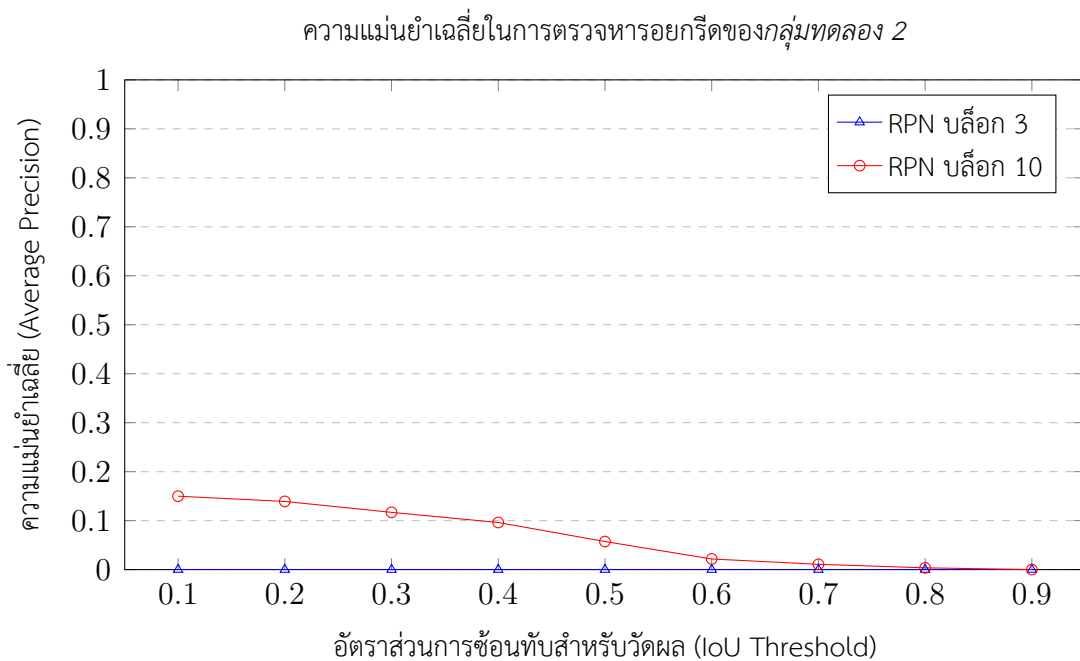
การตรวจหากรอบสี่เหลี่ยมของรอยกริด จะปรับเปลี่ยนจากเดิมที่เน็ตเวิร์กจะเข้าที่พุดค่าของกรอบที่มีค่าคะแนนสูงเกินกว่า 0.5 (Confident Score Threshold) เป็นการหาค่ากรอบที่มีค่าคะแนนสูงสุดเพียงกรอบเดียว (Max Confident Score) เนื่องจากในแต่ละภาพมีกรอบรอยกริดเพียงตำแหน่งเดียว

ตารางที่ 3.6: ความแม่นยำเฉลี่ย (Average Precision) ของการตรวจหารอยกริดที่อัตราส่วนการซ้อนทับสำหรับวัดผลระดับต่างๆ ด้วย Faster-RCNN ที่มี RPN ในตำแหน่งต่างๆ ของแต่ละกลุ่มทดลอง

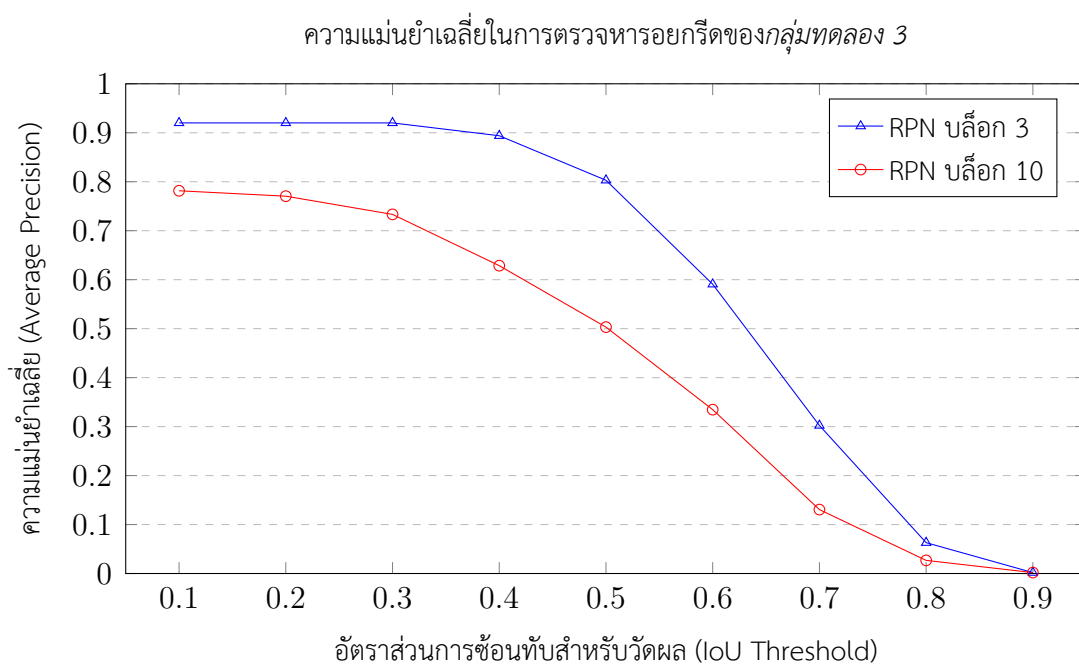
| IoU Threshold | กลุ่มทดลอง 1 (RGB) | | กลุ่มทดลอง 2 (DDD) | | กลุ่มทดลอง 3 (RGBxD) | |
|---------------|--------------------|-----------------|--------------------|-----------------|----------------------|-----------------|
| | RPN บล็อก 3 | RPN บล็อก 10 | RPN บล็อก 3 | RPN บล็อก 10 | RPN บล็อก 3 | RPN บล็อก 10 |
| 0.1 | .9176 | .7694 | 0 | .1499 | .9201 | .7815 |
| 0.2 | .9176 | .7386 | 0 | .1392 | .9201 | .7705 |
| 0.3 | .9176 | .6945 | 0 | .1169 | .9201 | .7333 |
| 0.4 | .9000 | .5868 | 0 | .0962 | .8938 | .6286 |
| 0.5 | .7833 | .4873 | 0 | .0573 | .8030 | .5031 |
| 0.6 | .5857 | .3523 | 0 | .0217 | .5907 | .3346 |
| 0.7 | .2806 | .1457 | 0 | .0108 | .3023 | .1305 |
| 0.8 | .0345 | .0315 | 0 | .0036 | .0630 | .0268 |
| 0.9 | .0009 | .0026 | 0 | .0001 | .0019 | .0019 |



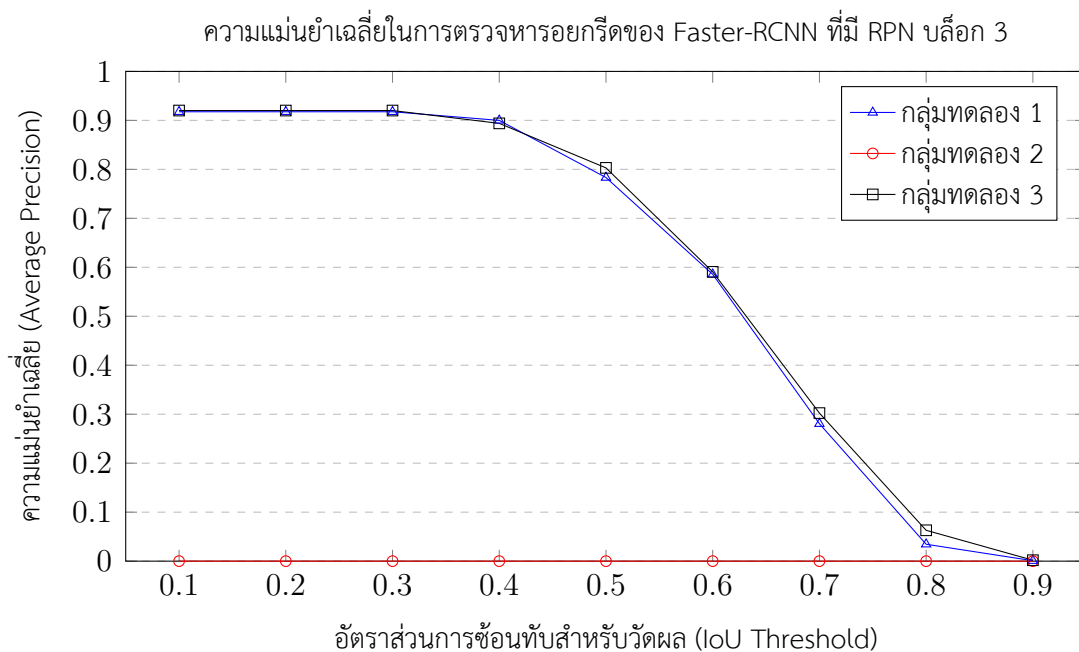
รูปที่ 3.18: ความแม่นยำเฉลี่ยในการตรวจหากรอยกึ่งด้วย Faster-RCNN กลุ่มทดลองที่ 1 ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2



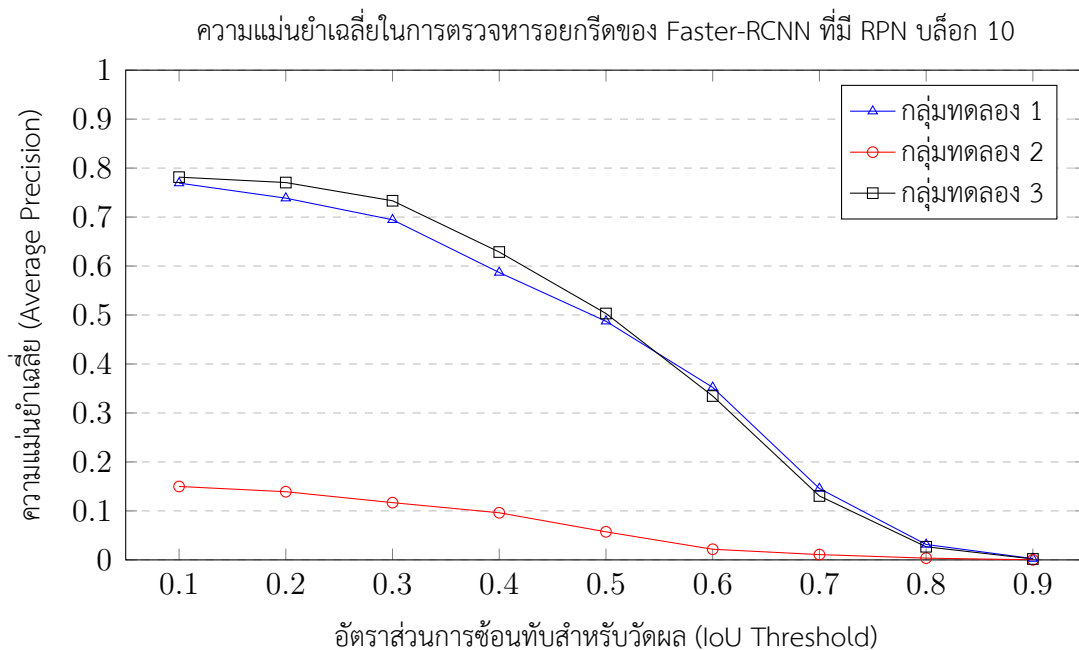
รูปที่ 3.19: ความแม่นยำเฉลี่ยในการตรวจหากรอยกึ่งด้วย Faster-RCNN กลุ่มทดลองที่ 2 ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2



รูปที่ 3.20: ความแม่นยำเฉลี่ยในการตรวจหากรอยกจริตด้วย Faster-RCNN กลุ่มทดลองที่ 3 ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2



รูปที่ 3.21: ความแม่นยำเฉลี่ยในการตรวจหากรอยกจริตด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 ของ MobileNetV2 แสดงกลุ่มทดลองที่ 1, 2, และ 3



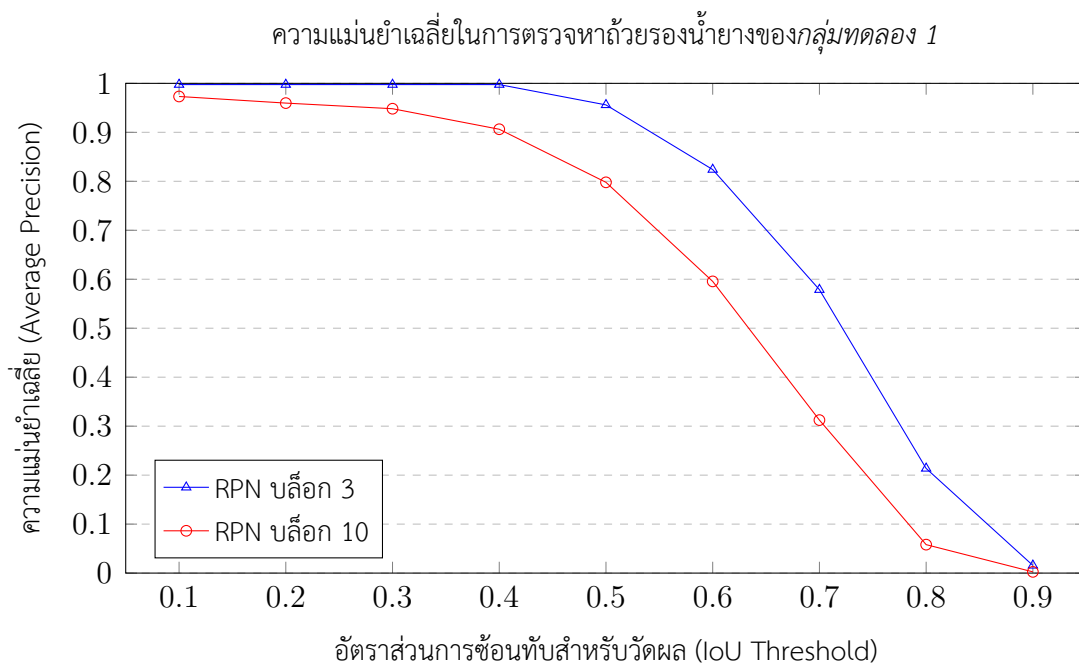
รูปที่ 3.22: ความแม่นยำเฉลี่ยในการตรวจหากรอยกจริตด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 10 ของ MobileNetV2 แสดงกลุ่มทดลองที่ 1, 2, และ 3

3.1.2.4 ความแม่นยำของการตรวจหาถ้วยร่อนน้ำยาง

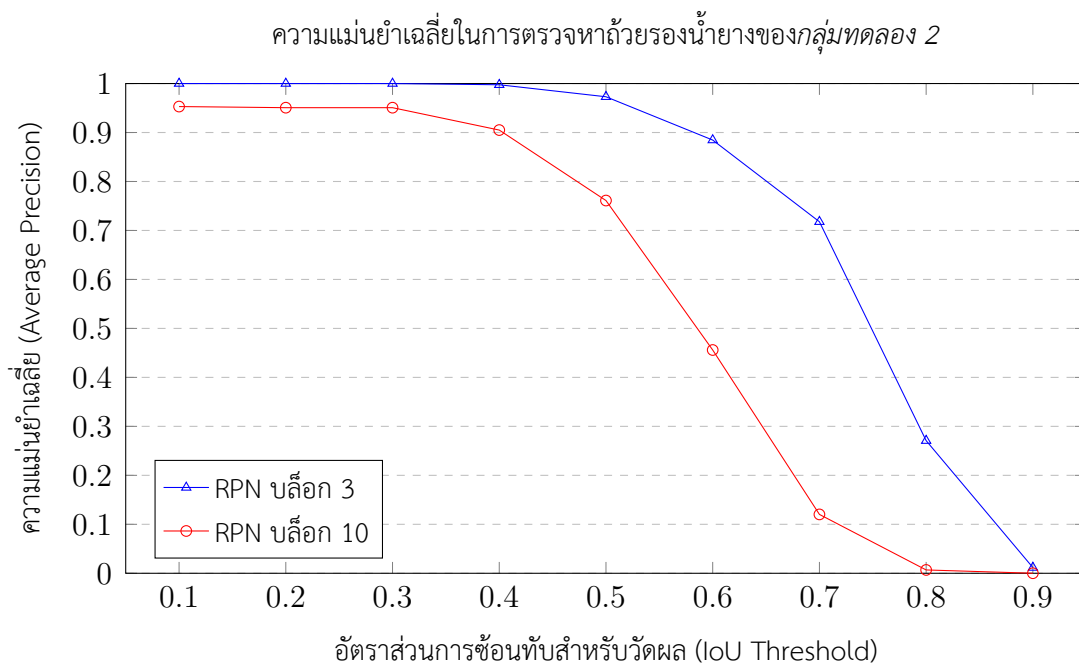
การตรวจหาถ้วยร่อนน้ำยางได้วัดผลด้วยค่าความแม่นยำเฉลี่ย (Average Precision) ที่อัตราส่วนการซ้อนทับของกรอบสี่เหลี่ยมล้อมรอบในหลายระดับ ตั้งแต่ 0.1 ถึง 0.9 เช่นเดียวกับการวัดผลการตรวจหารอยกรีด ผลการตรวจหาถ้วยร่อนน้ำยางในแต่ละกลุ่มทดลองแสดงในตารางที่ 3.7

ตารางที่ 3.7: ความแม่นยำเฉลี่ย (Average Precision) ของการตรวจหาถ้วยร่อนน้ำยางที่อัตราส่วนการซ้อนทับสำหรับวัดผลระดับต่างๆ ด้วย Faster-RCNN ที่มี RPN ในตำแหน่งต่างๆ ของแต่ละกลุ่มทดลอง

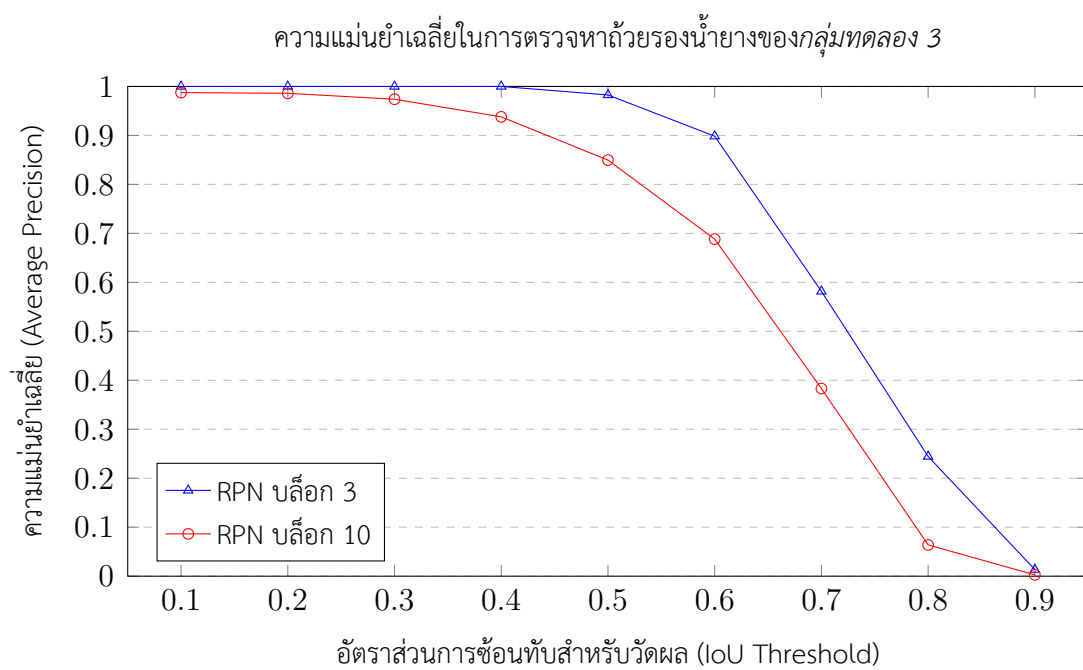
| IoU Threshold | กลุ่มทดลอง 1 (RGB) | | กลุ่มทดลอง 2 (DDD) | | กลุ่มทดลอง 3 (RGBxD) | |
|---------------|--------------------|-----------------|--------------------|-----------------|----------------------|-----------------|
| | RPN บล็อก 3 | RPN บล็อก 10 | RPN บล็อก 3 | RPN บล็อก 10 | RPN บล็อก 3 | RPN บล็อก 10 |
| 0.1 | .9975 | .9732 | 1.000 | .9530 | 1.000 | .9875 |
| 0.2 | .9975 | .9597 | 1.000 | .9507 | 1.000 | .9860 |
| 0.3 | .9975 | .9482 | 1.000 | .9507 | 1.000 | .9739 |
| 0.4 | .9975 | .9063 | .9975 | .9050 | 1.000 | .9379 |
| 0.5 | .9560 | .7979 | .9730 | .7611 | .9826 | .8496 |
| 0.6 | .8240 | .5956 | .8846 | .4559 | .8984 | .6883 |
| 0.7 | .5787 | .3124 | .7178 | .1202 | .5817 | .3831 |
| 0.8 | .2137 | .0581 | .2708 | .0068 | .2444 | .0639 |
| 0.9 | .0157 | .0025 | .0117 | .0001 | .0135 | .0027 |



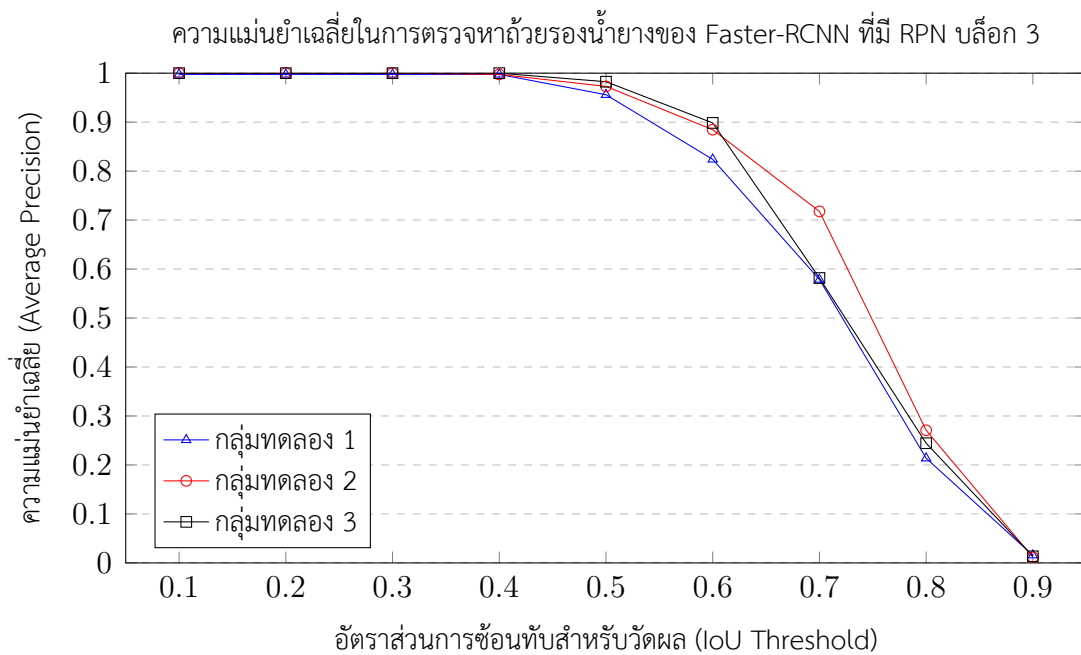
รูปที่ 3.23: ความแม่นยำเฉลี่ยในการตรวจหาถ้วยรองน้ำยางด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2 ในกลุ่มทดลองที่ 1



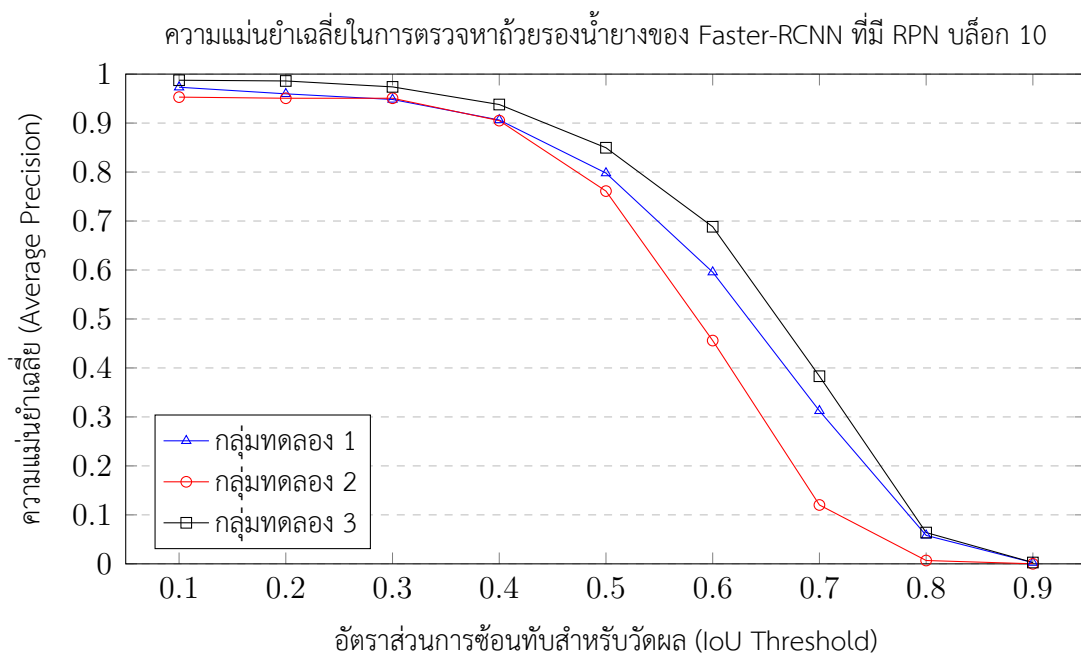
รูปที่ 3.24: ความแม่นยำเฉลี่ยในการตรวจหาถ้วยรองน้ำยางด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2 ในกลุ่มทดลองที่ 2



รูปที่ 3.25: ความแม่นยำเฉลี่ยในการตรวจหาถ้วยรองน้ำยางด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2 ในกลุ่มทดลองที่ 3



รูปที่ 3.26: ความแม่นยำเฉลี่ยในการตรวจหาถ้วยรองน้ำยางด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 ของ MobileNetV2 แสดงกลุ่มทดลองที่ 1, 2, และ 3



รูปที่ 3.27: ความแม่นยำเฉลี่ยในการตรวจหาถ้วยรองน้ำยางด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 10 ของ MobileNetV2 แสดงกลุ่มทดลองที่ 1, 2, และ 3

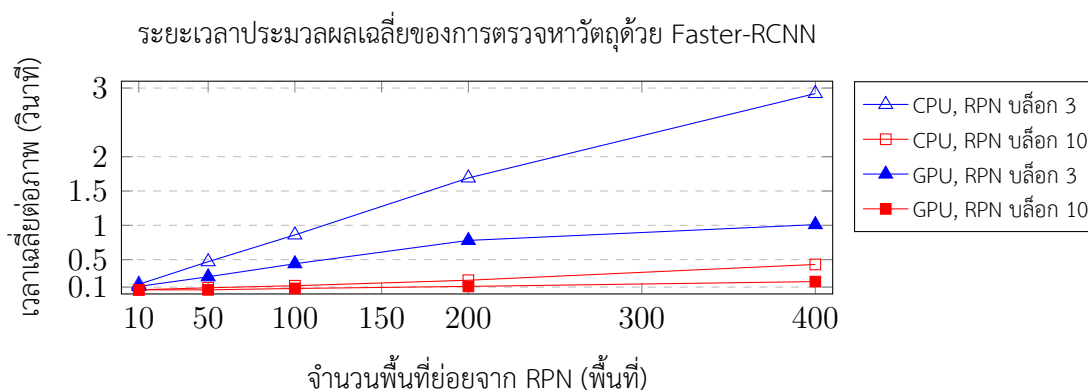
3.1.2.5 เวลาที่ใช้ประมวลผลต่อภาพ

จำนวนพื้นที่ย่อยที่ผลิตจาก RPN จะส่งผลต่อเวลาประมวลผลของเน็ตเวิร์กขณะตรวจหาวัตถุ โดยหากอนุญาตให้ RPN ผลิตพื้นที่ย่อยจำนวนมากจะทำให้ Faster-RCNN มีพื้นที่ในภาพเพื่อประมวลผลมากขึ้น และหากเน็ตเวิร์ก Faster-RCNN มีเลเยอร์ส่วนปลายของ RPN อยู่ในเลเยอร์ลำดับต้นของ MobileNetV2 (RPN บล็อก 3) พื้นที่ย่อยจาก RPN ทั้งหมดจะถูกคอนโวลูชันไปในเลเยอร์ลำดับสูงขึ้น ซึ่งเลเยอร์หลังจาก RPN มีจำนวนมากกว่าเน็ตเวิร์ก Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN อยู่ที่เลเยอร์ลำดับสูงกว่า (RPN บล็อก 10)

เวลาประมวลผลต่อภาพแสดงในตารางที่ 3.8 เมื่อกำหนดให้ Faster-RCNN คำนวณบนหน่วยประมวลผลกลางของเครื่อง (CPU) เพียงอย่างเดียว หรือคำนวณบนกราฟิกการ์ด (GPU) เพียงอย่างเดียว ทำการทดลองด้วยโปรแกรม MATLAB R2019a บนคอมพิวเตอร์ Core i7-8086K, RAM 32GB, GPU GTX1070Ti

ตารางที่ 3.8: เวลาเฉลี่ยที่ใช้ตรวจหาวัตถุต่อภาพของ Faster-RCNN ที่สร้างบน MobileNetV2 ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 3 และบล็อก 10 ในกรณีประมวลผลด้วย CPU และเมื่อประมวลผลด้วย GPU

| พื้นที่ย่อยต่อภาพ | CPU Time (วินาทีต่อภาพ) | | GPU Time (วินาทีต่อภาพ) | |
|-------------------|-------------------------|--------------|-------------------------|--------------|
| | RPN บล็อก 3 | RPN บล็อก 10 | RPN บล็อก 3 | RPN บล็อก 10 |
| 10 | 0.14 | 0.06 | 0.11 | 0.06 |
| 50 | 0.47 | 0.09 | 0.25 | 0.06 |
| 100 | 0.86 | 0.12 | 0.44 | 0.08 |
| 200 | 1.69 | 0.20 | 0.78 | 0.11 |
| 400 | 2.92 | 0.43 | 1.01 | 0.18 |



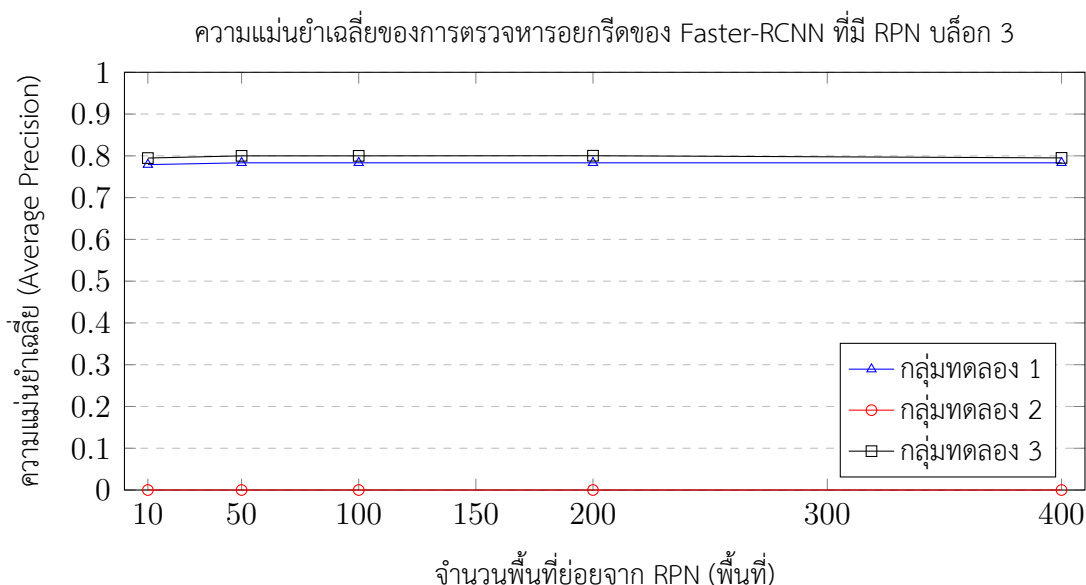
รูปที่ 3.28: ระยะเวลาประมวลผลเฉลี่ยต่อภาพของการตรวจหาวัตถุและถ่วงร่อนน้ำยาด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 และตำแหน่งบล็อก 10 ของ MobileNetV2 เมื่อประมวลผลด้วยหน่วยประมวลผลหลัก (CPU) หรือกราฟิกการ์ด (GPU)

3.1.2.6 ผลของจำนวนพื้นที่ย่อยจาก RPN หักกับความแม่นยำในการตรวจหารอยกริดและถ้วยรองน้ำยาง

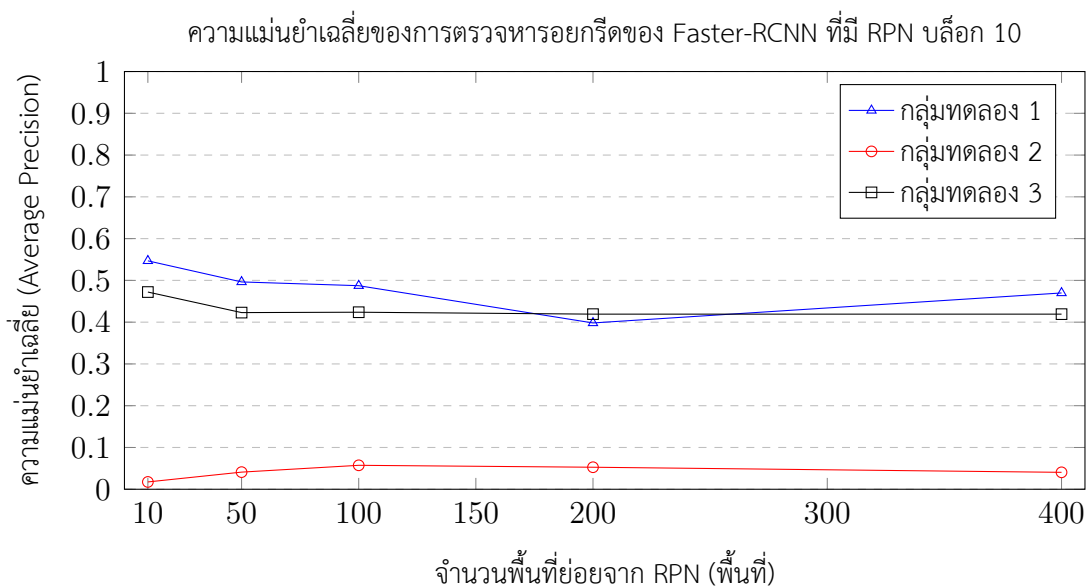
การทดสอบความแม่นยำการตรวจหารอยกริดและถ้วยรองน้ำยางตามข้อ 3.1.2.3 และ 3.1.2.4 ข้างต้น ได้กำหนดให้ Faster-RCNN แต่ละเน็ตเวิร์กมีจำนวนพื้นที่ย่อยจาก RPN ขณะตรวจหาวัตถุเท่ากับ 100 พื้นที่ย่อยต่อภาพ ทั้งนี้ได้ทำการทดลองเพิ่มเติมโดยปรับเปลี่ยนจำนวนพื้นที่ย่อยจาก RPN เพื่อสังเกตพฤติกรรมของเน็ตเวิร์กที่ส่งผลต่อความแม่นยำต่อการตรวจหา

กราฟค่าความแม่นยำเฉลี่ยของการตรวจหารอยกริดของแต่ละกลุ่มทดลอง เมื่อทดลองปรับจำนวนพื้นที่ย่อยจาก RPN ในขณะที่ Faster-RCNN ตรวจหาวัตถุเป็นจำนวนตั้งแต่ 10 ถึง 400 พื้นที่ย่อยต่อภาพ ของ Faster-RCNN ที่มีโครงสร้าง RPN บล็อก 3 และ RPN บล็อก 10 เมื่อวัดผลด้วยอัตราส่วนการซ้อนทับสำหรับวัตถุที่ 0.5 แสดงในรูปที่ 3.29 และ 3.30 และกราฟค่าความแม่นยำในการตรวจหาถ้วยรองน้ำยางเมื่อทดลองปรับจำนวนพื้นที่ย่อยจาก RPN ในขณะที่ Faster-RCNN ตรวจหาวัตถุเป็นจำนวนตั้งแต่ 10 ถึง 400 พื้นที่ย่อยต่อภาพ ของ Faster-RCNN ที่มี RPN บล็อก 3 และ RPN บล็อก 10 เมื่อวัดผลด้วยอัตราส่วนการซ้อนทับสำหรับวัตถุที่ 0.5 แสดงในรูปที่ 3.31 และ 3.32

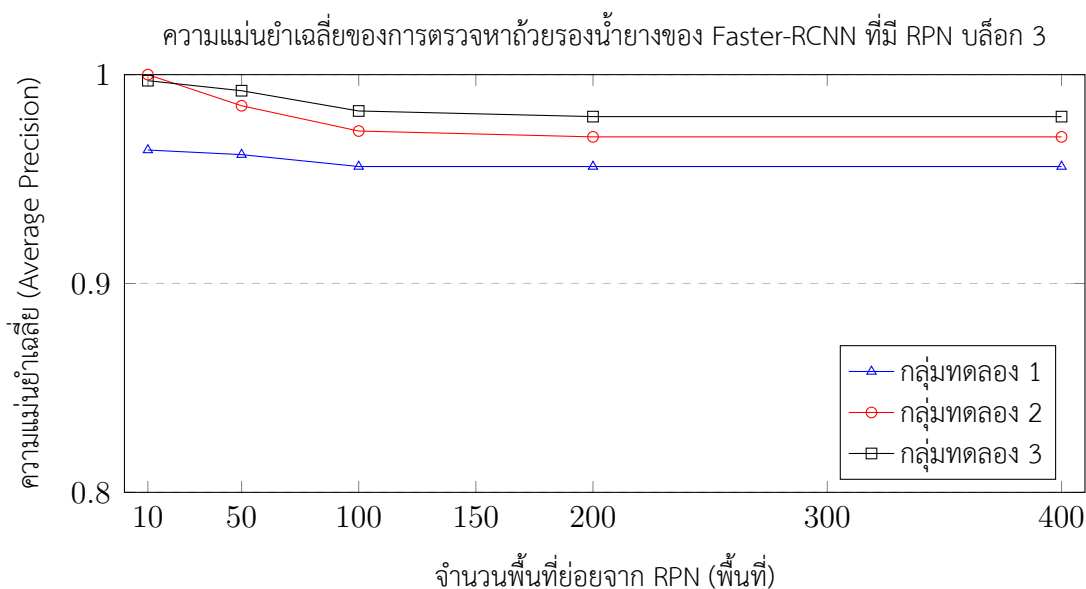
ผลการทดลองตรวจหารอยกริดในรูปที่ 3.29 และ 3.30, และการตรวจหาถ้วยรองน้ำยางในรูปที่ 3.31 และ 3.32 ของทั้ง 3 กลุ่มทดลอง ค่าความแม่นยำเฉลี่ยจะมีความแน่นอนมากขึ้นเมื่อให้จำนวนพื้นที่ย่อยจาก RPN มีค่ามากกว่า 100 พื้นที่ย่อยต่อภาพ



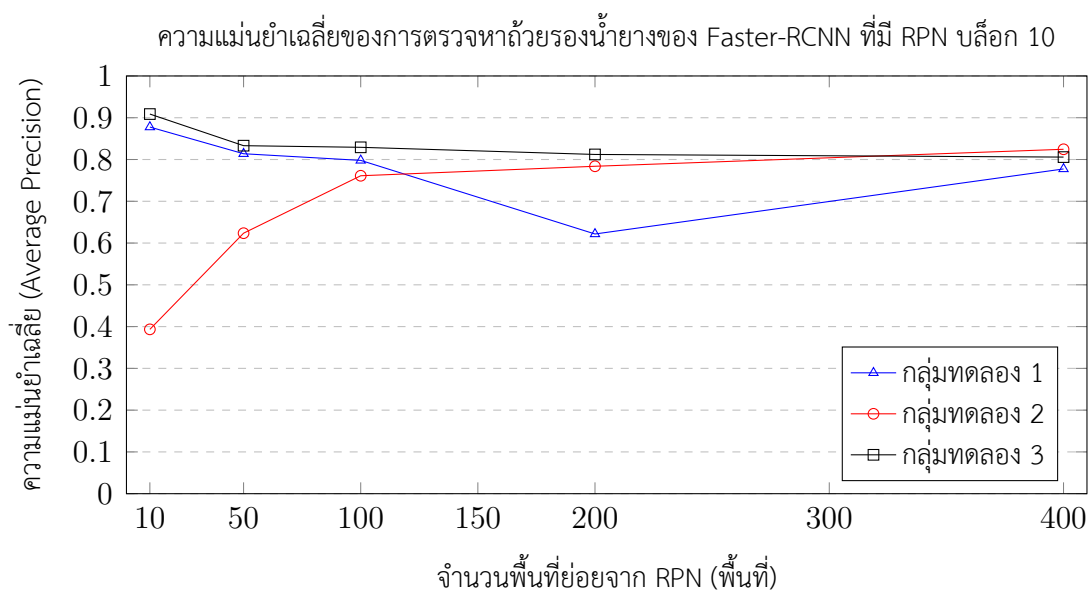
รูปที่ 3.29: ความแม่นยำเฉลี่ยของการตรวจหารอยกริดเมื่อเปลี่ยนจำนวนพื้นที่ย่อยจาก RPN ในขณะตรวจหาวัตถุด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 3 ของ MobileNetV2



รูปที่ 3.30: ความแม่นยำเฉลี่ยของการตรวจหารอยกริดเมื่อเปลี่ยนจำนวนพื้นที่ย่อยจาก RPN ในขณะที่ตรวจหาวัตถุด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 10 ของ MobileNetV2



รูปที่ 3.31: ความแม่นยำเฉลี่ยของการตรวจหาถ้วยรองน้ำยางเมื่อเปลี่ยนจำนวนพื้นที่ย่อยจาก RPN ในขณะที่ตรวจหาวัตถุด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 3 ของ MobileNetV2



รูปที่ 3.32: ความแม่นยำเฉลี่ยของการตรวจหาถ้วยร่อนน้ำยางเมื่อเปลี่ยนจำนวนพื้นที่ย่อยจาก RPN ในขณะที่ตรวจหาวัตถุด้วย Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 10 ของ MobileNetV2

3.1.2.7 สรุปผลการทดลอง: วิธีตรวจหารอยกริดและถ้วยร่องน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้างด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN

กลุ่มข้อมูลทดลอง ความแม่นยำเฉลี่ยในการตรวจหารอยกริดของแต่ละกลุ่มทดลองในรูปแบบที่ 3.21 และรูปที่ 3.22 แสดงให้เห็นว่าความแม่นยำเฉลี่ยของกลุ่มทดลองที่ 1 และ 3 ซึ่งคือภาพสี และภาพสีเมื่อถูกจำกัดพื้นที่ลำต้นด้วยภาพความลึก Faster-RCNN สามารถตรวจหารอยกริดได้สูงถึง 80.3% (โมเดลที่สร้างแบบ RPN บล็อก 3) ที่อัตราส่วนซ้อนทับสำหรับวัตถุผลที่ 0.5 ในขณะที่เน็ตเวิร์กที่ทำงานด้วยภาพถ่ายความลึกเพียงอย่างเดียวไม่สามารถตรวจหารอยกริดได้ เป็นผลจากเน็ตไม่สามารถสกัดหาพีเจอร์ที่บ่งชี้รอยกริดได้ หรือภาพถ่ายความลึกด้วยกล้องถ่ายภาพมุมข้างมีลักษณะของรอยกริดไม่เพียงพอต่อการตรวจหา สอดคล้องกับคุณสมบัติของโมดูลกล้องถ่ายภาพความลึก ของกล้องถ่ายภาพมุมกว้างซึ่งความผิดพลาดที่ระยะความลึกประมาณ 1 เมตร มีค่ามากกว่าความลึกเฉลี่ยของรอยกริดบนหน้ายาง (ความลึกระหว่างพื้นที่รอยกริดเดิมกับหน้ายางที่ยังไม่ผ่านการกริด)

การตรวจหาถ้วยร่องน้ำยางของทุกกลุ่มทดลองมีค่าความแม่นยำเฉลี่ยไม่ต่ำกว่า 95% โดยเฉพาะกลุ่มทดลองที่ 2 Faster-RCNN ที่ทำงานด้วยภาพถ่ายความลึกสามารถตรวจหารอยกริดได้สูงถึง 97.3% และกลุ่มทดลองที่ 3 ทำค่าความแม่นยำได้สูงสุดที่ 98.3% ที่อัตราส่วนการซ้อนทับสำหรับวัตถุผลที่ 0.5

หากเปรียบเทียบระหว่างกลุ่มทดลอง ความแม่นยำเฉลี่ยของการตรวจหารอยกริดและถ้วยร่องน้ำยางของกลุ่มทดลองที่ 3 มีค่าสูงที่สุด ทั้งนี้เป็นผลจากการจำกัดพื้นที่ลำต้นก่อนอินพุตภาพเข้า Faster-RCNN ทั้งในขั้นตอนการสอนและการตรวจหา การจำกัดพื้นที่ลำต้นของกลุ่มทดลองที่ 3 ทำให้ลดพื้นที่หลัง ซึ่งเป็นข้อมูลส่วนที่ไม่จำเป็นที่อาจมีความไม่แน่นอนสูง ทำให้เน็ตเวิร์กไม่ต้องเรียนรู้ลักษณะของพื้นที่หลังซับซ้อนเมื่อเปรียบเทียบกับกลุ่มทดลองที่ 1 ซึ่งคือภาพถ่ายสีตามปกติ

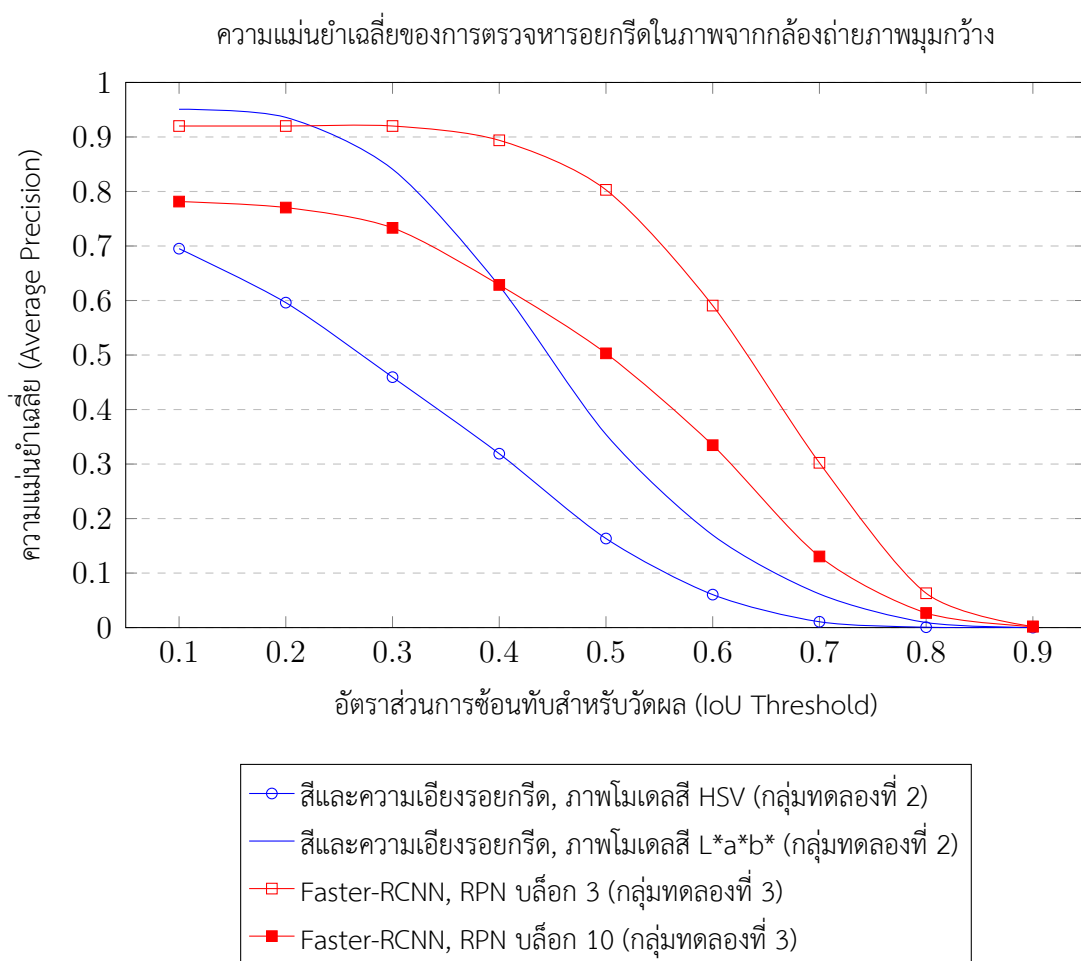
โครงสร้างของ Faster-RCNN ค่าลอสฟังก์ชันของการสอนด้วยวิธีการสอนสลับ 4 ขั้นตอน ตั้งแต่รูปที่ 3.12 ถึงรูปที่ 3.16 พบว่าค่าลอสฟังก์ชันของเน็ตเวิร์กตรวจหาวัตถุในขั้นตอนที่ 4 (หมายเลข 4) Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 3 (RPN บล็อก 3) มีค่าต่ำกว่าค่าลอสฟังก์ชันของ Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 10 (RPN บล็อก 10) ของกลุ่มทดลองที่ 1 และ 3 ที่จำนวนรอบการอัปเดตน้ำหนัก (Iteration) เท่ากัน บ่งชี้ถึงความสามารถในการทำนายรอยกริดและถ้วยร่องน้ำยางของเน็ตเวิร์กที่มี RPN บล็อก 3 ดีกว่าเน็ตเวิร์กที่มี RPN บล็อก 10

ผลการตรวจหารอยกริดเปรียบเทียบระหว่าง Faster-RCNN ที่มี RPN บล็อก 3 กับ Faster-RCNN ที่มี RPN บล็อก 10 แสดงในรูปที่ 3.18, 3.19, และ 3.20 และผลการตรวจหาถ้วยร่องน้ำยางแสดงในรูปที่ 3.23, 3.24, และ 3.25 ผลการทดลองแสดงให้เห็นว่า Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 3 สามารถตรวจหารอยกริดและถ้วยร่องน้ำยางได้ดีกว่า Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 10 เนื่องจาก Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 3 มีค่าความแม่นยำเฉลี่ยสูงกว่าในทุกอัตราส่วนการซ้อนทับสำหรับวัตถุผล

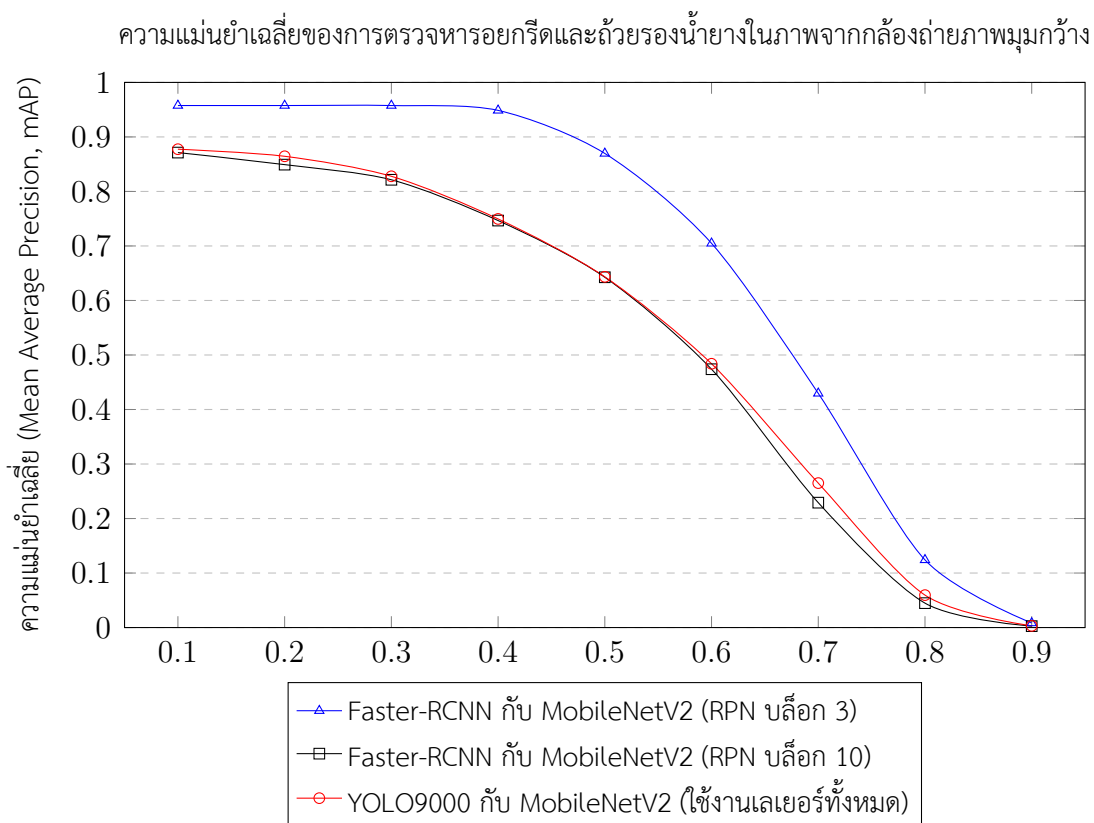
3.1.3 เปรียบเทียบผลการตรวจหารอยกรีดในภาพถ่ายจากกล้องถ่ายภาพมุมกว้างของทั้งสองวิธี

กราฟเปรียบเทียบความแม่นยำเฉลี่ยของการตรวจหารอยกรีดในภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง ในแต่ละอัตราส่วนการซ้อนทับสำหรับวัตถุด้วยวิธีตรวจหาทั้งสองวิธีที่นำเสนอในวิทยานิพนธ์ฉบับนี้ แสดงในรูปที่ 3.33 โดยแสดงผลของกลุ่มทดลองที่ให้ผลที่ดีที่สุด สามารถจัดอันดับความแม่นยำของการตรวจหารอยกรีดที่อัตราส่วนการซ้อนทับสำหรับวัตถุที่ 0.5 ได้ดังนี้

- 1 Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ที่ลำดับเลเยอร์บล็อกที่ 3 ของ MobileNetV2 ในกลุ่มข้อมูลทดลองที่ 3 ให้ความแม่นยำเฉลี่ย 80.30%
- 2 Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ที่ลำดับเลเยอร์บล็อกที่ 10 ของ MobileNetV2 ในกลุ่มข้อมูลทดลองที่ 3 ให้ความแม่นยำเฉลี่ย 50.31%
- 3 วิธีตรวจหารอยกรีดด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี เมื่อประมวลผลภาพโมเดลสี $L^*a^*b^*$ เมื่อกำหนดค่าพารามิเตอร์ที่เกี่ยวข้องตามกลุ่มทดลองที่ 2 ให้ความแม่นยำเฉลี่ย 35.43%
- 4 วิธีตรวจหารอยกรีดด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกรีดผ่านการจัดกลุ่มสี เมื่อประมวลผลภาพโมเดลสี HSV เมื่อกำหนดค่าพารามิเตอร์ที่เกี่ยวข้องตามกลุ่มทดลองที่ 2 ให้ความแม่นยำเฉลี่ย 16.34%



รูปที่ 3.33: ความแม่นยำเฉลี่ยในการตรวจหารอยกรีดของแต่ละวิธีในกลุ่มทดลองที่ดีที่สุด



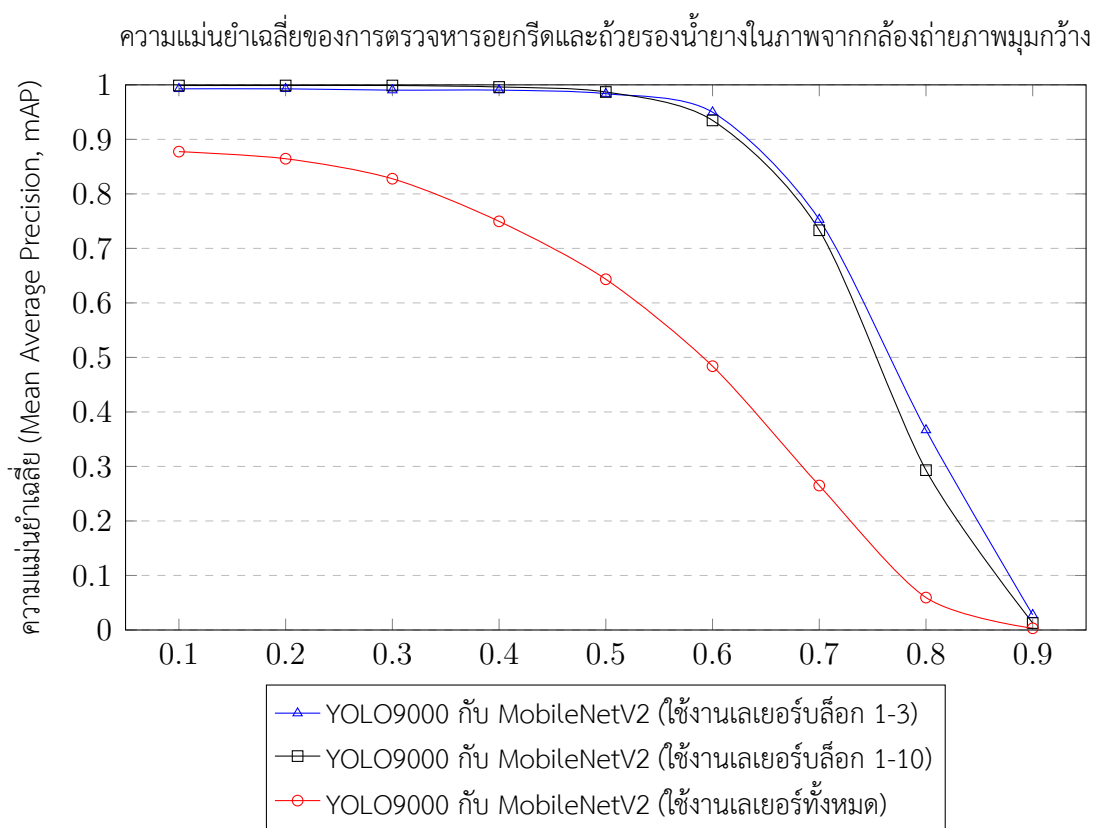
รูปที่ 3.34: ความแม่นยำเฉลี่ยในการตรวจหารอยกริดและถ้วยรองน้ำยาง (mAP) ระหว่าง Faster-RCNN กับ YOLOv2 ที่ใช้งานเลเยอร์ทั้งหมดของ MobileNetV2, เมื่อใช้ภาพสี RGB เป็นภาพอินพุต

3.1.4 กรณีเปรียบเทียบ: ผลการตรวจหารอยกริดและถ้วยรองน้ำยาง ระหว่างโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันชนิด YOLO กับชนิด Faster-RCNN เมื่อทำงานบน MobileNetV2

เน็ตเวิร์กที่มีโครงสร้างแบบส่วนเดียวถูกพัฒนาให้มีความสามารถในการตรวจหาวัตถุขนาดเล็กได้ดีขึ้น ตัวอย่างเช่น YOLO9000 [65] หรือ YOLOv2 โดยในกรณีเมื่อใช้เน็ตเวิร์กที่ผ่านการเรียนรู้มาก่อนเป็นฐาน สำหรับการปรับละเอียดเพื่อการตรวจหาวัตถุชุดข้อมูลเฉพาะ ดังการทดลองตัวอย่างการใช้งาน YOLO9000 ทำงานร่วมกับ MobileNetV2

กราฟเปรียบเทียบความแม่นยำเฉลี่ยในการตรวจหารอยกริดระหว่าง Faster-RCNN ที่มี RPN ในลำดับเลเยอร์บล็อก 3 ของ MobileNetV2 (RPN บล็อก 3) และ RPN ในลำดับเลเยอร์บล็อก 10 ของ MobileNetV2 (RPN บล็อก 10) กับ YOLO9000 ที่ทำงานร่วมกับเลเยอร์ของ MobileNetV2 ทั้งหมด แสดงในรูปที่ 3.34 ซึ่งทั้งสองชนิดโครงสร้างได้ใช้งานเลเยอร์ภายใน MobileNetV2 ครบทุกเลเยอร์ พบว่า Faster-RCNN ที่มี RPN บล็อก 3 มีความแม่นยำเฉลี่ยที่สูงกว่า YOLO9000 และ Faster-RCNN ที่มี RPN บล็อก 10 มีความแม่นยำเฉลี่ยใกล้เคียงกับ YOLO9000

แต่เมื่อลดจำนวนเลเยอร์ลำดับสูงของ MobileNetV2 โดยให้ YOLO9000 ตรวจสอบในเลเยอร์ในลำดับบล็อก 3 ของ MobileNetV2 และทดลองให้ตรวจสอบในเลเยอร์ในลำดับบล็อก 10 ของ MobileNetV2 กลับส่งผลให้ YOLO9000 ตรวจสอบรอยกริดและถ้วยร่อนน้ำยางมีความแม่นยำเฉลี่ยสูงกว่าการใช้เลเยอร์ทั้งหมดของ MobileNetV2 แสดงในรูปที่ 3.35 ทั้งนี้อาจเป็นผลจากจำนวนเลเยอร์ที่ลดลงอย่างมาก ทำให้สามารถสอนเน็ตเวิร์กได้รวดเร็วและแม่นยำมากขึ้น ทำให้เน็ตเวิร์กที่ผ่านการสอนมีความจำเพาะกับชุดข้อมูลรอยกริดและถ้วยร่อนน้ำยาง มากกว่าเน็ตเวิร์กที่มีโครงสร้างที่ใช้เลเยอร์ทั้งหมดของ MobileNetV2 (Overfitting) และอาจสรุปได้ว่าฟีเจอร์ที่บ่งบอกรอยกริดและถ้วยร่อนน้ำยางถูกสกัดออกได้ในเลเยอร์ลำดับต้นของ MobileNetV2 ทำให้การลดเลเยอร์ลำดับสูงของเน็ตเวิร์กเพื่อตรวจหาวัตถุในชุดข้อมูลเฉพาะ (Custom Dataset) เป็นประเด็นที่น่าสนใจในการศึกษาต่อไป



รูปที่ 3.35: ความแม่นยำเฉลี่ยของการตรวจหารอยกริดและถ้วยร่อนน้ำยาง (mAP) ด้วย YOLOv2 ที่ใช้งานเลเยอร์เฉพาะเลเยอร์บล็อก 1-3, เลเยอร์บล็อก 1-10 และเลเยอร์ทั้งหมดของ MobileNetV2 เมื่อภาพอินพุตคือภาพสี RGB

3.2 การทดลองภาพถ่ายจากกล้องระยะใกล้ (กล้อง 2)

การทดลองด้วยภาพถ่ายจากกล้องระยะใกล้ ประกอบด้วยการทดลองย่อย 2 การทดลอง ได้แก่ การทดลองวิธีปรับตำแหน่งแนวรอยกริดอ้างอิงจากแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ (ชื่อย่อย 2.4.2.2) และการทดลองวิธีตรวจหาแนวรอยกริดในภาพถ่ายจากกล้องระยะใกล้ (ชื่อย่อย 2.4.3)

3.2.1 ผลการทดลอง: วิธีปรับตำแหน่งแนวรอยกริดอ้างอิงจากแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ

การวัดผลการปรับตำแหน่งของแนวรอยกริดอ้างอิงด้วยค่า RSS , RLS , RQS ซึ่งอธิบายไว้ในขั้นตอนที่ 4 ของวิธีการปรับตำแหน่งรอยกริดอ้างอิงที่กำหนดด้วยมือ ชื่อย่อย 2.4.2.2 ตามสมการที่ 2.28, 2.29, และ 2.30 ตามลำดับ โดยกำหนดค่าคะแนนของแนวรอยกริดอ้างอิงที่ “ปรับตำแหน่งสำเร็จ” มีเงื่อนไขดังต่อไปนี้

- $RLS > 1.00$, เพื่อบ่งชี้ว่าแนวรอยกริดอ้างอิงที่ได้จากวิธีการปรับตำแหน่ง มีตำแหน่งตรงกับขอบตามแนวรอยกริดจริงในภาพถ่ายความลึกมากกว่าแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ และ
- $RSS \geq 1.00$, เพื่อบ่งชี้ว่าแนวรอยกริดอ้างอิงที่ได้จากวิธีการปรับตำแหน่ง มีรูปร่างตรงมากกว่าหรือเท่ากับแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ

3.2.1.1 ตัวแปรต้นของการทดลอง

วิธีการปรับตำแหน่งของแนวรอยกริดอ้างอิงที่กำหนดด้วยมือในขั้นตอนที่ 2 (อัลกอริทึม 2.4.2) และขั้นตอนที่ 3 (อัลกอริทึม 2.4.3) ต่างมีตัวแปร sr ควบคุมขนาดของกรอบการค้นหาที่ส่งผลต่อการทำงานของอัลกอริทึมทั้งสอง ดังนั้น sr คือตัวแปรต้นของการทดลอง

ผลการปรับตำแหน่งแนวรอยกริดอ้างอิงที่กำหนดด้วยมือด้วยอัลกอริทึมที่ใช้ค่า sr แบบค่าเดียว (ในขั้นตอนที่ 1 ถึง 4 ของวิธีการปรับตำแหน่งแนวรอยกริดอ้างอิง) แสดงในตารางที่ 3.9 และผลการปรับตำแหน่งแนวรอยกริดอ้างอิงที่กำหนดด้วยมือด้วยอัลกอริทึม 2.4.4 ซึ่งจะหาแนวรอยกริดที่ดีที่สุดในช่วงกรอบค้นหา ($sr = \{1, 2, 3, \dots, sr_{max}\}$) (ในขั้นตอนที่ 1 ถึง 5 ของวิธีการปรับตำแหน่งแนวรอยกริดอ้างอิง) แสดงในตารางที่ 3.10

3.2.1.2 ผลการปรับตำแหน่งแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ

กราฟจำนวนแนวรอยกริดอ้างอิงที่กำหนดด้วยมือที่ถูกปรับตำแหน่งได้สำเร็จ แสดงในรูปที่ 3.36 โดยแสดงผลในกรณีที่อัลกอริทึมใช้ค่า sr แบบค่าเดียว และผลในกรณีที่ใช้ค่า sr แบบช่วง ซึ่งจะหาแนวรอยกริดอ้างอิงที่ถูกปรับตำแหน่งที่ดีที่สุดในช่วง $[1, sr_{max}]$

ผลการทดลองการปรับตำแหน่งด้วยค่า sr แบบค่าเดียว พบว่าอัลกอริทึมสามารถปรับตำแหน่งแนวรอยกริดอ้างอิงที่กำหนดด้วยมือได้สำเร็จถึง 80% เมื่อค่า sr อยู่ในช่วง 2 ถึง 10 พิกเซล แต่หากให้ค่า sr เพิ่มสูงขึ้น ปริมาณแนวรอยกริดอ้างอิงที่ปรับตำแหน่งได้สำเร็จจะลดลงอย่างต่อเนื่อง แต่ในการค้นหา

แนวรอยกริดที่ดีที่สุดในช่วง 1 ถึง sr_{max} แนวรอยกริดสามารถถูกปรับแต่งได้สำเร็จถึง 90% ในช่วง 20 พิกเซล แรก และขยับเพิ่มขึ้นเล็กน้อยเท่านั้นหากเพิ่มค่า sr_{max} ให้มากขึ้น

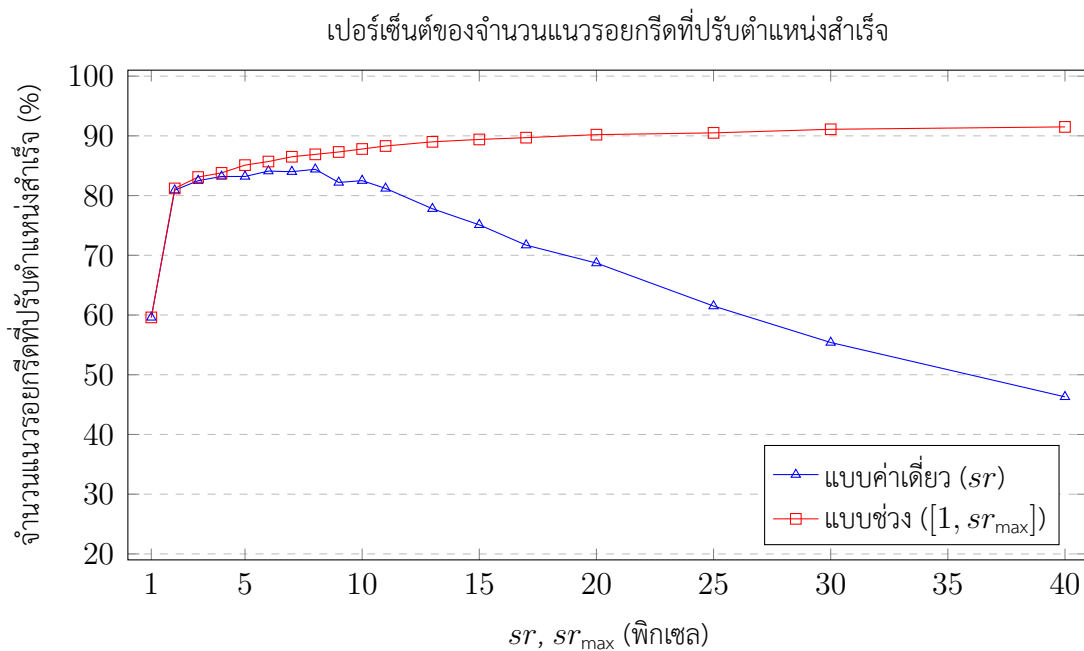
กราฟแสดงค่า RQS , RLS , RSS เฉลี่ยของแนวรอยกริดที่ปรับตำแหน่งสำเร็จ แสดงในรูปที่ 3.37, 3.38, และ 3.39 ตามลำดับ จากการทดลองพบว่าหากค่า sr สูงจะทำให้ค่า RSS เฉลี่ยลดลง เนื่องจากแนวรอยกริดอ้างอิงจะสูญเสียความตรง (Straightness) ในขณะเดียวกันจะทำให้ RLS มีค่าเพิ่มมากขึ้นเพราะสามารถย้ายแนวรอยกริดอ้างอิงให้ตรงตำแหน่งได้ดีขึ้น แต่ทั้งนี้เป็นผลเฉลี่ยจากจำนวนแนวรอยกริดอ้างอิงที่ปรับตำแหน่งได้สำเร็จลดลง

ตารางที่ 3.9: ผลการปรับตำแหน่งแนวรอยกริด ในกรณี sr แบบค่าเดียว

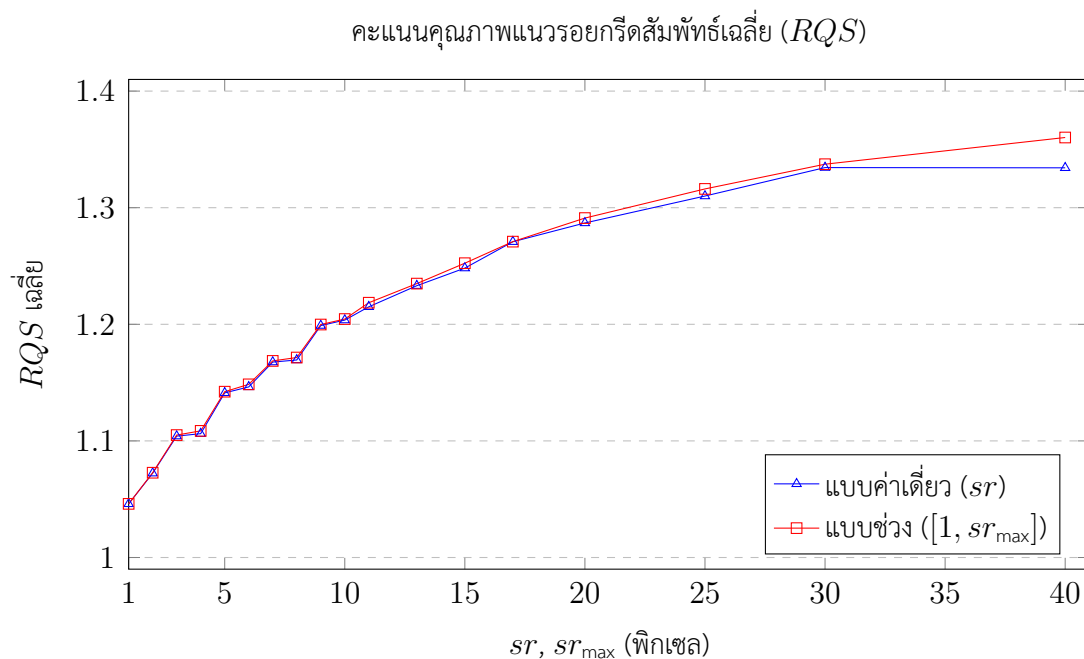
| sr (พิกเซล) | จำนวนปรับตำแหน่งสำเร็จ (%) | ในกรณีปรับตำแหน่งสำเร็จ | | |
|---------------|----------------------------|-------------------------|--------------|--------------|
| | | RSS เฉลี่ย | RLS เฉลี่ย | RQS เฉลี่ย |
| 1 | 59.6 | .9940 | 1.0391 | 1.0460 |
| 2 | 80.9 | .9950 | 1.0666 | 1.0723 |
| 3 | 82.5 | .9949 | 1.0980 | 1.1041 |
| 4 | 83.2 | .9949 | 1.1002 | 1.1062 |
| 5 | 83.6 | .9949 | 1.1348 | 1.1412 |
| 6 | 84.1 | .9948 | 1.1398 | 1.1463 |
| 7 | 84.0 | .9951 | 1.1613 | 1.1676 |
| 8 | 84.4 | .9951 | 1.1630 | 1.1694 |
| 9 | 82.2 | .9952 | 1.1924 | 1.1989 |
| 10 | 82.5 | .9952 | 1.1971 | 1.2036 |
| 11 | 81.2 | .9954 | 1.2086 | 1.2150 |
| 13 | 77.8 | .9959 | 1.2272 | 1.2331 |
| 15 | 75.1 | .9964 | 1.2427 | 1.2482 |
| 17 | 71.7 | .9966 | 1.2654 | 1.2708 |
| 20 | 68.7 | .9968 | 1.2820 | 1.2869 |
| 25 | 61.5 | .9982 | 1.3067 | 1.3100 |
| 30 | 55.4 | .9980 | 1.3309 | 1.3344 |
| 40 | 46.3 | .9995 | 1.3320 | 1.3342 |

ตารางที่ 3.10: ผลการปรับตำแหน่งแนวรอยกริด ในกรณี sr เป็นช่วง $[1, sr_{max}]$

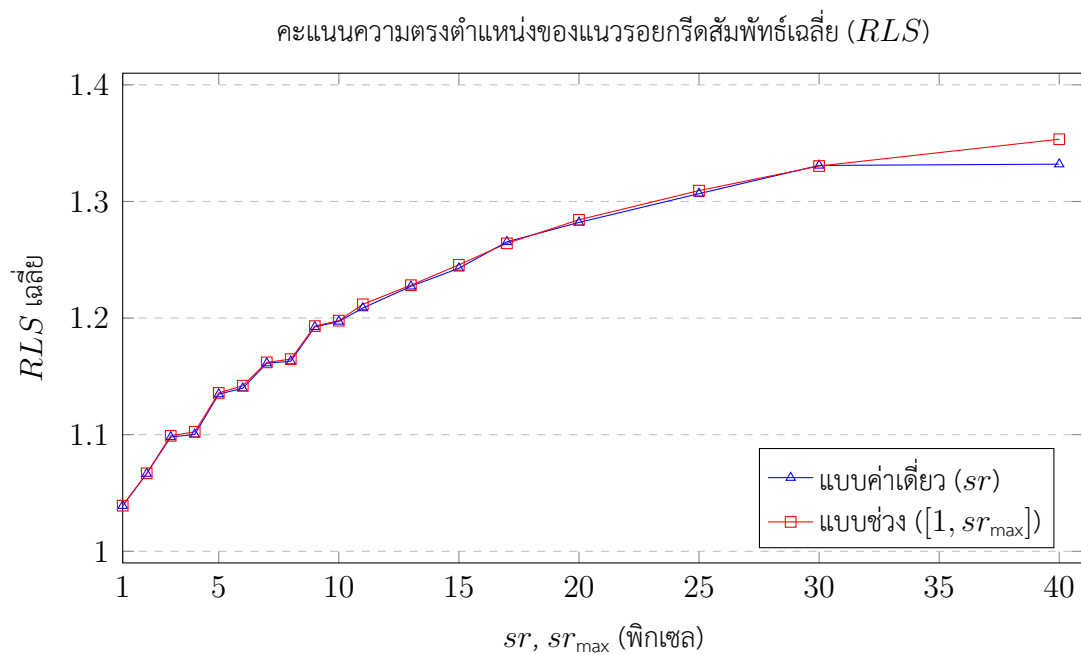
| sr_{max} (พิทเซล) | จำนวนปรับตำแหน่งสำเร็จ (%) | ในกรณีปรับตำแหน่งสำเร็จ | | |
|---------------------|----------------------------|-------------------------|--------------|--------------|
| | | RSS เฉลี่ย | RLS เฉลี่ย | RQS เฉลี่ย |
| 1 | 59.6 | .9940 | 1.0391 | 1.0460 |
| 2 | 81.2 | .9952 | 1.0669 | 1.0727 |
| 3 | 83.1 | .9952 | 1.0990 | 1.1050 |
| 4 | 83.8 | .9952 | 1.1025 | 1.1086 |
| 5 | 85.1 | .9953 | 1.1358 | 1.1422 |
| 6 | 85.7 | .9954 | 1.1420 | 1.1485 |
| 7 | 86.5 | .9955 | 1.1620 | 1.1686 |
| 8 | 86.9 | .9955 | 1.1648 | 1.1715 |
| 9 | 87.3 | .9959 | 1.1931 | 1.1998 |
| 10 | 87.8 | .9959 | 1.1978 | 1.2045 |
| 11 | 88.3 | .9961 | 1.2118 | 1.2185 |
| 13 | 89.0 | .9964 | 1.2282 | 1.2349 |
| 15 | 89.4 | .9968 | 1.2457 | 1.2524 |
| 17 | 89.7 | .9970 | 1.2641 | 1.2709 |
| 20 | 90.2 | .9972 | 1.2843 | 1.2911 |
| 25 | 90.5 | .9975 | 1.3093 | 1.3161 |
| 30 | 91.1 | .9977 | 1.3304 | 1.3373 |
| 40 | 91.5 | .9981 | 1.3534 | 1.3602 |



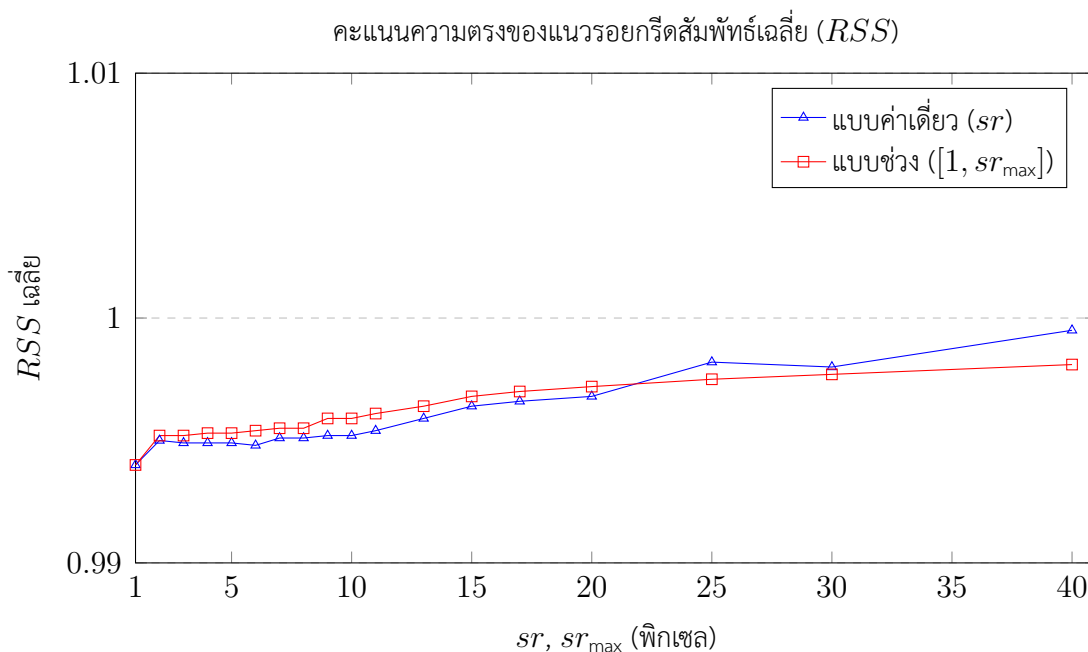
รูปที่ 3.36: เปอร์เซ็นต์ของจำนวนแนวรอยกริดที่ปรับตำแหน่งได้สำเร็จ ในกรณีเมื่อกำหนดค่า sr แบบค่าเดียว และผลการปรับตำแหน่งสำเร็จกรณีเมื่อกำหนดค่า sr แบบช่วง



รูปที่ 3.37: ค่าคะแนนคุณภาพแนวรอยกริดสัมพัทธ์เฉลี่ย (RQS) ของแนวรอยกริดที่ปรับตำแหน่งสำเร็จ ในกรณีการกำหนดค่า sr แบบค่าเดียว และผลการปรับตำแหน่งสำเร็จกรณีกำหนดค่า sr แบบช่วง



รูปที่ 3.38: คะแนนความตรงตำแหน่งแนวรอยกริดสัมพัทธ์เฉลี่ย (RLS) ของแนวรอยกริดที่ปรับตำแหน่งได้สำเร็จ ในกรณีกำหนดค่า sr แบบค่าเดียว และผลการปรับตำแหน่งสำเร็จกรณีกำหนดค่า sr แบบช่วง



รูปที่ 3.39: คะแนนความตรงของแนวรอยกรีดสัมพันธ์เฉลี่ย (RSS) ของแนวรอยกรีดที่ปรับตำแหน่งสำเร็จ ในกรณีกำหนดค่า sr แบบค่าเดี่ยว และผลการปรับตำแหน่งสำเร็จกรณีใช้ค่า sr แบบช่วง

3.2.1.3 สรุปผลการทดลอง: วิธีปรับตำแหน่งแนวรอยกรีดอ้างอิงจากแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือ

ผลการทดลองแสดงให้เห็นว่าการกำหนดแนวรอยกรีดอ้างอิงด้วยมือ แนวรอยกรีดที่ได้จะมีความคลาดเคลื่อนจากขอบความลึกตามแนวรอยกรีด ที่ปรากฏในภาพถ่ายความลึก ซึ่งปรากฏชัดเจนจากผลการทดลองที่อนุญาตให้แนวรอยกรีดที่จะถูกปรับตำแหน่งให้สามารถย้ายไปจากตำแหน่งเดิมที่กำหนดด้วยมือ ในระยะที่เป็นพิกเซลระหว่าง 1 ถึง 20 พิกเซลแรก อัลกอริทึมสามารถปรับตำแหน่งแนวรอยกรีดอ้างอิงได้สำเร็จในจำนวนสูงสุดถึง 83% เมื่อ sr มีค่าระหว่าง 2 ถึง 10 พิกเซล (แบบค่าเดี่ยว) และสามารถปรับตำแหน่งแนวรอยกรีดได้สำเร็จสะสมสูงถึง 90% (sr แบบช่วง) ซึ่งสอดคล้องกับผลการทดลองเมื่อในกรณีกำหนดกรอบการค้นหาที่ทำให้แนวรอยกรีดที่จะถูกปรับตำแหน่งสามารถเคลื่อนที่ไปจากเดิมมากที่สุดในระยะแบบคงที่ (sr ค่าเดี่ยว) ซึ่งผลการปรับตำแหน่งแนวรอยกรีดสำเร็จส่วนใหญ่จะมีค่าสูงในช่วง 1 ถึง 20 พิกเซลแรก โดยแนวรอยกรีดอ้างอิงใหม่ที่ได้ยังคงรักษารูปร่างของแนวรอยกรีดให้มีความตรงที่มีมากกว่าหรือเท่ากับแนวรอยกรีดอ้างอิงที่กำหนดด้วยมือ

ผลการทดลองไปเป็นตามสมมติฐานที่ว่า การกำหนดตำแหน่งแนวรอยกรีดอ้างอิงด้วยมือมีความคลาดเคลื่อนในระดับพิกเซล และผลของฟิลเตอร์เบลอภาพแบบเกาส์ ทำให้สัญญาณความถี่สูงในภาพถูกรวมเข้าด้วยกันทั้งในด้านตำแหน่งและขนาดของสัญญาณ และจากการทดลองเมื่อเพิ่ม sr ให้มีค่าสูงขึ้น จำนวนแนวรอยกรีดที่เข้าเงื่อนไขปรับตำแหน่งสำเร็จจะลดลง เป็นผลจากสัญญาณความถี่สูงบริเวณขอบตามแนวรอยกรีดในภาพจะถูกเกลี่ย (Smooth) มากจนเกินไป จนเพิ่มโอกาสที่จะทำให้พิกเซลของแนวรอยกรีดพลาดจากตำแหน่งที่ควร ซึ่งเป็นผลสืบเนื่องจากการค้นหาที่มีขนาดใหญ่ สังเกตได้จากค่าเฉลี่ย RSS ของแนวรอยกรีดอ้างอิงที่ปรับตำแหน่งสำเร็จเพิ่มสูงขึ้นเมื่อค่า sr มากขึ้น ในรูปที่ 3.39

3.2.2 ผลการทดลอง: วิธีตรวจหาแนวรอยกริดในภาพถ่ายระยะใกล้

3.2.2.1 ตัวแปรต้นของการทดลอง

วิธีตรวจหาแนวรอยกริดในภาพถ่ายจากกล้องระยะใกล้ (2.4.3) มีตัวแปรควบคุมการทำงานทั้งสิ้น 2 ค่า ได้แก่ T_{up} ซึ่งควบคุมความชันในทิศทางชี้ขึ้นของลำดับพิกเซลที่เป็นแนวรอยกริดที่ปรากฏในภาพ และ T_{down} ทำหน้าที่ควบคุมความชันในทิศทางชี้ลงของลำดับพิกเซลที่เป็นแนวรอยกริดที่ปรากฏในภาพ

3.2.2.2 ผลการตรวจหาแนวรอยกริด

ผลการตรวจหาแนวรอยกริดในภาพจะแบ่งออกเป็น 2 ขั้นตอนที่ต่อเนื่องกัน ในขั้นแรกแนวรอยกริดจะถูกวัดผลการตรวจหาด้วยกรอบสี่เหลี่ยมล้อมรอบ (ตัวอย่างเช่นเดียวกับรูปที่ 2.62) ด้วยอัตราส่วนการซ้อนทับสำหรับวัดผลที่ 0.5 ในการบ่งชี้การตรวจพบ ซึ่งสามารถวัดผลด้วยค่าความแม่นยำเฉลี่ย จากนั้นวัดค่าระยะห่างของแนวรอยกริดที่ตรวจพบกับแนวรอยกริดอ้างอิง ด้วยระยะทางแบบเฮาส์ดอร์ฟฟ์ (Hausdorff Distance) ผลการทดลองแสดงดังรูปที่ 3.40 และรูปที่ 3.41

ผลการทดลองในรูปที่ 3.41 พบว่าค่าของระยะทางแบบเฮาส์ดอร์ฟฟ์ของแนวรอยกริดที่ถูกตรวจพบ จะมีค่าน้อยที่สุด ประมาณ 13 พิกเซล ซึ่งเป็นค่าที่ตัวแปรควบคุมทั้งสองมีค่าอยู่ในช่วง $5 \leq T_{down} \leq 10$ และ $2 \leq T_{up} \leq 8$ และเป็นช่วงที่ทำให้ค่าความแม่นยำเฉลี่ยของการตรวจหาตำแหน่งแนวรอยกริดด้วยกรอบสี่เหลี่ยมล้อมรอบ เท่ากับ 0.89 ถึง 0.90 หรือเฉลี่ยประมาณ 89.5%

3.2.2.3 สรุปผลการทดลอง

ผลการทดลองพบว่าค่า T_{up} , T_{down} ที่เหมาะสมจะอยู่ในช่วงตั้งแต่ 3 พิกเซลจนถึง 10 พิกเซล ซึ่งเป็นช่วงที่ทำให้ความคลาดเคลื่อนระหว่างแนวรอยกริดที่ถูกตรวจพบกับแนวรอยกริดอ้างอิงมีค่าเฉลี่ยเท่ากับ 13 พิกเซล ซึ่งเป็นค่าน้อยกว่าบริเวณอื่นโดยรอบ และช่วงดังกล่าวเป็นค่าที่ทำให้ความแม่นยำเฉลี่ยของการตรวจพบแนวรอยกริดมีค่าประมาณ 89.9%

ความแม่นยำเฉลี่ยของการตรวจหาตำแหน่งแนวรอยกริดด้วยกรอบสี่เหลี่ยมล้อมรอบ

| | | | | | | | | | | | | | |
|-------------------------------|----|--------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----|
| T_{down} (พิกเซล) | 13 | 0.873 | 0.895 | 0.899 | 0.902 | 0.906 | 0.908 | 0.911 | 0.911 | 0.911 | 0.911 | 0.911 | |
| | 12 | 0.873 | 0.895 | 0.898 | 0.901 | 0.905 | 0.906 | 0.909 | 0.910 | 0.910 | 0.910 | 0.910 | |
| | 11 | 0.873 | 0.895 | 0.898 | 0.901 | 0.905 | 0.906 | 0.908 | 0.909 | 0.909 | 0.909 | 0.909 | |
| | 10 | 0.873 | 0.895 | 0.898 | 0.901 | 0.905 | 0.906 | 0.908 | 0.908 | 0.908 | 0.908 | 0.908 | |
| | 9 | 0.873 | 0.895 | 0.898 | 0.900 | 0.904 | 0.905 | 0.907 | 0.907 | 0.907 | 0.907 | 0.907 | |
| | 8 | 0.873 | 0.893 | 0.897 | 0.899 | 0.901 | 0.902 | 0.903 | 0.903 | 0.903 | 0.903 | 0.903 | |
| | 7 | 0.873 | 0.892 | 0.895 | 0.897 | 0.898 | 0.898 | 0.899 | 0.899 | 0.899 | 0.899 | 0.899 | |
| | 6 | 0.873 | 0.892 | 0.893 | 0.893 | 0.893 | 0.894 | 0.894 | 0.894 | 0.894 | 0.894 | 0.894 | |
| | 5 | 0.864 | 0.878 | 0.879 | 0.879 | 0.879 | 0.880 | 0.880 | 0.880 | 0.880 | 0.880 | 0.880 | |
| | 4 | 0.856 | 0.867 | 0.867 | 0.867 | 0.867 | 0.868 | 0.868 | 0.868 | 0.868 | 0.868 | 0.868 | |
| | 3 | 0.831 | 0.838 | 0.838 | 0.838 | 0.838 | 0.838 | 0.838 | 0.838 | 0.838 | 0.838 | 0.838 | |
| | 2 | 0.725 | 0.728 | 0.728 | 0.728 | 0.728 | 0.727 | 0.727 | 0.727 | 0.727 | 0.727 | 0.727 | |
| | 1 | 0.160 | 0.160 | 0.160 | 0.160 | 0.159 | 0.159 | 0.159 | 0.159 | 0.159 | 0.159 | 0.159 | |
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | | T_{up} (พิกเซล) | | | | | | | | | | | |

รูปที่ 3.40: ความแม่นยำเฉลี่ย (Average Precision) ของการตรวจหาตำแหน่งแนวรอยกริดด้วยกรอบสี่เหลี่ยมที่อัตราส่วนการซ้อนทับสำหรับวัดผลเท่ากับ 0.5

ระยะทางแบบเฮาส์ดอร์ฟฟ์เฉลี่ย (พิกเซล)

| | | | | | | | | | | | | | |
|-------------------------------|----|--------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| T_{down} (พิกเซล) | 13 | 15.87 | 14.68 | 14.35 | 14.34 | 14.24 | 14.29 | 14.49 | 14.53 | 14.57 | 14.65 | 14.84 | 14.90 |
| | 12 | 15.59 | 14.40 | 14.05 | 14.07 | 13.94 | 13.94 | 14.14 | 14.23 | 14.28 | 14.36 | 14.55 | 14.61 |
| | 11 | 15.42 | 14.20 | 13.85 | 13.86 | 13.73 | 13.73 | 13.87 | 13.97 | 14.02 | 14.09 | 14.27 | 14.34 |
| | 10 | 15.15 | 13.91 | 13.54 | 13.56 | 13.44 | 13.45 | 13.59 | 13.63 | 13.68 | 13.76 | 13.94 | 14.00 |
| | 9 | 14.88 | 13.70 | 13.33 | 13.30 | 13.20 | 13.21 | 13.35 | 13.39 | 13.44 | 13.52 | 13.71 | 13.78 |
| | 8 | 14.73 | 13.52 | 13.16 | 13.14 | 13.08 | 13.12 | 13.23 | 13.30 | 13.35 | 13.43 | 13.62 | 13.69 |
| | 7 | 14.56 | 13.34 | 12.95 | 12.95 | 12.88 | 12.91 | 13.07 | 13.14 | 13.19 | 13.27 | 13.46 | 13.53 |
| | 6 | 14.52 | 13.41 | 13.13 | 13.16 | 13.12 | 13.15 | 13.25 | 13.32 | 13.37 | 13.46 | 13.65 | 13.72 |
| | 5 | 14.63 | 13.82 | 13.56 | 13.63 | 13.58 | 13.60 | 13.71 | 13.78 | 13.84 | 13.94 | 14.13 | 14.19 |
| | 4 | 16.18 | 15.59 | 15.40 | 15.46 | 15.45 | 15.46 | 15.58 | 15.62 | 15.67 | 15.78 | 15.99 | 16.03 |
| | 3 | 20.35 | 20.11 | 19.99 | 20.03 | 20.06 | 19.99 | 20.09 | 20.13 | 20.18 | 20.28 | 20.43 | 20.47 |
| | 2 | 30.09 | 29.90 | 29.90 | 29.94 | 29.96 | 29.92 | 29.90 | 29.95 | 30.00 | 30.04 | 30.12 | 30.16 |
| | 1 | 57.22 | 57.22 | 57.22 | 57.22 | 57.10 | 57.07 | 57.07 | 57.07 | 57.07 | 57.07 | 57.07 | 57.07 |
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | | T_{up} (พิกเซล) | | | | | | | | | | | |

รูปที่ 3.41: ระยะทางแบบเฮาส์ดอร์ฟฟ์ (Hausdorff Distance) โดยเฉลี่ย ในหน่วยพิกเซลระหว่างแนวรอยกริดที่ตรวจพบกับแนวรอยกริดอ้างอิง

บทที่ 4

บทวิจารณ์

ระบบมองเห็นของหุ่นยนต์ใช้งานกล้องถ่ายภาพเป็นอุปกรณ์ตรวจหาตำแหน่งรอยกริดและถ้วยรองน้ำยาง โดยกำหนดกล้องถ่ายภาพในสองระยะ ได้แก่ ภาพถ่ายมุมกว้างเป็นภาพถ่ายที่มีรอยกริดและครอบคลุมหน้ายางได้ทั้งหมดในพื้นที่ลำต้น และภาพถ่ายรอยกริดระยะใกล้เพื่อชดเชยความละเอียดของรอยกริดที่สูญเสียไปจากภาพมุมกว้าง

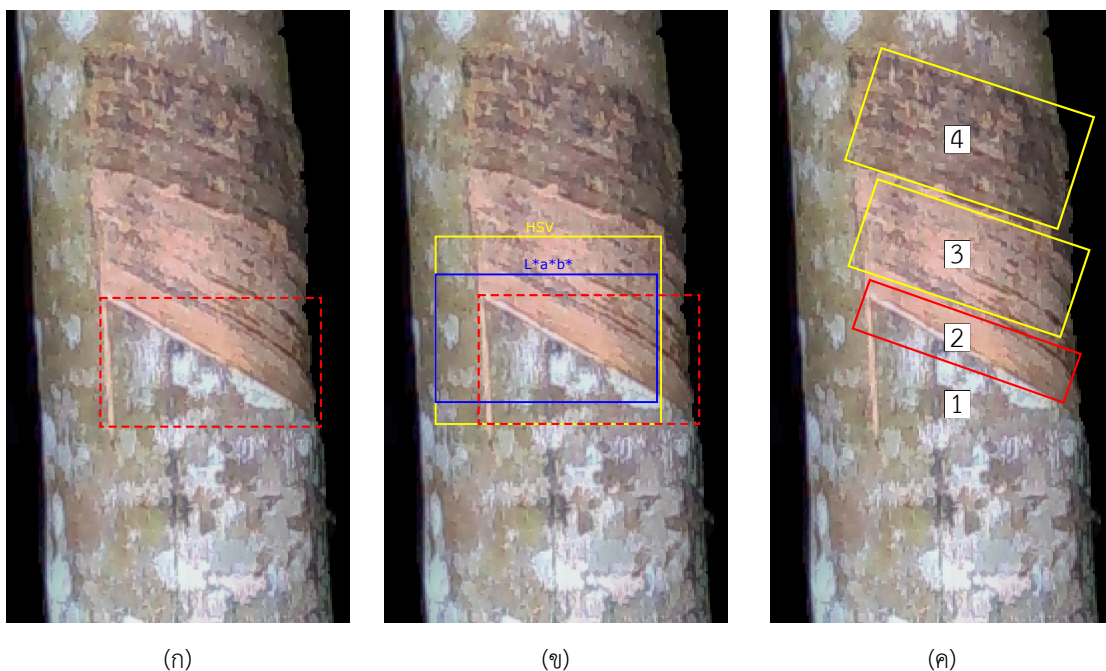
4.1 การตรวจหาแนวรอยกริดในภาพถ่ายมุมกว้าง

การตรวจหารอยกริดและถ้วยรองน้ำยางในภาพมุมกว้าง คือการตรวจหากรอบสี่เหลี่ยมที่ล้อมรอบรอยกริดและกรอบสี่เหลี่ยมล้อมรอบถ้วยรองน้ำยางในภาพ ซึ่งวิทยานิพนธ์นี้ได้นำเสนอวิธีตรวจโดย 2 วิธีที่แตกต่างกัน ได้แก่ 1.วิธีตรวจหารอยกริดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี่ และ 2.วิธีตรวจหารอยกริดและถ้วยรองน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้างด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN

4.1.1 การตรวจหารอยกริดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี่

การสร้างกรอบสี่เหลี่ยมเริ่มต้นเพื่อตรวจหารอยกริดที่ต่างกันอย่างสิ้นเชิงในแต่ละภาพ โดยอ้างอิงกับความกว้างของลำต้นที่ปรากฏในภาพนั้นๆ และทำให้กรอบสี่เหลี่ยมเริ่มต้นครอบคลุมพื้นที่เฉพาะภายในลำต้น ส่งผลให้การวัดผลด้วยอัตราส่วนการซ้อนทับสำหรับวัดผลที่ระดับค่ามีประโยชน์มากขึ้น เพราะสามารถยืนยันได้ว่าตำแหน่งกรอบสี่เหลี่ยมล้อมรอบรอยกริดที่ได้ อาจผิดพลาดตามความสูงของลำต้นเท่านั้น เมื่อเทียบกับการสร้างกรอบสี่เหลี่ยมเริ่มต้นแบบทั่วทั้งภาพ

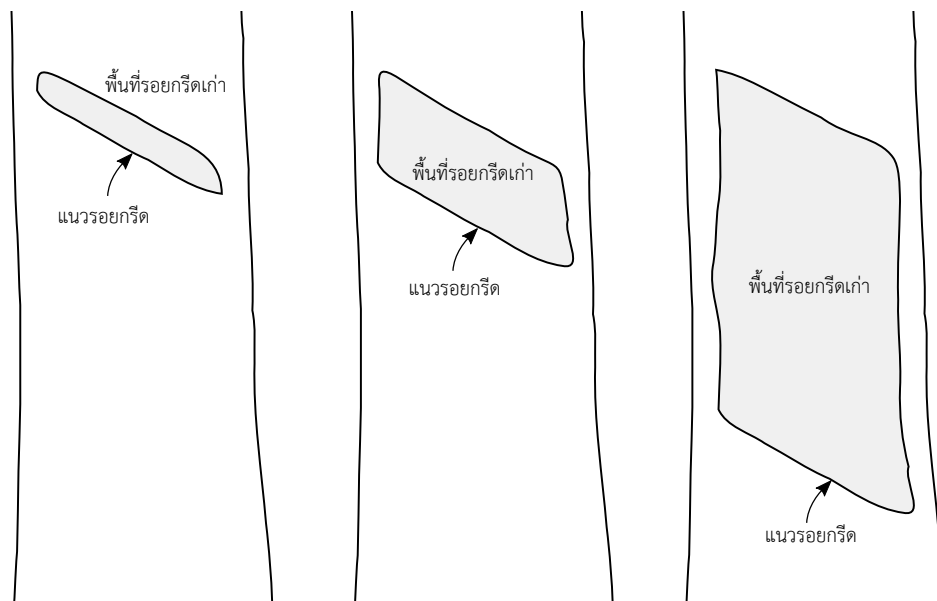
แม้ว่าวิธีการตรวจหารอยกริดการจัดกลุ่มสี่และความเอียงของรอยกริดร่วมกับการใช้งานกรอบสี่เหลี่ยมล้อมรอบสามารถระบุตำแหน่งของรอยกริดบนลำต้นได้ แต่ยังพบข้อพร่องในการระบุตำแหน่งของกรอบสี่เหลี่ยมล้อมรอบรอยกริด (Localization) ซึ่งเป็นผลจากการใช้ค่าคะแนนตรวจหา (*boxScore*) ซึ่งจะพิจารณาพื้นที่ทั้งหมดในกรอบสี่เหลี่ยม สังเกตได้จากผลการทดลองในกรณีการแยกสี่ด้วยภาพโมเดลสี่ $L \times a \times b$ ในตารางที่ 3.4 ซึ่งการตรวจหาให้ความแม่นยำเฉลี่ยที่ดีในช่วงอัตราส่วนการซ้อนทับสำหรับวัดผลในช่วงต่ำ โดยเฉพาะในช่วง 0.1 ถึง 0.3 การตรวจหาที่มีความแม่นยำเฉลี่ยสูงสุดในช่วง 94.75% ลดลงถึง 82.34% และความแม่นยำเฉลี่ยจะลดลงอย่างรวดเร็วเมื่อเพิ่มอัตราส่วนการซ้อนทับสำหรับวัดผลให้สูงขึ้น โดยเมื่อค่าอัตราส่วนการซ้อนทับสำหรับวัดผลช่วง 0.5 ถึง 0.7 วิธีตรวจหารอยกริดให้ความแม่นยำเฉลี่ยสูงสุดในช่วง 35.52% ลดลงถึง 4.29% ตามลำดับ



รูปที่ 4.1: ภาพขยายบริเวณพื้นที่รอยกรีดเก่า, รอยกรีดล่าสุด, และพื้นที่ที่ยังไม่ถูกกรีด ของภาพถ่ายมุมกว้าง โดยในรูป (ก) แสดงกรอบสี่เหลี่ยมอ้างอิงในรูปแบบเส้นประ, (ข) แสดงกรอบสี่เหลี่ยมผลลัพธ์จากการตรวจหาด้วยภาพโมเดลสี HSV กับภาพโมเดลสี $L^*a^*b^*$ พร้อมทั้งกรอบสี่เหลี่ยมอ้างอิง, และ (ค) แสดงบริเวณที่มีสีแตกต่างกัน โดยพื้นที่ในกรอบหมายเลข 1, 2, 3, และ 4 คือพื้นที่ที่ยังไม่ถูกกรีด, พื้นที่รอยกรีดเก่าที่ติดกับรอยกรีดล่าสุด, พื้นที่รอยกรีดเก่า, และพื้นที่รอยกรีดที่เก่ากว่า ตามลำดับ

ปัญหาดังกล่าวเกิดจากความไม่แน่นอนของสีหน้ายาง ตัวอย่างตามรูปที่ 4.1 โดยในรูปที่ 4.1 (ข) ผลลัพธ์กรอบสี่เหลี่ยมล้อมรอบรอยกรีดจากการตรวจหาทั้งภาพโมเดลสี HSV และภาพโมเดลสี $L^*a^*b^*$ มีการคลาดตำแหน่งจากกรอบสี่เหลี่ยมล้อมรอบอ้างอิง และมีบางส่วนอยู่ในพื้นที่ลำดับต้น ทั้งนี้เกิดจากบริเวณโดยรอบพื้นที่รอยกรีดเก่าถูกพิจารณาเป็นพื้นที่กลุ่มที่สีแดงในการจัดกลุ่มสี หรือเกิดจากความ เข้มหรือโทนของสีเกิดการเปลี่ยนแปลงไปจากโทนสีแดงของหน้ายางบริเวณใกล้รอยกรีดล่าสุด สามารถสังเกตเห็นได้ในหมายเลข 2, 3, และ 4 ในรูปที่ 4.1 (ค) ซึ่งหน้ายางที่เก่ามากกว่า (หมายเลข 4) จะมีสีที่คล้ำกว่าบริเวณใกล้รอยกรีดล่าสุด (หมายเลข 2) ค่อนข้างมากและสังเกตเห็นได้ชัดเจน เป็นผลให้ค่ากลางของกลุ่มสีโทนแดงเมื่อผ่านการจัดกลุ่ม มีโอกาสได้รับผลการเฉลี่ยค่าจากกลุ่มสีของพื้นที่บริเวณหมายเลข 4 มากกว่าพื้นที่หมายเลข 2

เมื่อหน้ายางที่ผ่านการกรีดมานานวันขึ้น ตัวอย่างตามรูปที่ 4.2 พื้นที่รอยกรีดเก่าของหน้ายางที่ผ่านการกรีดมานานกว่าจะมีขนาดใหญ่ ทำให้สีโทนแดงในพื้นที่หน้ายางจะมีความหลากหลายมากขึ้น และเป็นปัญหาต่อการตรวจหาตำแหน่งรอยกรีดด้วยวิธีการจัดกลุ่มสี หรือในกรณีพื้นที่รอยกรีดเก่ามีขนาดเล็กเนื่องจากเป็นหน้ายางเปิดกรีดใหม่ อาจมีสีโทนแดงที่ไม่ชัดเจนกว่าบริเวณโดยรอยกรีด การจัดกลุ่มสีโทนแดงอาจถูกรบกวนจากพื้นที่บริเวณอื่นในหน้ายาง นอกจากนี้ยังมีปัจจัยอื่นที่ส่งผลต่อความไม่แน่นอนของสีในพื้นที่หน้ายาง เช่น ระยะห่างระหว่างกล้องถ่ายภาพกับต้นยางพารา, ทิศทางหรือมุมมองกล้อง, ความสว่าง, หรือทิศทางของแสงที่ใช้ในการถ่ายภาพ เป็นต้น



รูปที่ 4.2: ภาพวาดแสดงลักษณะหน้าอย่างที่พื้นที่รอยกริดเก่ามีการขยายขนาดเมื่อผ่านการกริดมากขึ้น

4.1.2 การตรวจหารอยกริดและถ้วยรอนน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้างด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN

4.1.2.1 จุดเด่นของสถาปัตยกรรมเน็ตเวิร์กคอนโวลูชันสำหรับตรวจหาวัตถุแบบ Faster-RCNN

หากเปรียบเทียบเน็ตเวิร์ก Faster-RCNN ที่มีโครงสร้างแบบสองส่วนกับเน็ตเวิร์กตรวจหาวัตถุโครงสร้างแบบส่วนเดียว พบว่าเน็ตเวิร์ก Faster-RCNN มีข้อได้เปรียบที่สามารถตรวจหาวัตถุขนาดเล็กได้แม่นยำมากกว่าอ้างอิงจาก [64, 65] โดยยกตัวอย่างเปรียบเทียบกับเน็ตเวิร์ก YOLO [64] และเน็ตเวิร์ก SSD [50]) เนื่องจาก Faster-RCNN มีเน็ตเวิร์กย่อยที่ทำหน้าที่ค้นหาพื้นที่ย่อยโดยเฉพาะ (Region Proposal Network, RPN) ซึ่งในขั้นตอนการสอนจะอาศัยกรอบสี่เหลี่ยมเริ่มต้น (Anchor Box) โดยกรอบสี่เหลี่ยมดังกล่าวจะถูกกำหนดขนาดตามลักษณะของข้อมูล และผู้พัฒนาสามารถเลือกขนาดพีเจอร์เลเยอร์ของการตรวจหาให้เหมาะสมกับขนาดวัตถุที่สนใจได้ ซึ่งเป็นการกำหนดของเลเยอร์ส่วนปลายของ RPN ในตำแหน่งต่างๆ ของ MobileNetV2

โดยจากขนาดของรอยกริดและถ้วยรอนน้ำยางในภาพถ่ายมุมกว้างมีขนาดที่ค่อนข้างเล็กเมื่อเทียบกับขนาดของภาพ จึงทำให้เลือกทดลองเพิ่มเลเยอร์ส่วนปลายของ RPN ในเลเยอร์ภายใน MobileNetV2 ที่มีขนาดเทนเซอร์ที่ประมวลผลแตกต่างกัน ได้แก่ตำแหน่งบล็อก 3 (ประมวลผลเทนเซอร์ขนาด 28x28) และเลเยอร์บล็อก 10 (ประมวลผลเทนเซอร์ที่ขนาด 14x14) ซึ่งลักษณะต่างๆ เหล่านี้จะแตกต่างจากเน็ตเวิร์กแบบส่วนเดียว YOLO [64] และ SSD [50] ที่ไม่สามารถเลือกเลเยอร์ที่จำทำการตรวจหาวัตถุได้เหมือนกับการเลือกเลเยอร์ส่วนปลายให้เป็น RPN ของ Faster-RCNN แต่ใช้การแบ่งพื้นที่เป็นกริดในภาพ (Grid Cell) เพื่อสอนและตรวจหากรอบสี่เหลี่ยมล้อมรอบวัตถุ โดยไม่คำนึงถึงขนาดของกรอบสี่เหลี่ยมที่เหมาะสมกับชุดข้อมูล

แม้ว่า YOLO รุ่นถัดมา (YOLO9000 [65]) ได้เพิ่มความละเอียดของกริดให้มากขึ้นจากรุ่นก่อนหน้า (YOLO, [64]) เป็นผลให้การตรวจหาวัตถุขนาดเล็กมีความแม่นยำมากขึ้น ซึ่งในกรณีการตรวจหา

รอยกรีดและถ้วยร่อนน้ำยางในภาพถ่ายจากกล้องมุมกว้าง ข้อย่อยที่ 3.1.4 ได้แสดงตัวอย่างการทดลองเปรียบเทียบระหว่าง YOLO9000 กับ Faster-RCNN ที่ทำงานร่วมกับ MobileNetV2 ที่ผ่านการสอนมาก่อน ผลการทดลองตามรูปที่ 3.34 พบว่าเน็ตเวิร์ก YOLO9000 ที่ใช้เลเยอร์จาก MobileNetV2 ครบทุกเลเยอร์เช่นเดียวกับ Faster-RCNN มีความแม่นยำใกล้เคียงกับ Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN บล็อก 10 แต่ความแม่นยำน้อยกว่า Faster-RCNN ที่มี RPN บล็อก 3 อย่างเห็นได้ชัดในทุกอัตราส่วนการซ้อนทับสำหรับวัตถุ

เน็ตเวิร์ก YOLO9000 จะสิ้นสุดการทำคอนโวลูชันทันทีในเลเยอร์สำหรับระบุพื้นที่ย่อยบนเน็ตเวิร์ก MobileNetV2 ต่างจาก Faster-RCNN ที่เอาท์พุตจาก RPN จะถูกคอนโวลูชันต่อไปจนครบทุกเลเยอร์ของ MobileNetV2 เป็นผลให้พีเจอร์ระดับสูงจะไม่ถูกสกัดต่อไป ดังตัวอย่างการทดลองเปรียบเทียบ เมื่อกำหนดให้เน็ตเวิร์ก YOLO9000 ทำงานร่วมกับ MobileNetV2 ไม่ครบทุกเลเยอร์ เช่น การสร้าง YOLO9000 ด้วย MobileNetV2 ด้วยเลเยอร์ตั้งแต่บล็อก 1 จนถึงบล็อก 3 และเลเยอร์ตั้งแต่บล็อก 1 จนถึงบล็อก 10 กลับทำให้ความแม่นยำเฉลี่ยของการตรวจหารอยกรีดและถ้วยร่อนน้ำยางมีค่าสูงมากกว่า 95% ในทั้งสองกรณี ซึ่งอาจเกิดจากพีเจอร์ที่เป็นประโยชน์ต่อการตรวจหารอยกรีดและถ้วยร่อนน้ำยางถูกสกัดได้หมดในเลเยอร์ส่วนดังกล่าว การใช้งานเลเยอร์จาก MobileNetV2 ทั้งหมดอาจทำให้ข้อมูลพีเจอร์ที่เป็นประโยชน์ดังกล่าวถูกกลดทอนลงในเลเยอร์ระดับสูงกว่าชั้นต่อๆ มา เป็นผลให้ความสามารถในการรู้จำลดลง ซึ่งเป็นประเด็นการศึกษาการพัฒนาการตรวจหารอยกรีดและถ้วยร่อนน้ำยางด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันต่อไป

4.1.2.2 ตำแหน่งของ RPN ภายใน MobileNetV2

การสร้าง Faster-RCNN ร่วมกับ MobileNetV2 โดยให้มีเลเยอร์ส่วนปลายของ RPN ในตำแหน่งบล็อก 3 ของ MobileNetV2 ให้ค่าความแม่นยำเฉลี่ยในการตรวจหารอยกรีดและถ้วยร่อนน้ำยางที่ 80% ซึ่งสูงกว่า Faster-RCNN ที่มีเลเยอร์ส่วนปลาย RPN ในตำแหน่งบล็อก 10 ที่ให้ค่าความแม่นยำเฉลี่ยที่ 50% ที่อัตราส่วนการซ้อนทับสำหรับวัตถุที่ 0.5 ทั้งนี้เป็นผลจากการกำหนด RPN ในเลเยอร์บล็อก 3 จะทำให้ขนาดของรอยกรีดและถ้วยร่อนน้ำยางที่ปรากฏในภาพ (เทนเซอร์) ณ เลเยอร์ตำแหน่งบล็อก 3 ยังคงเป็นขนาดที่ Faster-RCNN สามารถรู้จำได้ ต่างจากการกำหนดเลเยอร์ส่วนปลายในตำแหน่งบล็อก 10 ซึ่งขนาดของรอยกรีดและถ้วยร่อนน้ำยางในภาพ ณ เลเยอร์ตำแหน่งดังกล่าวมีขนาดเล็กเกินไปจนไม่เพียงพอต่อการรู้จำด้วยเน็ตเวิร์กส่วนที่เหลือที่อยู่ถัดจาก RPN ได้

4.1.2.3 การแบ่งกลุ่มข้อมูลทดลองตามช่องสัญญาณภาพอินพุต

โครงข่ายประสาทเชิงลึกแบบคอนโวลูชันที่ผ่านการสอนมาก่อนโดยส่วนมาก ต่างเรียนรู้ด้วยข้อมูลรูปภาพสีแบบ RGB รวมไปถึง MobileNetV2 มีขนาดเท่ากับ 3 ช่องสัญญาณ ทำให้เน็ตเวิร์กดังกล่าวจะเรียนรู้พีเจอร์ต่างๆ ที่เป็นประโยชน์ต่อการรู้จำวัตถุ บนพื้นฐานข้อมูลรูปภาพสี RGB อาทิ ขอบภาพ, ค่าสี, การรวมกลุ่มของขอบภาพ, หรือรูปร่างของวัตถุในภาพซึ่งเป็นพีเจอร์ในระดับที่สูงขึ้นไป เป็นต้น

ภาพถ่ายมุมกว้างประกอบไปด้วยภาพถ่ายภาพสีและภาพความลึก ซึ่งมีจำนวน 4 ช่องสัญญาณ เพื่อใช้งานเน็ตเวิร์กขนาด 3 ช่องสัญญาณกับภาพขนาด 4 ช่องสัญญาณ จึงจัดกลุ่มข้อมูลภาพขนาด 4 ช่องสัญญาณให้เหลือ 3 ช่องสัญญาณตามขนาดอินพุตของเน็ตเวิร์ก ได้แก่ ภาพถ่ายสีตามปกติเพียงอย่างเดียว,

ภาพความลึกเพียงอย่างเดียว, และภาพถ่ายสีเมื่อคุณกับภาพถ่ายความลึก โดยผลการทดลองแสดงให้เห็นว่า ความแม่นยำในการตรวจหารอยกรีดและถ้วยรองน้ำยางของกลุ่มภาพถ่ายสี RGB เมื่อถูกกรองด้วยภาพความลึก (Depth Segmentation) มีค่าสูงกว่าการใช้ภาพถ่ายสี RGB ตามปกติ ทั้งนี้เกิดจากการคูณด้วยภาพความลึก ทำให้พื้นที่พื้นหลังที่ไม่ใช่ลำต้นถูกขจัดออกจากภาพไปก่อนป้อนเข้าสู่เน็ตเวิร์ก ทำให้พื้นที่พื้นหลังเหลือเป็นบริเวณพิกเซลที่มีค่าเท่ากับศูนย์ในภาพ ทำให้เน็ตเวิร์กไม่ต้องเรียนรู้ลักษณะของพื้นหลัง ซึ่งมีความหลากหลายหรือความแปรปรวนมากกว่าพื้นที่ลำต้น หรือกล่าวคือทำให้เน็ตเวิร์กเรียนรู้ลักษณะของพื้นที่ลำต้นได้ง่ายขึ้น วิธีการดังกล่าวเป็นการนำภาพถ่ายความลึกมาใช้ประโยชน์เพื่อการตรวจหารอยกรีดบนหน้ายางได้

ในขณะเดียวกันการตรวจหาด้วยภาพถ่ายความลึกเพียงอย่างเดียว ไม่สามารถตรวจหาตำแหน่งรอยกรีดได้ เนื่องจากลักษณะของรอยกรีดไม่ปรากฏในภาพถ่ายความลึกมุมกว้าง แต่การตรวจหาถ้วยรองน้ำยางด้วยภาพถ่ายความลึกเพียงอย่างเดียว มีความแม่นยำที่ค่อนข้างสูง ทั้ง Faster-RCNN ที่มีเลเยอร์ส่วนปลายบล็อก 3 และบล็อก 10 มีค่าประมาณเท่ากับ 97% และ 76% ตามลำดับ ดังนั้นการตรวจหาถ้วยรองน้ำยางในภาพถ่ายมุมกว้าง สามารถทำได้ในเฉพาะภาพถ่ายความลึก เมื่อเปรียบเทียบกับกรณีการใช้ภาพถ่ายสี การตรวจหาจึงไม่ต้องการแสงสว่าง หรือสามารถหลีกเลี่ยงแสงจากสภาพแวดล้อมโดยรอบได้

4.1.2.4 การแบ่งข้อมูลเพื่อสอนและทดสอบเน็ตเวิร์กของแต่ละกลุ่มทดลอง

การแบ่งข้อมูลเพื่อใช้สอนและทดสอบเน็ตเวิร์กในกรณีที่มีข้อมูลมีปริมาณมากมักนิยมใช้วิธีแบ่งข้อมูลออกเป็น 2 กลุ่ม (Hold-out Method) ในปริมาณที่เป็นอัตราส่วน เช่น 70:30, 80:20 เป็นต้น สำหรับสอนและทดสอบเน็ตเวิร์กตามลำดับ การแบ่งปริมาณข้อมูลในสัดส่วนดังกล่าวกับชุดข้อมูลภาพถ่ายรอยกรีดมุมกว้าง ซึ่งมีจำนวน 500 ชุดภาพ อาจทำให้ข้อมูลสำหรับสอนหรือทดสอบมีปริมาณน้อยเกินไป ส่งผลต่อความน่าเชื่อถือในการอนุมานผลการตรวจหาด้วยวิธีตรวจหาที่พัฒนาขึ้นไปยังกลุ่มข้อมูลอื่น

ดังนั้นจึงเลือกวิธีการแบ่งข้อมูลทดสอบ K กลุ่ม (K-Fold Cross Validation) ซึ่งเป็นวิธีมาตรฐานเพื่อใช้แบ่งส่วนข้อมูลที่มีปริมาณน้อยสำหรับสอนและทดสอบเน็ตเวิร์ก เป็นวิธีที่เพิ่มความน่าเชื่อถือในการทดลองด้วยข้อมูลภาพถ่ายจากกล้องถ่ายภาพมุมกว้างที่มีปริมาณเพียง 500 ภาพเมื่อเปรียบเทียบกับวิธีการแบ่งข้อมูลแบบ 2 กลุ่มข้างต้น โดยการทดลองได้เลือกค่า $K=5$ ซึ่งจะแบ่งข้อมูลชุดภาพในแต่ละกลุ่มทดลองออกเป็น 5 ส่วน โดยข้อมูลจำนวน 4 ส่วน จะถูกใช้สอนเน็ตเวิร์ก และข้อมูล 1 ส่วนที่เหลือใช้สำหรับทดสอบหรือวัดผลการตรวจหา ทำให้แต่ละกลุ่มทดลองสามารถสอนเน็ตเวิร์กทั้งหมด 5 โมเดล ด้วยข้อมูลที่ไม่ซ้ำกัน จากนั้นเฉลี่ยความแม่นยำของแต่ละเน็ตเวิร์กที่ได้เพื่อเป็นค่าความแม่นยำเฉลี่ยของกลุ่มทดลองนั้น

4.1.2.5 การเพิ่มข้อมูล (Data Augmentation)

ภาพถ่ายมุมกว้างมีจำนวน 500 ภาพ ซึ่งถือว่ามีความที่ค่อนข้างน้อยต่อการสอนเน็ตเวิร์กแบบคอนโวลูชันเพื่อตรวจหาวัตถุ จึงเหมาะที่จะเพิ่มปริมาณข้อมูล (Augmentation) ซึ่งเทคนิคที่ใช้อย่างแพร่หลายคือการประมวลผลภาพเพื่อสร้างภาพใหม่ ที่แตกต่างจากภาพต้นฉบับ โดยในการทดลองได้เลือกการประมวลผลภาพต้นฉบับให้ต่างออกไปจากเดิมในสองรูปแบบได้แก่ การสุมหมุนภาพ และการสุมเพิ่มหรือลดความสว่างโดยรวมของภาพ

การเลือกแปลงภาพให้ต่างไปจากเดิมในรูปแบบดังกล่าว มีแนวโน้มที่จะเกิดขึ้นได้จากการทำงานระบบมองเห็นของหุ่นยนต์ภายในสวนยางพารา เช่น ต้นยางพารามีระยะห่างจากหุ่นยนต์มากหรือลดลงต่างกันออกไปในแต่ละต้น ส่งผลให้แสงสว่างอาจมากขึ้นหรือลดลง หรือในกรณีที่ต้นยางพาราเอนไปจากเดิมมากกว่าปกติ การสุมหมุนภาพให้ต่างไปจากเดิมเล็กน้อยจึงเป็นการจำลองเหตุการณ์ในลักษณะดังกล่าว

ทั้งนี้หากประมวลผลภาพเพื่อเพิ่มปริมาณของข้อมูลในลักษณะอื่น เช่น การสุมลบบางพื้นที่ในภาพ, การยืดหรือหดภาพ, หรือการสุมเพิ่มสัญญาณรบกวนในภาพ (Noise) สามารถทำได้เช่นกันแต่อาจทำให้ค่าความแม่นยำเฉลี่ยที่ได้ในแต่ละกลุ่มข้อมูลทดลองเปลี่ยนไปจากเดิม และภาพจากการประมวลผลดังกล่าวไม่ได้มีแนวโน้มที่จะเกิดขึ้น หรือมีโอกาสเกิดขึ้นได้น้อยในสภาพการทำงานของหุ่นยนต์ในสวนยางพารา

4.1.2.6 ขนาดภาพอินพุต

เน็ตเวิร์ก Faster-RCNN รองรับรูปภาพอินพุตในหลายขนาด อ้างอิงจากงานต้นฉบับ Faster-RCNN [66] ที่ทำงานบนเน็ตเวิร์ก VGG-16 มีภาพอินพุตขนาด 600x600 พิกเซล ซึ่งภาพอินพุตที่มีขนาดใหญ่กว่าจะให้ผลการตรวจหาของวัตถุขนาดเล็กได้แม่นยำมากขึ้น โดยจากข้อจำกัดของหน่วยความจำและทรัพยากรเครื่องคอมพิวเตอร์ที่ใช้ทำการทดลอง อีกทั้งเพื่อให้สอดคล้องกับระบบคอมพิวเตอร์ที่สามารถติดตั้งให้กับหุ่นยนต์ที่อาจมีศักยภาพในการประมวลผลไม่สูงนัก วิทยานิพนธ์ฉบับนี้จึงเลือกใช้เน็ตเวิร์ก MobileNetV2 ที่มีขนาดภาพอินพุตเท่ากับ 224x224 พิกเซล ซึ่งผลการทดลองแสดงให้เห็นว่าขนาดดังกล่าวเพียงพอต่อการตรวจหากรอบของรอยกรีดและถ้วยรองน้ำยางในภาพถ่ายมุมกว้าง ซึ่งได้ความแม่นยำเฉลี่ย (mAP) สูงสุดถึง 89.28%

ทั้งนี้สามารถเลือกขนาดของภาพอินพุตที่ใหญ่มากขึ้นเพื่อเพิ่มความแม่นยำในการตรวจหารอยกรีดและถ้วยรองน้ำยาง แต่ต้องแลกกับการประมวลผลที่มากขึ้น ซึ่งจะทำให้ระยะเวลาประมวลผลต่อภาพนานมากขึ้น

4.2 การตรวจหาแนวรอยกริดในภาพถ่ายระยะใกล้

4.2.1 การกำหนดแนวรอยกริดอ้างอิงในภาพ

การกำหนดแนวรอยกริดอ้างอิงในภาพ มีขึ้นเพื่อสร้างชุดข้อมูลสำหรับการพัฒนาวิธีตรวจหาแนวรอยกริดในภาพ ประกอบไปด้วย วิธีการกำหนดตำแหน่งแนวรอยกริดอ้างอิงด้วยมือ และวิธีการปรับตำแหน่งแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ

4.2.1.1 การกำหนดแนวรอยกริดอ้างอิงด้วยมือ

การประมาณแนวรอยกริดด้วยเส้นทแยงมุมที่เชื่อมมุมซ้ายบนและมุมขวาล่างของกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริด ทำให้พิกเซลมุมซ้ายบนของกรอบอยู่ที่ตำแหน่งซ้ายสุดของรอยกริดที่ปรากฏในภาพ (จุดเริ่มกริด) และทำให้พิกเซลมุมขวาล่างของกรอบ อยู่ที่ตำแหน่งขวาสุดของรอยกริดที่ปรากฏในภาพ (จุดสิ้นสุดการกริด) เช่นเดียวกับในกรณีการกำหนดตำแหน่งรอยกริดในภาพถ่ายจากกล้องถ่ายภาพมุมกว้าง วิธีการดังกล่าวมีความละเอียดที่ไม่เพียงพอต่อการใช้เป็นแนวรอยกริดอ้างอิงสำหรับการตรวจหาในภาพถ่ายระยะใกล้ เนื่องจากแนวเส้นทแยงมุมจะคลาดเคลื่อนไปจากตำแหน่งแนวรอยกริดจริงที่ปรากฏในภาพ ความคลาดเคลื่อนดังกล่าวเกิดจากเส้นทแยงมุมของกรอบที่เป็นเส้นตรงซึ่งไม่มีการโค้งตามลักษณะโค้งของแนวรอยกริด เพื่อแก้ปัญหาดังกล่าวจึงนำเสนอการใช้เส้นโค้งหลายเหลี่ยม (Polygonal Curve) แทนแนวรอยกริดในภาพ ซึ่งเส้นโค้งหลายเหลี่ยมสามารถปรับความเอียงไปตามแนวรอยกริดได้ดีกว่าเส้นทแยงมุมที่เป็นเพียงเส้นตรง

โปรแกรมกำหนดตำแหน่งในภาพโดยทั่วไปไม่มีเครื่องมือสำหรับวาดเส้นโค้งหลายเหลี่ยมโดยตรง และไม่มีรูปแบบการเก็บข้อมูลเส้นโค้งหลายเหลี่ยมให้กับภาพ เหมือนกับกรอบสี่เหลี่ยมที่สามารถแทนได้ด้วยชุดตัวเลขกำกับ หรือรูปปิดที่ใช้ภาพขาวดำในการเก็บข้อมูลตำแหน่ง ดังนั้นจึงออกแบบให้ผู้กำหนดตำแหน่งใช้งานโปรแกรมกำหนดตำแหน่ง¹ วาดกรอบสี่เหลี่ยมล้อมรอบรอยกริด และวาดรูปหลายเหลี่ยมที่มีด้านหรือขอบ ทับตามแนวรอยกริด แล้วใช้วิธีหาเส้นขอบล่างกับพื้นที่รูปหลายเหลี่ยมเพื่อแทนแนวรอยกริดในภาพ โดยรูปแบบการเก็บข้อมูลจะเป็นตำแหน่งพิกเซล เรียงลำดับจากตำแหน่งพิกเซลซ้ายสุดมือไปยังพิกเซลขวามือสุด ซึ่งลักษณะของเส้นขอบล่างได้รับรองว่า มีพิกเซลในทุกคอลัมน์ที่อยู่ติดกันในระหว่างพิกเซลที่อยู่ทางซ้ายสุดกับพิกเซลขวาสุดของแนวรอยกริดในภาพ

ความแม่นยำของการกำหนดแนวรอยกริดอ้างอิงด้วยวิธีข้างต้น จะขึ้นอยู่กับความสามารถของผู้กำหนดรอยกริดที่จะสังเกตเห็นแนวรอยกริดในภาพ และสามารถวางจุดมุมของรูปหลายเหลี่ยมได้เพียงพอและแม่นยำเพียงใด ซึ่งหากผู้กำหนดตำแหน่งเพิ่มจุดมุมของรูปหลายเหลี่ยมให้มากขึ้น อาจทำให้แนวรอยกริดอ้างอิงที่ดีตรงกับแนวรอยกริดจริงมากขึ้น แต่ในขณะเดียวกันแนวรอยกริดอาจเกิดความคลาดเคลื่อนหรือสูญเสียคุณสมบัติความเรียบมากขึ้นเช่นกัน

¹Microsoft VoTT (Visual Object Tagging Tool)

4.2.1.2 การปรับตำแหน่งแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ

การทดลองปรับตำแหน่งแนวรอยกริดอ้างอิงที่กำหนดด้วยมือเกิดขึ้นจากสมมติฐานที่ว่า แนวรอยกริดอ้างอิงในระดับพิกเซลด้วยที่กำหนดด้วยมือ ที่สร้างจากการวาดกรอบสี่เหลี่ยมล้อมรอบและรูปหลายเหลี่ยมในภาพถ่ายสีตามวิธีในข้อ 2.4.2.1 อาจมีความแม่นยำหรือความถูกต้องที่ไม่เพียงพอเพื่อใช้เป็นแนวรอยกริดอ้างอิง

การปรับตำแหน่งแนวรอยกริดอ้างอิงที่กำหนดด้วยมือจะใช้ลักษณะขอบซึ่งเป็นขอบประเภทขั้นบันได (Step Edge) [60] ตามแนวรอยกริดที่ปรากฏในภาพถ่ายความลึก ลักษณะขอบดังกล่าวเป็นผลมาจากการวางกล่องให้มีแกน Y ขนานเส้นสัมผัสกับผิวหน้ายางในแนวตั้ง ดังนั้นอัลกอริทึมปรับตำแหน่งแนวรอยกริดที่กำหนดด้วยมือจะปรับตำแหน่งแนวรอยกริดที่กำหนดด้วยมือ ให้ตรงตำแหน่งขอบตามแนวรอยกริดในภาพถ่ายความลึกมากขึ้น ด้วยอัลกอริทึมปรับตำแหน่งที่ออกแบบขึ้น ซึ่งเป็นประเภทวนรอบทำซ้ำเพื่อผลลัพธ์ที่ดีที่สุด (Exhaustive Search) อัลกอริทึมอาศัยหลักการหาขอบตามแนวรอยกริดในภาพถ่ายความลึกด้วยการกรองความถี่สูงในภาพด้วยฟิลเตอร์โซเบลเฉพาะแนวแกนตั้ง (Y-Sobel Kernel) และการเบลอภาพขนาดของเกรเดียนต์ด้วยฟิลเตอร์เบลอภาพแบบเกาส์ที่มีค่าพารามิเตอร์ σ แทนค่าการกระจายของน้ำหนักรายในเคอร์เนล ค่าดังกล่าวส่งผลต่ออัตราการเบลอของภาพ และส่งผลต่อตำแหน่งจุดยอดของสัญญาณภาพขนาดของเกรเดียนต์ตามแนวแกนตั้งของภาพความลึกตามแนวรอยกริด

การปรับตำแหน่งแนวรอยกริดอ้างอิงของวิธีปรับตำแหน่งตามข้อย่อย 2.4.2.2 จะพยายามย้ายตำแหน่งพิกเซลของแนวรอยกริดที่กำหนดด้วยมือ ไปยังจุดยอดภายในภาพขนาดเกรเดียนต์ของความลึกตามแนวแกนตั้งที่อยู่ใกล้เคียงและอยู่ภายในกรอบค้นหา R_{sr} ที่ครอบคลุมแนวรอยกริดที่กำหนดด้วยมือเป็นระยะ sr พิกเซล ตามแนวแกนตั้งของภาพ อัลกอริทึมปรับตำแหน่งจะวนรอบเบลอภาพขนาดของเกรเดียนต์ด้วยฟิลเตอร์เบลอภาพแบบเกาส์ที่มีค่า σ เพิ่มขึ้นเรื่อยๆ จนภาพเบลอมากพอที่จะทำให้ขนาดของเกรเดียนต์ตามแนวแกนตั้งภายในกรอบค้นหาแต่ละคอลัมน์ เป็นสัญญาณจุดยอดเดี่ยวในมุมมองของคอลัมน์ ซึ่งเป็นตำแหน่งที่เหมาะสมของแนวรอยกริดตามสมมติฐาน แนวรอยกริดในตำแหน่งแห่งสัญญาณที่เป็นจุดยอดในภาพความถี่สูงของขอบความลึกจะถูกปรับรูปร่างโดยรวมให้เรียบและเอียงทำมุม 30 ถึง 45 องศากับแกนนอนของภาพอ้างอิงตามรูปร่างแนวรอยกริดตามระบบกริดลงแบบเวียนขวา ดังนั้นวิธีปรับตำแหน่งแนวรอยกริดอ้างอิงที่กำหนดด้วยมือ มีตัวแปรที่จะต้องควบคุมคือ sr เพื่อกำหนดกรอบการค้นหา และคอยควบคุมค่า σ ของฟิลเตอร์เบลอภาพแบบเกาส์ในทางอ้อม ซึ่งในขั้นตอนสุดท้ายของอัลกอริทึมปรับตำแหน่งแนวรอยกริดอ้างอิง อัลกอริทึมจะวนหาแนวรอยกริดที่มีคุณภาพหรือมีลักษณะที่ดีที่สุดภายในกรอบการค้นหา sr ที่มีค่าตั้งแต่ 1 พิกเซล ดังนั้นผลของการปรับตำแหน่งจะขึ้นอยู่กับลักษณะการวางจุดมุมของรูปหลายเหลี่ยมเพื่อแทนแนวรอยกริดอ้างอิง จากวิธีกำหนดแนวรอยกริดอ้างอิงด้วยมือ

ผลการทดลองได้ยืนยันสมมติฐานว่าการกำหนดแนวรอยกริดอ้างอิงด้วยมือมีความคลาดเคลื่อน โดยวิธีปรับตำแหน่งสามารถปรับตำแหน่งแนวรอยกริดอ้างอิงได้สำเร็จสูงสุดถึง 91% เมื่อใช้การค้นหาแนวรอยกริดที่ดีที่สุดในการค้นหาที่กำหนด ในขณะที่หากพิจารณาขนาดกรอบการค้นหาแบบค่าเดียว (sr แบบค่าเดียว) พบว่าแนวรอยกริดอ้างอิงที่กำหนดด้วยมือถูกปรับตำแหน่งได้สำเร็จในการค้นหาที่ขนาด 2 ถึง 12 พิกเซล มากที่สุดประมาณ 83% จากจำนวนแนวรอยกริดอ้างอิงทั้งหมดจำนวน 1,405 แนวรอยกริด โดยในขณะที่หากใช้ปรับตำแหน่งด้วยกรอบการค้นหาที่มีขนาดใหญ่มากขึ้น เช่นกรอบการค้นหาขนาด 20 พิก-

เซล วิธีปรับตำแหน่งสามารถปรับตำแหน่งแนวรอยกริดอ้างอิงได้สำเร็จลดลง เหลือเพียง 70%, หรือที่ขนาด 40 พิกเซล เหลือเพียง 37% จากจำนวนแนวรอยกริดอ้างอิงทั้งหมด

ผลการทดลองที่ได้ ค่าขนาดของกรอบค้นหาความสัมพันธ์กับขนาดของภาพที่ประมวลผลซึ่งในการทดลองได้ประมวลผลบนภาพขนาดกว้าง 1280 สูง 720 พิกเซลตลอดทั้งการทดลอง วิธีการปรับตำแหน่งแนวรอยกริดอ้างอิงที่ได้พัฒนาขึ้นยังสามารถนำไปปรับใช้กับภาพที่มีขนาดอื่นได้เช่นกัน

4.2.2 การตรวจหาตำแหน่งแนวรอยกริดในภาพถ่ายระยะใกล้

4.2.2.1 วิธีตรวจหาแนวรอยกริด

วิธีการตรวจหาแนวรอยกริดในภาพระยะใกล้ประกอบไปด้วย 3 ขั้นตอนย่อย โดยในขั้นแรกอัลกอริทึม 2.4.5 จะคำนวณหาพื้นที่เงาตามแนวรอยกริด ด้วยการหาความแตกต่างระหว่างภาพถ่ายสีที่ 1 และ 2 ด้วยการลบค่าพิกเซลในตำแหน่งเดียวกันระหว่างภาพโทนเทา (Grayscale) ของภาพถ่ายสีทั้งสอง ดังนั้นภาพถ่ายทั้งสองจะต้องถูกถ่ายจากกล้องที่มีตำแหน่งคงที่ เพื่อไม่ให้แต่ละพิกเซลในภาพทั้งสองคลาดตำแหน่งกัน อีกทั้งภาพถ่ายสีที่ 2 จะต้องเป็นภาพที่มีเงาตามแนวรอยกริดเท่านั้น ภาพผลลัพธ์จากขั้นตอนนี้อาจมีพื้นที่เงาที่ไม่ใช่เงาตามแนวรอยกริดหลงเหลืออยู่ ซึ่งอาจบดบังการหาเส้นขอบล่างด้วยวิธีสแกนตามแนวคอลัมน์ พื้นที่ที่เป็นปัญหานี้จะถูกนำออกจากภาพในขั้นตอนถัดไป

ในขั้นตอนสุดท้าย อัลกอริทึม 2.4.7 จะคำนวณหาแนวรอยกริดจากเส้นขอบล่างของภาพด้วยการทำคอลัมน์สแกน จากนั้นจะพิจารณาส่วนของเส้นขอบล่างที่ยาวที่สุดและมีลักษณะที่เรียบเป็นแนวรอยกริดผลลัพธ์ โดยมีพารามิเตอร์ T_{up} , T_{down} ทำหน้าที่ควบคุมการแกว่งของค่าอนุพันธ์ซึ่งใช้พิจารณาแทนความเรียบของแนวรอยกริด การทดลองได้ใช้วิธีการกำหนดค่าของตัวแปรทั้งสองแบบค่าเดียวตลอดทั้งชุดข้อมูล (Global Threshold) ซึ่งเป็นวิธีที่ไม่มีประสิทธิภาพมากนัก เพราะค่าดังกล่าวไม่ได้อาจทำให้แนวรอยกริดที่ตรวจหาได้ไม่ครบสมบูรณ์หากเปลี่ยนไปใช้ค่าอื่นที่ใกล้เคียง เนื่องจากการเลือกค่า T_{up} , T_{down} ไม่ได้พิจารณาลักษณะความเอียงของแนวรอยกริดที่ปรากฏในภาพนั้นๆ

วิธีการตรวจหา รอยกริดที่เป็นขั้นตอนย่อยทำงานตามลำดับต่อเนื่องกัน ดังนั้นหากเกิดปัญหาขึ้นในขั้นตอนใดขั้นตอนหนึ่ง วิธีการตรวจหาทั้งหมดก็จะไม่สามารถตรวจหาแนวรอยกริดได้ อีกทั้งวิธีการตรวจหาแนวรอยกริดไม่ได้รับรองว่าจุดปลายทั้งสองของแนวรอยกริดที่ตรวจหาได้ในภาพ เป็นจุดเริ่มกริดและจุดสิ้นสุดการกริดของแนวรอยกริดจริง

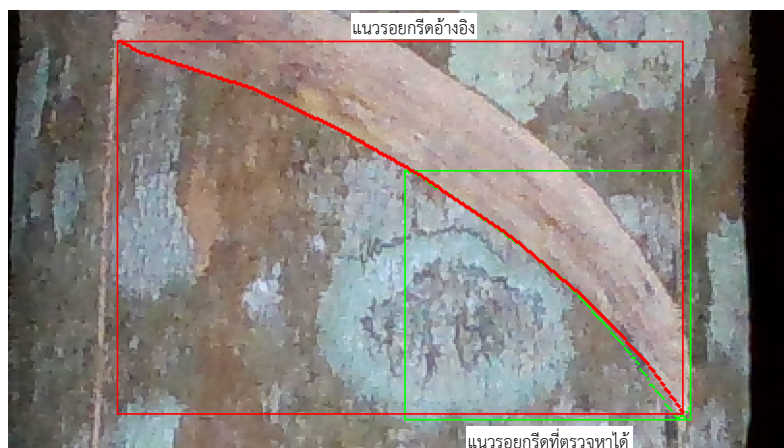
4.2.2.2 ผลการตรวจหาแนวรอยกริดในภาพ

การวัดผลการตรวจหาแนวรอยกริดประกอบด้วย 2 ขั้นตอนที่ต้องเนื่องกัน โดยในขั้นแรกจะวัดการตรวจพบ ด้วยอัตราส่วนการซ้อนทับของกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริดที่อัตราส่วนการซ้อนทับสำหรับวัดผลเท่ากับ 0.5 ซึ่งเป็นค่าที่นิยมใช้สำหรับวัดผลการตรวจหาวัตถุในภาพ [16] จากนั้นวัดระยะห่างของแนวรอยกริดที่ตรวจพบ กับแนวรอยกริดอ้างอิงด้วยระยะทางแบบเฮาส์ดอร์ฟฟ์ วิธีการวัดผลข้างต้นแตกต่างจากวิธีการวัดผลการตรวจหาวัตถุโดยทั่วไปที่ใช้อัตราส่วนการซ้อนทับระหว่างกรอบสี่เหลี่ยมเพียงอย่างเดียว เนื่องจากลักษณะของกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริดไม่ได้อธิบายลักษณะของแนวรอยกริดภายใน ซึ่งได้ประเด็นดังกล่าวได้อธิบายในบทที่ 2 ซ้อย่อย 2.5.2

ค่า T_{up} และ T_{down} จะส่งผลต่อความแม่นยำทั้งการวางตำแหน่งกรอบสี่เหลี่ยม และระยะห่างระหว่างแนวรอยกริดอ้างอิงกับแนวรอยกริดที่ตรวจพบ ยกตัวอย่างผลการตรวจหาแนวรอยกริดในรูปที่ 4.3 โดยรูปที่ 4.3 (ก) , 4.3 (ข) , และ 4.3 (ค) แสดงกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริดและแนวรอยกริดภายในกรอบ เมื่อใช้ค่า (T_{up}, T_{down}) เท่ากับ (2, 5), (2, 3), และ (2, 2) ตามลำดับ โดยพบว่า T_{up} และ T_{down} ต่างกันจะส่งผลต่ออัตราส่วนการซ้อนทับของกรอบสี่เหลี่ยมที่ล้อมรอบแนวรอยกริด และยังทำให้ความสมบูรณ์ของแนวรอยกริดที่ตรวจหาได้เปลี่ยนไปเช่นกัน ถึงแม้ว่าแนวรอยกริดที่ตรวจหาได้ในทั้งสามภาพตัวอย่างจะซ้อนทับกับแนวรอยกริดอ้างอิงบางส่วนได้ดี แต่หาก T_{up} และ T_{down} มีค่าที่ไม่เหมาะสม จะทำให้ผลการตรวจหาถูกระบุเป็นการตรวจหาที่ตรวจไม่พบ เนื่องจากอัตราส่วนการซ้อนทับของกรอบสี่เหลี่ยมของแนวรอยกริดที่ตรวจหาได้ กับกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริดอ้างอิงมีค่าน้อยกว่า 0.5

ผลการทดลองที่กำหนดค่าแปรทั้งสองแบบค่าเดียวตลอดทั้งชุดข้อมูล แสดงให้เห็นว่าค่าของตัวแปรทั้งสองในช่วง 3 ถึง 10 พิกเซล ทำให้การตรวจหาแนวรอยกริดมีความแม่นยำเฉลี่ยสูงสุดที่ประมาณ 89.9% ในรูปที่ 3.40 ในขณะที่ระยะห่างแบบเฮาส์ดอร์ฟฟ์เฉลี่ยระหว่างแนวรอยกริดที่ตรวจพบกับแนวรอยกริดอ้างอิงมีค่าเฉลี่ยที่ 13 พิกเซล แสดงในกรอบเส้นประในรูปที่ 4.4 ซึ่งเป็นค่าที่น้อยกว่าการตรวจหาเมื่อใช้ค่า T_{up} และ T_{down} อื่นในบริเวณใกล้เคียง

ค่า T_{up} และ T_{down} จากการทดลอง เป็นค่าที่สัมพันธ์กับขนาดของภาพและขนาดของแนวรอยกริดที่ปรากฏในภาพ ซึ่งหากถ่ายภาพระยะใกล้โดยให้กล้องเข้าใกล้แนวรอยกริดมากขึ้นหรือลดลง อาจทำให้ค่า T_{up} และ T_{down} ที่สามารถตรวจหาแนวรอยกริดได้แม่นยำที่สุดเปลี่ยนไปเช่นกัน ทั้งนี้เป็นผลจากมุมมองของกล้องระยะใกล้กับรอยกริดบนหน้ายาง ดังนั้นค่าจากการทดลองในช่วง 3 ถึง 10 พิกเซล ดังกล่าวจึงเป็นผลจากระยะการวางกล้องระยะใกล้ที่ 20 ถึง 40 เซนติเมตรจากหน้ายาง และขนาดของภาพถ่ายที่ 1280x720 พิกเซล ดังนั้นวิธีการตรวจหาแนวรอยกริดในภาพระยะใกล้สามารถนำไปปรับใช้กับภาพถ่าย และแนวรอยกริดที่ปรากฏในภาพได้หลายขนาด โดยการหาค่า T_{up} และ T_{down} ที่เหมาะสมกับแนวรอยกริดในแต่ละภาพ

(ก) $T_{up} = 2, T_{down} = 5$ (ข) $T_{up} = 2, T_{down} = 3$ (ค) $T_{up} = 2, T_{down} = 2$

รูปที่ 4.3: ภาพขยายของตัวอย่างผลการตรวจหาแนวรอยกรีดในภาพถ่ายระยะใกล้เมื่อใช้ค่า T_{up} และ T_{down} ต่างกัน

ระยะทางแบบเฮาส์คอฟฟ์เฉลี่ย (พิกเซล)

| | | | | | | | | | | | | |
|----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 13 | 15.87 | 14.68 | 14.35 | 14.34 | 14.24 | 14.29 | 14.49 | 14.53 | 14.57 | 14.65 | 14.84 | 14.90 |
| 12 | 15.59 | 14.40 | 14.05 | 14.07 | 13.94 | 13.94 | 14.14 | 14.23 | 14.28 | 14.36 | 14.55 | 14.61 |
| 11 | 15.42 | 14.20 | 13.85 | 13.86 | 13.73 | 13.73 | 13.87 | 13.97 | 14.02 | 14.09 | 14.27 | 14.34 |
| 10 | 15.15 | 13.91 | 13.54 | 13.56 | 13.44 | 13.45 | 13.59 | 13.63 | 13.68 | 13.76 | 13.94 | 14.00 |
| 9 | 14.88 | 13.70 | 13.33 | 13.30 | 13.20 | 13.21 | 13.35 | 13.39 | 13.44 | 13.52 | 13.71 | 13.78 |
| 8 | 14.73 | 13.52 | 13.16 | 13.14 | 13.08 | 13.12 | 13.23 | 13.30 | 13.35 | 13.43 | 13.62 | 13.69 |
| 7 | 14.56 | 13.34 | 12.95 | 12.95 | 12.88 | 12.91 | 13.07 | 13.14 | 13.19 | 13.27 | 13.46 | 13.53 |
| 6 | 14.52 | 13.41 | 13.13 | 13.16 | 13.12 | 13.15 | 13.25 | 13.32 | 13.37 | 13.46 | 13.65 | 13.72 |
| 5 | 14.63 | 13.82 | 13.56 | 13.63 | 13.58 | 13.60 | 13.71 | 13.78 | 13.84 | 13.94 | 14.13 | 14.19 |
| 4 | 16.18 | 15.59 | 15.40 | 15.46 | 15.45 | 15.46 | 15.58 | 15.62 | 15.67 | 15.78 | 15.99 | 16.03 |
| 3 | 20.35 | 20.11 | 19.99 | 20.03 | 20.06 | 19.99 | 20.09 | 20.13 | 20.18 | 20.28 | 20.43 | 20.47 |
| 2 | 30.09 | 29.90 | 29.90 | 29.94 | 29.96 | 29.92 | 29.90 | 29.95 | 30.00 | 30.04 | 30.12 | 30.16 |
| 1 | 57.22 | 57.22 | 57.22 | 57.22 | 57.10 | 57.07 | 57.07 | 57.07 | 57.07 | 57.07 | 57.07 | 57.07 |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

T_{down} (พิกเซล)

T_{up} (พิกเซล)

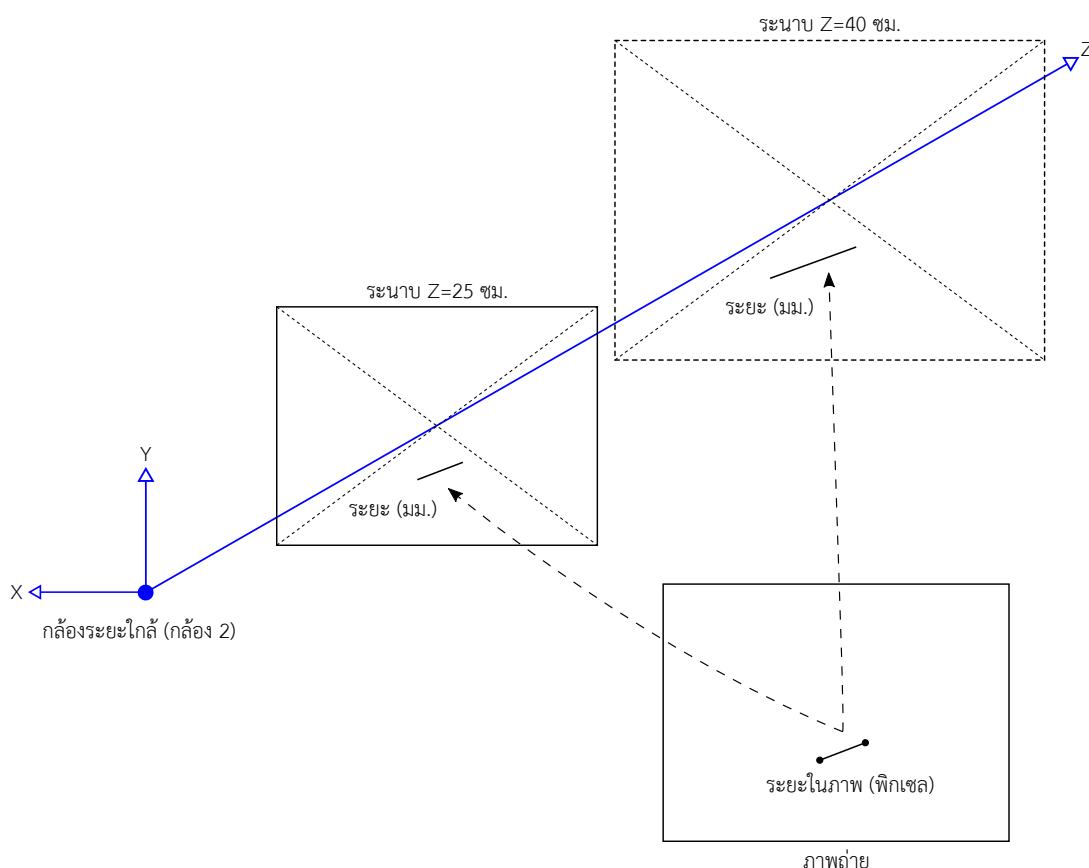
รูปที่ 4.4: ระยะห่างเฉลี่ยระหว่างแนวรอยกริดที่ตรวจพบกับแนวรอยกริดอ้างอิง โดยแสดง ช่วงของค่า T_{up} , T_{down} ที่ทำให้ระยะห่างมีค่าน้อยกว่าตำแหน่งอื่นโดยรอบ (ในกรอบเส้นประ)

ระยะคลาดเคลื่อนของแนวรอยกริดตามรูปที่ 4.4 เป็นระยะในหน่วยพิกัดภาพหรือพิกเซลเท่านั้น โดยหากต้องการทราบระยะคลาดเคลื่อนในหน่วยเมตริกซ์ (มิลลิเมตร) จะต้องทำการแปลงระยะพิกเซลไปเป็นระยะในหน่วยเมตริกซ์ ซึ่งระยะในภาพถ่ายจะสัมพันธ์กับระยะในระนาบ XY ใดๆ ของกล้องระยะใกล้ ตัวอย่างตามรูปที่ 4.5 โดยหากระยะที่ปรากฏในภาพมีตำแหน่งในระนาบความลึกที่ต่างกัน จะส่งผลให้ระยะในหน่วยเมตริกซ์ไม่เท่ากัน อีกทั้งตำแหน่งต่างๆ ในภาพ เช่นบริเวณกลางภาพหรือบริเวณขอบภาพ ต่างส่งผลถึงการแปลงค่าระยะพิกเซลไปเป็นระยะในหน่วยเมตริกซ์เช่นกัน

ระยะคลาดเคลื่อนของแนวรอยกริดจากผลการทดลองในรูปที่ 4.4 เป็นค่าเฉลี่ยของแนวรอยกริดที่ตรวจพบซึ่งไม่ได้ระบุตำแหน่งของระยะนั้นภายในภาพ และไม่มีระยะความลึกของภาพรวมอยู่ ซึ่งหากต้องการทราบระยะคลาดเคลื่อนในหน่วยเมตริกซ์จะต้องแปลงค่าระยะพิกเซลไปเป็นระยะในหน่วยมิลลิเมตร ในแต่ละภาพ ทั้งนี้เพื่อเป็นการประมาณอย่างคร่าวๆของผลการทดลองในรูปที่ 4.4 จึงเลือกคำนวณระยะคลาดเคลื่อนของแนวรอยกริดอย่างง่าย โดยประมาณระยะคลาดเคลื่อนของแนวรอยกริดให้อยู่ในตำแหน่งกึ่งกลางภาพ และเป็นระยะที่เอียงทำมุม 45 องศาทั้งแกนนอน X และแกน Y และมีระยะความลึกเฉลี่ยที่ 32.5 เซนติเมตร

ซึ่งเป็นค่ากลางของระยะถ่ายภาพของกล้องระยะใกล้ ดังนั้นระยะคลาดเคลื่อนของแนวรอยกริดโดยเฉลี่ยที่ 13 พิกเซล ตามกลุ่มในเส้นประตามรูปที่ 4.4 สามารถคำนวณเป็นระยะในระนาบ XY เท่ากับ 9.0 มิลลิเมตร ซึ่งสามารถคำนวณได้ด้วยสมการที่ 2.35

เนื่องจากการวัดผลการตรวจพบของแนวรอยกริดเลือกใช้อัตราส่วนการซ้อนทับที่ 0.5 ซึ่งทำให้ระยะคลาดเคลื่อนของแนวรอยกริดเป็นไปตามรูปที่ 4.4 โดยหากการวัดผลได้เพิ่มอัตราส่วนการซ้อนทับสำหรับวัดผลการตรวจพบด้วยกรอบสี่เหลี่ยมให้มีความมากขึ้น จะทำให้ระยะคลาดเคลื่อนมีค่าเฉลี่ยลดลง แต่ขณะเดียวกันจะทำให้จำนวนแนวรอยกริดที่ตรวจพบลดลงเช่นกัน



รูปที่ 4.5: ตำแหน่งในภาพถ่ายเมื่อถูกแปลงไปเป็นตำแหน่งในพิกัดสามมิติ ในระนาบความลึกที่ 25 เซนติเมตร กับ 40 เซนติเมตร จากกล้องระยะใกล้

บทที่ 5

บทสรุปและข้อเสนอแนะ

วิทยานิพนธ์ฉบับนี้นำเสนอระบบมองเห็นสำหรับหุ่นยนต์ทำสวนยางพาราอัตโนมัติด้วยกระบวนการกริดและเก็บน้ำยางจากถ้วยรองน้ำยางผ่านแขนหุ่นยนต์ โดยใช้งานกล้องถ่ายภาพสีและภาพความลึก (RGB-D) รุ่นอินเทลเรียลเซนส์โมเดล D400 จำนวน 2 ตำแหน่ง ซึ่งกล้องถ่ายภาพจะทำงานในสองขั้นตอนที่ต่อเนื่องกัน โดยในขั้นแรก กล้องระยะใกล้หรือกล้องถ่ายภาพมุมกว้าง (กล้อง 1) ทำหน้าที่ตรวจหาตำแหน่งรอยกริดและถ้วยรองน้ำยางในมุมมองภาพรวม ซึ่งเป็นระยะที่ทำให้มุมรับภาพครอบคลุมหน้ายางทั้งหมด จากนั้นกล้องในระยะใกล้ (กล้อง 2) จะเคลื่อนเข้าหาตำแหน่งที่ได้จากกล้องมุมกว้างเพื่อตรวจหาแนวรอยกริดในรายละเอียดที่มากกว่ากล้องมุมกว้าง

การเก็บข้อมูลภาพถ่ายรอยกริดได้ถ่ายภาพต้นยางพาราจากสวนยางพาราในช่วงเวลากลางวันหรือในช่วงเวลาเช้าตรู่ที่มีสภาพแสงน้อยเท่านั้น ซึ่งสอดคล้องกับสภาพแวดล้อมในสวนยางพาราขณะกริดของเกษตรกรทางภาคใต้ของประเทศไทย วิทยานิพนธ์ได้นำเสนออุปกรณ์ถ่ายภาพหน้ายางในทั้งสองระยะ ซึ่งจะใช้งานไฟส่องสว่างร่วมกับกล้องถ่ายภาพเพื่อการทำงานของกล้องในที่มืดต่างกันออกไป โดยวิทยานิพนธ์ฉบับนี้ได้รวบรวมข้อมูลภาพถ่ายหน้ายางในรูปแบบชุดภาพซึ่งประกอบด้วยภาพถ่ายสีและภาพความลึก ทั้งสิ้น 2 ชุดข้อมูล ได้แก่ ชุดภาพถ่ายด้วยกล้องถ่ายภาพมุมกว้างจำนวน 500 ชุดภาพ โดยใช้ไฟส่องสว่างแบบคงที่ที่วางในตำแหน่งฐานของหุ่นยนต์ในการถ่ายภาพ และชุดภาพถ่ายด้วยกล้องระยะใกล้จำนวน 1,405 ชุดภาพ ซึ่งใช้ชุดไฟส่องสว่างระยะใกล้ร่วมในการถ่ายภาพ

การกำหนดตำแหน่งอ้างอิงของรอยกริดและถ้วยรองน้ำยางในภาพ วิทยานิพนธ์ได้นำเสนอการใช้กรอบสี่เหลี่ยมล้อมรอบ (Bounding Box) เพื่อกำหนดตำแหน่งอ้างอิงให้กับรอยกริดและถ้วยรองน้ำยางในภาพถ่ายมุมกว้าง และใช้งานกรอบสี่เหลี่ยมล้อมรอบพร้อมทั้งเส้นโค้งหลายมุม สำหรับกำหนดแนวรอยกริดอ้างอิงในภาพถ่ายระยะใกล้ โดยในการกำหนดแนวรอยกริดอ้างอิงด้วยเส้นโค้งหลายมุมได้นำเสนอวิธีการกำหนดตำแหน่งแนวรอยกริดอ้างอิงด้วยมือ พร้อมทั้งวิธีการสร้างแนวรอยกริดอ้างอิงใหม่จากการปรับตำแหน่งแนวรอยกริดอ้างอิงที่กำหนดด้วยมือโดยอาศัยข้อมูลจากภาพถ่ายความลึก ซึ่งให้ผลการปรับตำแหน่งได้สำเร็จสูงสุดที่ 91%

การตรวจหาตำแหน่งรอยกริดในชุดภาพมุมกว้างได้นำเสนอ 2 วิธีที่ต่างกัน ได้แก่ 1.วิธีตรวจหารอยกริดในภาพจากกล้องถ่ายภาพมุมกว้าง ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี ซึ่งงานใช้ทั้งภาพถ่ายสีและภาพถ่ายความลึกในการประมวลผลวิธีตรวจหาให้ความแม่นยำเฉลี่ยสูงสุดที่ 35.4% เมื่อวัดผลด้วยอัตราส่วนซ้อนทับสำหรับวัดผลที่ 0.5 และในกรณีเมื่อประมวลผลภาพด้วยโมเดลสี $L^*a^*b^*$ และ 2.วิธีตรวจหารอยกริดและถ้วยรองน้ำยางในภาพจากกล้องถ่ายภาพมุมกว้างด้วยโครงข่ายประสาทจำลองเชิงลึกแบบคอนโวลูชันชนิด Faster-RCNN ที่ทำงานร่วมกับ MobileNetV2 ครบทุกเลเยอร์ โดยประมวลผลภาพขนาด 3 ช่องสัญญาณ โดยเพิ่มการคำนวณภาพถ่ายสีและภาพความลึกก่อนป้อนเป็นอินพุตให้กับเน็ตเวิร์ก ให้ความแม่นยำเฉลี่ยสูงสุดที่ 80.3%

วิธีการตรวจหาแนวรอยกริดในชุดภาพถ่ายระยะใกล้ ใช้เทคนิคการประมวลผลภาพถ่ายสี 2 ภาพ ซึ่งภาพทั้งสองจะต้องเป็นภาพที่มีเงาเกิดขึ้นตามแนวรอยกริด และเป็นภาพหน้ายางที่มีความสว่างที่สม่ำเสมอและที่ไม่มีเงาเกิดขึ้นตามแนวรอยกริด เพื่อให้สามารถใช้การหาความแตกต่างระหว่าง 2 ภาพ ในการตรวจหาแนวรอยกริดได้ วิธีการตรวจหาแนวรอยกริดสามารถตรวจหาแนวรอยกริดด้วยกรอบสี่เหลี่ยมล้อมรอบแนวรอยกริดได้แม่นยำเฉลี่ยสูงสุดที่ 89.9% เมื่อวัดผลด้วยอัตราส่วนซ้อนทับสำหรับวัตถุที่ 0.5 ในขณะเดียวแนวรอยกริดที่ได้มีความคลาดเคลื่อนจากแนวรอยกริดอ้างอิงเฉลี่ยที่ 13 พิกเซล ในภาพขนาดความละเอียด 1280x720 พิกเซล เมื่อวัดผลด้วยระยะห่างแบบเฮาส์ดอพอพ์ ซึ่งคิดเป็นระยะคลาดเคลื่อนเฉลี่ยประมาณ 9.0 มิลลิเมตรในระนาบความลึกที่ขนานกับระนาบรับภาพของกล้อง

การปรับปรุงวิธีการตรวจหาแนวรอยกริดในภาพถ่ายจากกล้องระยะใกล้ด้วยวิธีการตรวจหาแนวรอยกริดในระยะใกล้ด้วยเทคนิคการเลื่อนกรอบสี่เหลี่ยมล้อมรอบภายในภาพ และความเอียงของแนวรอยกริดผ่านการจัดกลุ่มสี สามารถทำได้โดย เปลี่ยนรูปแบบการคำนวณการจัดกลุ่มสีในขั้นตอนการแยกพื้นที่โทนสีแดงออกจากพื้นที่อื่นในลำดับ ซึ่งในวิทยานิพนธ์นี้ได้ใช้วิธีจัดกลุ่มค่าสีของทุกๆ พิกเซลในพื้นที่ลำดับ ทำให้การประมวลผลภาพขนาดใหญ่จะมีพิกเซลที่จะต้องคำนวณเป็นจำนวนมาก ซึ่งอาจได้รับผลกระทบจากค่าพิกเซลที่เป็นสัญญาณรบกวนส่งผลต่อความแม่นยำ และทำให้เวลาประมวลผลต่อภาพมาก ทั้งนี้สามารถปรับปรุงได้โดยเปลี่ยนจากการพิจารณาค่าสีในทุกพิกเซล เป็นการพิจารณาค่าสีที่ตำแหน่งพิกเซลข้ามพิกเซล ซึ่งคล้ายกับการทำคอนโวลูชันแบบข้ามพิกเซลที่อยู่ติดกัน (Dilated Convolution) ซึ่งเป็นการลดจำนวนพิกเซลที่จะต้องจัดกลุ่มค่าสี หรือการเปลี่ยนรูปแบบการสร้างกรอบสี่เหลี่ยมล้อมรอบเริ่มต้น จากเดิมที่กำหนดกรอบสี่เหลี่ยมล้อมรอบทั่วพื้นที่ลำดับที่ปรากฏในภาพ อาจลดจำนวนได้โดยสร้างกรอบสี่เหลี่ยมเฉพาะพื้นที่บริเวณกลุ่มสีโทนแดง ซึ่งอาจส่งผลให้ความแม่นยำและเวลาที่ใช้ประมวลผลต่อภาพดีขึ้น เป็นต้น

สำหรับการปรับปรุงวิธีการตรวจหาแนวรอยกริดในกล้องระยะใกล้ อาจเปลี่ยนวิธีการเลือกค่าพารามิเตอร์ T_{up} และ T_{down} ให้เหมาะกับแนวรอยกริดในแต่ละภาพได้ หรือการใช้งานคอนโวลูชันเน็ตเวิร์กร่วมกับวิธีการตรวจหาที่ได้นำเสนอเพื่อเพิ่มความแม่นยำ พร้อมทั้งเพิ่มเป้าหมายของการตรวจหาให้สามารถระบุตำแหน่งจุดเริ่มต้นและจุดสิ้นสุดของแนวรอยกริดได้ โดยอาศัยข้อมูลความลึกจากภาพถ่ายความลึก

การพัฒนาต่อยอดระบบมองเห็นของระบบหุ่นยนต์ทำสวนยางพาราอัตโนมัติในประเด็นอื่นๆ สามารถทำได้ทั้งขั้นตอนการเก็บข้อมูลภาพถ่ายด้วยกล้องแต่ละระยะ เช่น การปรับปรุงระบบให้แสงสว่างกับหน้ายาง, หรือการใช้กล้องถ่ายภาพย่านความถี่อื่นที่ไม่ต้องพึ่งพาการใช้ไฟส่องสว่าง เป็นต้น หรือในประเด็นการปรับปรุงวิธีการตรวจหาแนวรอยกริดในภาพถ่ายแต่ละระยะ เช่น การทำให้วิธีการตรวจหาแนวรอยกริดในกล้องถ่ายภาพมุมกว้างให้มีเวลาประมวลผลต่อภาพที่ดีขึ้น, มีความแม่นยำมากขึ้น, หรือทนทานต่อสภาพแสงอื่นนอกเหนือจากสภาพแสงในที่มืด เป็นต้น

บรรณานุกรม

- [1] Abraham, P. (1992). Tapping of *Hevea brasiliensis*. *Natural Rubber*, page 263.
- [2] Alt, H. and Scharf, L. (2008). Computing the Hausdorff distance between curved objects. *International Journal of Computational Geometry & Applications*, 18(04):307–320.
- [3] Anucha Watanapa, S. S., Wisitsree Wiyaratn, P. P., and Rawungsook, N. (2010). COATING LATEX CUP WITH PTFE TO DECREASE LATEX ATTACHMENT. *วารสารวิศวกรรมศาสตร์ ปีที่ 2 ฉบับที่ 2 วิศวกรรมเทคโนโลยีไทย*, 2(9):70.
- [4] Arbeláez, P., Maire, M., Fowlkes, C., and Malik, J. (2011). Contour Detection and Hierarchical Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5): 898–916.
- [5] Benhamou, S. (2004). How to reliably estimate the tortuosity of an animal’s path: Straightness, sinuosity, or fractal dimension? *Journal of theoretical biology*, 229:209–20.
- [6] Bianco, S., Cadène, R., Celona, L., and Napoletano, P. (2018). Benchmark analysis of representative deep neural network architectures. *IEEE Access*, 6:64270–64277.
- [7] Caesar, H., Uijlings, J. R. R., and Ferrari, V. (2016). COCO-Stuff: Thing and Stuff Classes in Context. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1209–1218.
- [8] Chambon, B., Angthong, S., Kongmanee, C., Somboonsuke, B., Mazon, S., Puengcharoen, A., Martin, C., and Lacote, R. (2014). A Comparative Analysis of Smallholders Tapping Practices in Four Rubber Producing Regions of Thailand. In *Advances in Rubber*, volume 844 of *Advanced Materials Research*, pages 34–37. Trans Tech Publications Ltd.
- [9] Chantuma, P., Lacote, R., Leconte, A., and Gohet, E. (2011). An innovative tapping system, the double cut alternative, to improve the yield of *hevea brasiliensis* in thai rubber plantations. *Field crops research*, 121(3):416–422.
- [10] Corp, C. T. P. I. (2019). Intel® RealSense™ Camera 400 Series (DS5) Product Family Datasheet. [Online].
- [11] Corp, I. (2018). Production CAD Files for Intel® RealSense™ D400 Series. [Online].
- [12] Dai, J., Li, Y., He, K., and Sun, J. (2016). R-FCN: Object Detection via Region-based Fully Convolutional Networks. *CoRR*, abs/1605.06409.

- [13] Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893 vol. 1.
- [14] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Li, F.-F. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255.
- [15] Di Stefano, L. and Bulgarelli, A. (1999). A simple and efficient connected components labeling algorithm. In *Proceedings 10th International Conference on Image Analysis and Processing*, pages 322–327. IEEE.
- [16] Everingham, M., Eslami, S. M. A., Van Gool, L., Williams, C. K. I., Winn, J., and Zisserman, A. (2015). The Pascal Visual Object Classes Challenge: A Retrospective. *International Journal of Computer Vision*, 111(1):98–136.
- [17] Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., and Zisserman, A. (2010). The Pascal Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision*, 88(2):303–338.
- [18] Fernández, R., Montes, H., Surdilovic, J., Surdilovic, D., Gonzalez-De-Santos, P., and Armada, M. (2018). Automatic detection of field-grown cucumbers for robotic harvesting. *IEEE Access*, 6:35512–35527.
- [19] Fujioka, H., Kano, H., Egerstedt, M., and Martin, C. (2005). Smoothing Spline Curves and Surfaces for Sampled Data. *Int. J. of Innovative Computing, Information and Control*, 1.
- [20] Girshick, R. (2015). Fast R-CNN. In *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), ICCV '15*, pages 1440–1448, Washington, DC, USA. IEEE Computer Society.
- [21] Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR '14*, page 580–587, USA. IEEE Computer Society.
- [22] Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2016). Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 38(1):142–158.

- [23] Gohet, E. and Chantuma, P. (2003). Reduced tapping frequency and DCA tapping systems Research towards improvement of Thailand rubber plantations productivity. *IRRDB annual meeting, 15 - 16 September, Chiang Mai, Thailand.*
- [24] Gohet, E., Lacote, R., Leconte, A., Chapuset, T., Rivano, F., and Chambon, B. (2016). Improving rubber smallholdings productivity and resilience through adoption of good agricultural practices. In *Focus Forum on Natural Rubber Sustainability. International Rubber Study Group, Singapore.*
- [25] Gongal, A., Silwal, A., Amatya, S., Karkee, M., Zhang, Q., and Lewis, K. (2016). Apple crop-load estimation with over-the-row machine vision system. *Computers and Electronics in Agriculture*, 120:26–35.
- [26] Grana, C., Borghesani, D., and Cucchiara, R. (2009). Connected Component Labeling Techniques on Modern Architectures. In Foggia, P., Sansone, C., and Vento, M., editors, *Image Analysis and Processing – ICIAP 2009*, pages 816–824, Berlin, Heidelberg. Springer Berlin Heidelberg.
- [27] Grunnet-Jepsen, A., Sweetser, J. N., and Woodfill, J. (2018). Best-Known-Methods for Tuning Intel® RealSense™ D400 Depth Cameras for Best Performance. Technical report.
- [28] Hamuda, E., Glavin, M., and Jones, E. (2016). A survey of image processing techniques for plant extraction and segmentation in the field. *Computers and Electronics in Agriculture*, 125:184–199.
- [29] Haun, A. M. and Peli, E. (2013). Perceived contrast in complex images. *Journal of Vision*, 13(13):3–3.
- [30] He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep Residual Learning for Image Recognition. *CoRR*, abs/1512.03385.
- [31] Hou, X., Yuille, A., and Koch, C. (2013). Boundary Detection Benchmarking: Beyond F-Measures. In *Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2123–2130.
- [32] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *CoRR*, abs/1704.04861.

- [33] Huang, J., Rathod, V., Sun, C., Zhu, M., Balan, A. K., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S., and Murphy, K. (2016). Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3296–3297.
- [34] Huttenlocher, D. P., Klanderman, G. A., and Rucklidge, W. J. (1993). Comparing images using the Hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence*, 15(9):850–863.
- [35] Ibraheem, N., Hasan, M., Khan, R. Z., and Mishra, P. (2012). Understanding Color Models: A Review. *ARPN Journal of Science and Technology*, 2.
- [36] Irie, N., Taguchi, N., Horie, T., and Ishimatsu, T. (2009). Asparagus harvesting robot coordinated with 3-d vision sensor. In *2009 IEEE International Conference on Industrial Technology*, pages 1–6.
- [37] J Green, P. and Silverman, B. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach.*, volume 50.
- [38] Jain, A. K. and Farrokhnia, F. (1990). Unsupervised texture segmentation using Gabor filters. In *1990 IEEE International Conference on Systems, Man, and Cybernetics Conference Proceedings*, pages 14–19.
- [39] Juman, M. A., Wong, Y. W., Rajkumar, R. K., and Goh, L. J. (2016). A novel tree trunk detection method for oil-palm plantation navigation. *Computers and Electronics in Agriculture*, 128:172–180.
- [40] Kang, H. and Chen (2019). Fruit detection and segmentation for apple harvesting using visual sensor in orchards. *Sensors*, 19:4599.
- [41] Keselman, L., Iselin Woodfill, J., Grunnet-Jepsen, A., and Bhowmik, A. (2017). Intel Realsense Stereoscopic Depth Cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 1–10.
- [42] Krizhevsky, A. (2009). Learning Multiple Layers of Features from Tiny Images.
- [43] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *NIPS*.
- [44] Kunghun, W. and Tantrapiwat, A. (2018). Development of a vision based mapping in rubber tree orchard. In *2018 International Conference on Engineering, Applied Sciences, and Technology (ICEAST)*, pages 1–4. IEEE.

- [45] Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., Pont-Tuset, J., Kamali, S., Popov, S., Mallocci, M., Duerig, T., and Ferrari, V. (2018). The Open Images Dataset V4: Unified image classification, object detection, and visual relationship detection at scale. *arXiv:1811.00982*.
- [46] Leu, A., Razavi, M., Langstädtler, L., Ristić-Durrant, D., Raffel, H., Schenck, C., Gräser, A., and Kuhfuss, B. (2017). Robotic green asparagus selective harvesting. *IEEE/ASME Transactions on Mechatronics*, 22(6):2401–2410.
- [47] Lin, G., Tang, Y., Zou, X., Xiong, J., and Li, J. (2019). Guava detection and pose estimation using a low-cost rgb-d sensor in the field. In *Sensors*.
- [48] Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In Fleet, D., Pajdla, T., Schiele, B., and Tuytelaars, T., editors, *Computer Vision – ECCV 2014*, pages 740–755, Cham. Springer International Publishing.
- [49] Linker, R. and Kelman, E. (2015). Apple detection in nighttime tree images using the geometry of light patches around highlights. *Computers and Electronics in Agriculture*, 114:154–162.
- [50] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S. E., Fu, C., and Berg, A. C. (2015). SSD: Single Shot MultiBox Detector. *CoRR*, abs/1512.02325.
- [51] Liu, X., Zhao, D., Jia, W., Ruan, C., Tang, S., and Shen, T. (2016). A method of segmenting apples at night based on color and position information. *Computers and Electronics in Agriculture*, 122:118–123.
- [52] Longsheng, F., Bin, W., Yongjie, C., Shuai, S., Gejima, Y., and Kobayashi, T. (2015). Kiwifruit recognition at nighttime using artificial lighting based on machine vision. *International Journal of Agricultural and Biological Engineering*, 8(4):52–59.
- [53] Mahathaninwong, N., Chucheep, T., Muangdee, N., Kongtim, P., Ananchaoenwong, A., Marthosa, S., and Sumangkay, K. (2016). Para rubber tapping behavior, using of para rubber tapping knife behavior and automatic para rubber tapping machine concept evaluation of rubber farmers. *Journal of Agricultural Research and Extension*, 33(1):66–76.
- [54] Maliackal, J. V. C., Asif, K. A., Sajith, P. A., and Joseph, S. K. (2017). Advanced Rubber Tree Tapping Machine. *Int. J. Res. Innov. Eng. Sci. Technol.*, 2(5):253–263.
- [55] Martin, D., Fowlkes, C., Tal, D., and Malik, J. (2001). A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring

- Ecological Statistics. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 2, pages 416–423 vol.2.
- [56] Microsoft (2018). Visual Object Tagging Tool (VoTT): An electron app for building end to end Object Detection Models from Images and Videos. <https://github.com/microsoft/VoTT>.
- [57] Monta, M., Kondo, N., and Shibano, Y. (1995). Agricultural robot in grape production system. In *Proceedings of 1995 IEEE International Conference on Robotics and Automation*, volume 3, pages 2504–2509 vol.3.
- [58] Nair, K. P. (2010). Rubber (*hevea brasiliensis*). *The agronomy and economy of important tree crops of the developing world*, pages 237–273.
- [59] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66.
- [60] Perona, P. and Malik, J. (1990). Detecting and localizing edges composed of steps, peaks and roofs. *[1990] Proceedings Third International Conference on Computer Vision*, pages 52–57.
- [61] Pethin, D., Nakkanong, K., and Nualsri, C. (2015). Performance and genetic assessment of rubber tree clones in Southern Thailand. *Scientia Agricola*, 72:306 – 313.
- [62] Plataniotis, K. and Venetsanopoulos, A. (2000). *Color Image Processing and Applications*.
- [63] P.Thala, N.Kaewhgam, K.Kumnornaew, and K.Satjawattana (2014). Effects of tapping time period and tapping system on latex yield of rubber trees (*Hevea brasiliensis*) at University of Phayao, Phayao, Thailand. *Khon Kaen Agriculture Journal*, Vol.42 SUPPLEMENT 4.
- [64] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 779–788. IEEE.
- [65] Redmon, J. and Farhadi, A. (2017). YOLO9000: Better, Faster, Stronger. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6517–6525.
- [66] Ren, S., He, K., Girshick, R., and Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1137–1149.

- [67] Rosell-Polo, J., Gregorio Lopez, E., Gené-Mola, J., Llorens Calveras, J., Torrent, X., Arnó, J., and Escolà, A. (2017). Kinect v2 sensor-based mobile terrestrial laser scanner for agricultural outdoor applications. *IEEE/ASME Transactions on Mechatronics*, In press.
- [68] S. Hou, H. and C. Andrews, H. (1979). Cubic Splines for Image Interpolation and Digital Filtering. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 26:508–517.
- [69] Sandler, M., Howard, A. G., Zhu, M., Zhmoginov, A., and Chen, L. (2018). Inverted Residuals and Linear Bottlenecks: Mobile Networks for Classification, Detection and Segmentation. *CoRR*, abs/1801.04381.
- [70] Silwal, A., Davidson, J. R., Karkee, M., Mo, C., Zhang, Q., and Lewis, K. (2017). Design, integration, and field evaluation of a robotic apple harvester. *Journal of Field Robotics*, 34(6):1140–1159.
- [71] Simien, A. and Penot, E. (2011). Current Evolution of Smallholder Rubber-Based Farming Systems in Southern Thailand. *Journal of Sustainable Forestry*, 30:247–260.
- [72] Soumya, S. J., Vishnu, R. S., Arjun, R. N., and Bhavani, R. R. (2016). Design and testing of a semi automatic rubber tree tapping machine (sart). In *2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, pages 1–4. IEEE.
- [73] Sural, S., Qian, G., and Pramanik, S. (2002). Segmentation and histogram generation using the hsv color space for image retrieval. *Proceedings. International Conference on Image Processing*, 2:II–II.
- [74] Susanto, H., Ali, S., et al. (2019). The Design of Flexible Rubber Tapping Tool with Settings the Depth and Thickness Control. In *IOP Conference Series: Materials Science and Engineering*, volume 506. IOP Publishing.
- [75] Vijayakumar, K., Gohet, E., Thomas, K., Xiaodi, W., Lakshman, R., Sopchoke, P., Karunaichamy, K., Mohd Akbar, S., et al. (2009). Revised international notation for latex harvest technology. *Journal of Rubber Research*.
- [76] Wang, C., Tang, Y., Zou, X., SiTu, W., and Feng, W. (2017). A robust fruit image segmentation algorithm against varying illumination for vision system of fruit harvesting robot. *Optik*, 131:626–631.
- [77] Wongtanawijit, R. and Kaorapapong, T. (2018). Rubber Tapped Path Detection using K-means Color Segmentation and Distance to Boundary Feature. In *2018 15th International*

- Conference on Electrical Engineering / Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 126–129.
- [78] Wongtanawijit, R. and Khaorapapong, T. (2019). Rubber Tapping Position and Harvesting Cup Detection Using Faster-RCNN with MobileNetV2. In *2019 23rd International Computer Science and Engineering Conference (ICSEC)*, pages 335–339.
- [79] Xiang, R. (2018). Image segmentation for whole tomato plant recognition at night. *Computers and Electronics in Agriculture*, 154:434–442.
- [80] Xiang, R., Ying, Y., and Jiang, H. (2013). A recognition algorithm for occluded tomatoes based on circle regression. In *2013 6th International Congress on Image and Signal Processing (CISP)*, volume 2, pages 713–717.
- [81] Yatawara, Y., Brito, W., Perera, M., and Balasuriya, D. (2019). “Appuhamy”-The Fully Automatic Rubber Tapping Machine. *ENGINEER*, 27:1.
- [82] Zhang, C., Yong, L., Chen, Y., Zhang, S., Ge, L., Wang, S., and Li, W. (2019). A Rubber-Tapping Robot Forest Navigation and Information Collection System Based on 2D LiDAR and a Gyroscope. *Sensors*, 19(9):2136.
- [83] Zhao, Y., Gong, L., Huang, Y., and Liu, C. (2016). A review of key techniques of vision-based control for harvesting robot. *Computers and Electronics in Agriculture*, 127:311–323.

ภาคผนวก ก

ผลงานตีพิมพ์เผยแพร่จากวิทยานิพนธ์

ก.1 บทความวิจัยเสนอในที่ประชุมวิชาการและมีการพิมพ์รวมเล่ม

- ก.1.1 Wongtanawijit, R. and Kaorapapong, T. (2018). Rubber Tapped Path Detection using K-means Color Segmentation and Distance to Boundary Feature. In *2018 15th International Conference on Electrical Engineering / Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 126–129
- ก.1.2 Wongtanawijit, R. and Khaorapapong, T. (2019). Rubber Tapping Position and Harvesting Cup Detection Using Faster-RCNN with MobileNetV2. In *2019 23rd International Computer Science and Engineering Conference (ICSEC)*, pages 335–339

ก.2 บทความวิชาการ

ก.2.1 -