

การตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท

นายธีรพัฒน์ เฉลยวุฒิ



จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)

เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)

are the thesis authors' files submitted through the University Graduate School.

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2557

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

BUILDING DETECTION FROM STREET-SIDE IMAGES IN RURAL AREA

Mr. Teerapat Chaloeivoot



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

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2014

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท
โดย	นายธีรพัฒน์ เฉลยวุฒิ
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.สีบสกุล พิภพมงคล

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วน
หนึ่งของการศึกษาตามหลักสูตรปริญญาโทบัณฑิต

.....คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.บัณฑิต เอื้ออาภรณ์)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(รองศาสตราจารย์ นงลักษณ์ โควาวิสารัช)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.สีบสกุล พิภพมงคล)

.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ธนรัตน์ ชลิตาพงศ์)

.....กรรมการ
(รองศาสตราจารย์ ดร.ไพศาล สันติธรรมนนท์)

.....กรรมการภายนอกมหาวิทยาลัย
(ดร.เดชาวุฒิชิต กัตถัญญูทวีทิพย์)

ธีรพัฒน์ เฉลยวุฒิ : การตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท (BUILDING DETECTION FROM STREET-SIDE IMAGES IN RURAL AREA) อ.ที่ปริกษานิพนธ์
 หลัก: ผศ. ดร.สืบสกุล พิภพมงคล, 80 หน้า.

ในงานวิจัยนี้ได้เสนอวิธีการตรวจหาอาคารที่อยู่ในภาพโดยใช้เทคนิคการประมวลผลภาพ ดิจิทัลและโครงข่ายประสาทเทียม โดยทั่วไปภาพถ่ายภาคพื้นดินความละเอียดสูงและข้อมูลในแต่ละ ท้องที่เป็นที่ต้องการในหลายหน่วยงาน เช่น ข้อมูลการใช้ประโยชน์จากที่ดินในพื้นที่ชนบท เป็นต้น วิธีการที่เสนอนี้ได้ทดลองกับภาพถ่ายข้างทางในชนบทของประเทศไทย โดยแบ่งเป็น 2 ขั้นตอน หลัก ขั้นตอนแรกคือการตรวจหาอาคาร เริ่มจากการกำจัดวัตถุที่ไม่เกี่ยวข้องในภาพถ่าย เช่น ต้นไม้ และท้องฟ้า จากนั้นแบ่งส่วนภาพเป็นหลายบริเวณด้วยสีของวัตถุและหาบริเวณที่สนใจ วิธีการที่ใช้ ได้แก่ การกำจัดสีเขียวต้นไม้ออกจากภาพ การกำจัดท้องฟ้า การแบ่งส่วนภาพด้วยสีและการตรวจหา รูปร่างเรขาคณิต วิธีการเหล่านี้ถูกนำมาเพื่อใช้ตรวจหาส่วนของอาคาร หลังจากได้ส่วนของอาคาร ก็ จะใช้วิธีการโปรเจกชันในการรวมกลุ่มของวัตถุเพื่อกำหนดเป็นบริเวณที่สนใจ ขั้นตอนที่สองคือการ รู้จำอาคาร เริ่มจากการทำให้ข้อมูลเป็นบรรทัดฐานเดียวกันจากนั้นนำเข้าสู่โครงข่ายประสาทเทียม แบบคอนโวลูชันนัล ข้อมูลฝึกของเครือข่ายประสาทเทียมใช้ภาพที่มีอาคาร 3,995 ภาพ ภาพที่ไม่มี อาคาร 3,348 ภาพและได้ทดสอบกับภาพที่มีอาคาร 1,832 ภาพและภาพที่ไม่มีอาคาร 2,104 ภาพ ภาพทั้งหมดมาจากกูเกิลสตรีทวิว โดยมนุษย์เป็นผู้ตรวจสอบความถูกต้อง ผลลัพธ์ที่ได้มีความถูกต้อง สำหรับภาพที่มีอาคารเฉลี่ยร้อยละ 87.50 และความถูกต้องสำหรับภาพที่ไม่มีอาคารเฉลี่ยร้อยละ 97.60

จุฬาลงกรณ์มหาวิทยาลัย
 CHULALONGKORN UNIVERSITY

ภาควิชา วิศวกรรมคอมพิวเตอร์

ลายมือชื่อนิสิต

สาขาวิชา วิศวกรรมคอมพิวเตอร์

ลายมือชื่อ อ.ที่ปรึกษาหลัก

ปีการศึกษา 2557

5470230221 : MAJOR COMPUTER ENGINEERING

KEYWORDS: BUILDING DETECTION / TERRESTRIAL IMAGE / IMAGE PROCESSING / IMAGE SEGMENTATION / CONVOLUTIONAL NEURAL NETWORK / MACHINE LEARNING

TEERAPAT CHALOEIVOOT: BUILDING DETECTION FROM STREET-SIDE IMAGES IN RURAL AREA. ADVISOR: ASST. PROF. SUEBSKUL PHIPHOBMONGKOL, Ph.D., 80 pp.

This research presents a method to detect and recognize buildings in terrestrial images. High-resolution terrestrial images are normally taken from land survey vehicles. These images and other surveyed data along roads are needed by many agencies that require new data as time passes by. Land use in rural area is an example that needs information about buildings and can benefit from terrestrial images. The proposed method was aimed to detect and recognize buildings in terrestrial images to benefit the above needs. The method consists of two stages. The first stage is building detection. It removes unwanted objects, performs image segmentation and finds regions of interest. Image processing techniques such as greenness removal, sky removal, color segmentation, color detection, shape detection are used. After that, building parts detection, projection profiles finding and the building determination are performed. The second stage is building recognition. It begins with image normalization, and uses convolutional neural network to recognize buildings. The method can identify a partial building if the whole building is not shown in an image. The training set contained 3,995 building images and 3,348 images without building from many provinces in Thailand. The proposed method was tested on 3,936 images (1,832 images with buildings and 2,104 images without buildings). The images were from Google Street View. The accuracy was determined by human inspection. The method gave promising results with an average of 87.50% accuracy for images with buildings and an average of 97.60% accuracy for images without building.

Department: Computer Engineering Student's Signature

Field of Study: Computer Engineering Advisor's Signature

Academic Year: 2014

กิตติกรรมประกาศ

วิทยานิพนธ์นี้เสร็จสมบูรณ์ได้ด้วยคำแนะนำช่วยเหลือจาก ผศ.ดร.สืบสกุล พิภพมงคล
ซึ่งได้สละเวลาและให้แง่คิดใหม่ๆรวมทั้งวิธีการแก้ไขปัญหาต่างๆ

ขอขอบคุณคณาจารย์ที่เป็นกรรมการในการสอบวิทยานิพนธ์ฉบับนี้ ที่ให้คำแนะนำ
และเสนอแนวทางในการพัฒนางานวิจัยให้ดียิ่งขึ้น ตลอดจนการแก้ไขปรับปรุงข้อบกพร่องต่างๆ
ของวิทยานิพนธ์ฉบับนี้

สุดท้ายนี้ขอขอบพระคุณครอบครัวที่ให้การสนับสนุนและเป็นกำลังใจให้กับผู้วิจัยเสมอ
มา



สารบัญ

หน้า

บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฅ
สารบัญภาพ	ฉ
บทที่ 1 บทนำ	1
1.1 ที่มาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์.....	2
1.3 ขอบเขตงานวิจัย.....	3
1.4 ขั้นตอนดำเนินงานวิจัย	3
1.5 ประโยชน์ที่ได้รับ.....	3
บทที่ 2 ทฤษฎีและหลักการที่เกี่ยวข้อง.....	4
2.1 ทบทวนงานวิจัยที่เกี่ยวข้อง.....	4
2.1.1 การตรวจหาอาคารจากภาพถ่ายภาคพื้นดิน	4
2.1.1.1 การตรวจหาอาคาร.....	4
2.1.1.2 การตรวจหาส่วนของอาคาร	5
2.1.2 การตรวจหาอาคารจากภาพถ่ายทางอากาศ.....	5
2.2 ทฤษฎีที่เกี่ยวข้อง	6
2.2.1 ความหมายของอาคาร.....	6
2.2.1.2 ราชบัณฑิตยสถาน.....	6
2.2.1.3 กรมโยธาธิการทหารบก.....	6

2.2.3 การแปลงภาพจากภาพสี (RGB) เป็นแบบจำลองสี HSV.....	7
2.2.4 การแบ่งส่วนภาพด้วยค่าขีดแบ่ง (Thresholding)	9
2.2.5 การแปลงภาพจากภาพสี (RGB) เป็นแบบจำลองสี $YCbCr$	9
2.2.6 Automatic Seeded Region Growing.....	10
2.2.6.1 การเลือกเมล็ดพันธุ์โดยอัตโนมัติ (The automatic seed selection).....	10
2.2.6.2 ทฤษฎีการเติบโตของบริเวณเมล็ดพันธุ์ (Seed region growing)	12
2.2.6.3 ขั้นตอนการผสานบริเวณเข้าด้วยกัน (Region-merging).....	12
2.2.7 การตรวจหาเส้นตรงด้วยวิธีการแปลงฮัฟ (Line Hough Transform).....	13
2.2.8 การเพิ่มจำนวนของภาพในชุดของการฝึก	15
2.2.9 การทำให้ข้อมูลในแต่ละแกนแนลเป็นบรรทัดฐานเดียวกัน (Data normalization)...	15
2.2.10 โครงข่ายประสาทเทียมแบบคอนโวลูชันนัล	16
2.2.10.1 ชั้นคอนโวลูชัน.....	16
2.2.10.2 ชั้น ReLU.....	16
2.2.10.3 ชั้น Pooling.....	17
2.2.10.4 ชั้น Dropout	17
บทที่ 3 แนวคิดและกระบวนการทำงาน	18
3.1 การกำจัดสีท้องฟ้า	20
3.2 การกำจัดสีเขียวใบไม้.....	20
3.3 การแบ่งส่วนภาพบริเวณของวัตถุด้วยสีและการใส่ป้ายหมายเลขให้แก่ละบริเวณ	20
3.4 การลบบริเวณที่ไม่ต้องการ.....	20
3.4.1 การลบบริเวณที่มีขนาดเล็ก.....	20
3.4.2 การลบบริเวณที่มีส่วนที่เหลือของต้นไม้	21
3.4.3 การลบบริเวณที่ติดกับขอบล่างของภาพ	21

3.5 การตรวจหาส่วนประกอบของอาคาร	21
3.5.1 การตรวจหาเส้นตรง	21
3.5.1.1 การหาเส้นตรงยาวในแนวราบ	21
3.5.1.2 การหาเส้นตรงที่ขนานกันในแนวตั้ง	21
3.5.1.3 การหาสี่เหลี่ยม	22
3.5.1.4 การหาจั่วหลังคา	22
3.5.2 การตรวจหาสีอาคาร	22
3.6 การหาบริเวณของวัตถุที่สนใจ	22
3.7 การตรวจสอบขนาดและอัตราส่วนกว้างยาว	22
3.8 การรู้จำอาคารด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันนัล	23
3.8.1 การเพิ่มข้อมูลการฝึก	23
3.8.2 การเตรียมข้อมูลเพื่อเข้าสู่โครงข่ายประสาทเทียม	24
3.8.3 การฝึกโครงข่ายประสาทเทียมโครงข่ายประสาทเทียมแบบคอนโวลูชันนัล	24
บทที่ 4 การทดลองและผลการทดลอง	25
4.1 ภาพที่ใช้ในการทดลอง	25
4.2 สภาพแวดล้อมด้านซอฟต์แวร์และฮาร์ดแวร์	25
4.3 การหาค่าพารามิเตอร์ที่เหมาะสมสำหรับการประมวลผลภาพดิจิทัล	25
4.3.1 การลบบริเวณท้องฟ้า	26
4.3.2 การลบสีเขียวใบไม้	27
4.3.3 การแบ่งส่วนภาพบริเวณของวัตถุด้วยสีและการติดป้ายหมายเลขให้แก่ละบริเวณ	28
4.3.4 การลบบริเวณที่ไม่ต้องการ	29
4.3.5 การทดลองหาเส้นตรงยาวในแนวราบ	30
4.3.6 การทดลองหาเส้นตรงที่ขนานกันในแนวตั้ง	31

4.3.7 การหาสีเหลืองในภาพถ่าย	31
4.3.8 การหาจั่วหลังคา.....	32
4.3.9 การตรวจหาสีอาคาร.....	32
4.3.10 การหาขอบเขตของบริเวณของวัตถุที่สนใจ	33
4.3.11 การตรวจสอบขนาดและอัตราส่วนของบริเวณ.....	33
4.5 การนำพารามิเตอร์มาทดสอบกับภาพถ่าย.....	34
4.5.1 การลบบริเวณท้องฟ้าและการลบสีเขียวของใบไม้.....	34
4.5.2 การทดสอบหาอาคารที่มีสิ่งบดบังและไม่มีสิ่งบดบัง	34
4.5.3 การทดสอบกับภาพด้านข้างของรถสำรวจจุลีสตรีทวิว.....	36
4.6 การรู้จำอาคารด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันนัล	37
4.6.1 การทดลองเปรียบเทียบโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลกับโครงข่าย ประสาทเทียมอื่น	37
4.6.2 การทดลองหาโมเดลที่ดีที่สุดในการรู้จำอาคาร.....	39
4.7 ผลการทดลอง	41
4.7.1 ทดสอบทั้งระบบกับภาพถ่ายในชนบท.....	41
4.7.2 ทดสอบกับภาพในหัวข้อ 4.5.3	42
4.7.3 การทดลองเปรียบเทียบความถูกต้องของโครงข่ายประสาทเทียมที่ใช้ภาพที่ผ่าน กระบวนการทำงานในงานวิจัยนี้เทียบกับการนำภาพต้นฉบับเข้าสู่โครงข่าย ประสาทเทียมโดยตรง	43
4.8 การวิเคราะห์ผลการทดลอง	44
บทที่ 5 สรุปผลและข้อเสนอแนะ	46
5.1 สรุปผล.....	46
5.2 ข้อเสนอแนะ	46
รายการอ้างอิง	48

ภาคผนวก ก ภาพตัวอย่างผลลัพธ์ที่ได้จากการประมวลผลภาพดิจิทัล	52
ภาคผนวก ข ตัวอย่างการใช้ประโยชน์จากผลลัพธ์ที่ได้จากการประมวลผลภาพดิจิทัล.....	62
ภาคผนวก ค ส่วนเชื่อมต่อผู้ใช้ของโปรแกรมตรวจหาบริเวณที่สนใจในภาพถ่ายข้างทางในชนบท....	65
ภาคผนวก ง ส่วนเชื่อมต่อผู้ใช้ของโปรแกรมรู้จำอาคารจากภาพถ่ายข้างทางในชนบท.....	71
ภาคผนวก จ บทความที่เสนอในงานประชุมวิชาการ.....	73
ประวัติผู้เขียนวิทยานิพนธ์	80



สารบัญตาราง

ตารางที่ 2.1 ตารางแสดงความสัมพันธ์ระหว่าง r และ θ	14
ตารางที่ 4.1 ค่า HSV ที่เหมาะสมจากการทดลองเพื่อตรวจหาอาคาร	32
ตารางที่ 4.2 ความถูกต้องที่ได้จากการทดลองในแต่ละโมเดล.....	35
ตารางที่ 4.3 ความผิดพลาดเชิงบวกจากการทดลอง	36
ตารางที่ 4.4 ความผิดพลาดเชิงลบจากการทดลอง	37
ตารางที่ 4.5 ผลการทดลองเปรียบเทียบโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลกับโครงข่ายประสาทเทียมอื่นๆ	37
ตารางที่ 4.6 โมเดลที่ใช้ในการทดลองหาโมเดลโครงข่ายประสาทเทียมแบบคอนโวลูชันแนลที่ดีที่สุดในการรู้จำอาคาร.....	40
ตารางที่ 4.7 ความถูกต้องที่ได้จากการทดลองในแต่ละโมเดล.....	41
ตารางที่ 4.8 ความถูกต้องของโครงข่ายประสาทเทียมที่ได้จากการการผ่านกระบวนการทำงานในงานวิจัยนี้เทียบกับการนำภาพต้นฉบับเข้าสู่โครงข่ายประสาทเทียมโดยตรง	43

สารบัญภาพ

รูปที่ 1-1 ตัวอย่างอาคารที่พบได้ทั่วไปในชนบท	2
รูปที่ 2-1 ค่าความเข้มของสีจากค่า 0 ไปจนถึง 255	7
รูปที่ 2-2 แบบจำลองสี HSV	8
รูปที่ 2-3 แบบจำลองสี YC _r C _b	9
รูปที่ 2-4 แสดงหลักการทำงานของวิธี Automatic Seeded Region Growing	10
รูปที่ 2-5 ตัวอย่างผลลัพธ์แต่ละขั้นตอนของวิธีการ Automatic Seeded Region Growing.....	13
รูปที่ 2-6 รูปแสดงตัวแปรสำหรับสมการ 2.20	14
รูปที่ 2-7 ตัวอย่างผลลัพธ์เมื่อใช้สมการ 2.20 ลงคะแนนลงในตารางความสัมพันธ์ระหว่าง r และ θ	15
รูปที่ 2-8 ขั้นตอนย่อยในแต่ละชั้นของโครงข่ายประสาทเทียมแบบคอนโวลูชันนัล	16
รูปที่ 3-1 ภาพถ่ายจากกูเกิลสตรีทวิว	18
รูปที่ 3-2 แผนภูมิสายงานของการตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท.....	19
รูปที่ 3-3 แผนภูมิสายงานขั้นตอนการฝึกการเรียนรู้จำอาคาร.....	23
รูปที่ 3-4 ตัวอย่างภาพผลลัพธ์จากขั้นตอนตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท	24
รูปที่ 4-1 ภาพตัวอย่างเมื่อแปลงภาพสี RGB เป็นขาวดำแบบไล่ระดับ.....	26
รูปที่ 4-2 ฮิสโตแกรมของภาพตัวอย่าง	26
รูปที่ 4-3 ตัวอย่างผลของขั้นตอนการทดลองแยกท้องฟ้าออกจากภาพถ่าย	27
รูปที่ 4-4 ภาพนำเข้า (ก) และผลที่ได้จากการทดลองการกำจัดใบไม้สีเขียวในภาพถ่าย (ข)	28
รูปที่ 4-5 ภาพนำเข้า (ก) และผลที่ได้จากการทดลองแบ่งส่วนภาพบริเวณของวัตถุด้วยสี (ข).....	29
รูปที่ 4-6 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองกำจัดบริเวณที่ไม่ต้องการ (ข).....	30
รูปที่ 4-7 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองหาเส้นตรงยาวในแนวราบ (ข).....	30
รูปที่ 4-8 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองหาเส้นตรงที่ขนานกันในแนวตั้ง (ข)	31

รูปที่ 4-9 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองหาสีเหลืองในภาพถ่าย (ข)	31
รูปที่ 4-10 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองหาจ้วหลังคา (ข).....	32
รูปที่ 4-11 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองตรวจหาสีอาคาร (ข).....	33
รูปที่ 4-12 ความผิดพลาดที่เกิดจากไม่ได้ตรวจสอบขนาดและอัตราส่วนของบริเวณ.....	34
รูปที่ 4-13 รอยต่อของภาพที่เกิดขึ้นจากการอัลกอริทึมการต่อภาพของกูเกิล	35
รูปที่ 4-14 (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์จากการดำเนินการตรวจหาอาคารจากภาพถ่าย....	36
รูปที่ 4-15 แผนภูมิแสดงความถูกต้องที่เพิ่มขึ้นในแต่ละรอบของการฝึก	38
รูปที่ 4-16 แผนภูมิแสดงค่าความถูกต้องของการทดลองกับชุดทดสอบ	39
รูปที่ 4-17 ตัวอย่างภาพที่ตรวจหาผิดพลาดของโครงข่ายประสาทเทียมคอนโวลูชันนัล	42
รูปที่ 4-18 ตัวอย่างภาพที่ตรวจหาไม่ได้เนื่องจากรายละเอียดของอาคารได้น้อยเกินไป	42
รูปที่ ก-1 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง	53
รูปที่ ก-2 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ).....	54
รูปที่ ก-3 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ).....	55
รูปที่ ก-4 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ).....	56
รูปที่ ก-5 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ).....	57
รูปที่ ก-6 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ).....	58
รูปที่ ก-7 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ).....	59
รูปที่ ก-8 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ).....	60
รูปที่ ก-9 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ).....	61
รูปที่ ข-1 ตัวอย่างชื่อภาพจากกูเกิลสตรีทวิวที่นำมาใช้ในการทดลอง.....	63
รูปที่ ข-2 ตัวอย่างการใช้ประโยชน์หลังการประมวลผลภาพดิจิทัลโดยการวาดพิกัดจีพีเอสลงใน แผนที่ของกูเกิล	64
รูปที่ ค-1 หน้าเชื่อมต่อผู้ใช้ของโปรแกรมตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท	66
รูปที่ ค-2 แผงควบคุมโปรแกรมตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท.....	67

รูปที่ ค-3 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์ของขั้นตอน Automatic Seeded Region Growing 68

รูปที่ ค-4 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์ของขั้นตอนการตรวจหาสีของอาคาร..... 68

รูปที่ ค-5 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์ของขั้นตอนการตรวจหาสีเหลี่ยมในภาพถ่าย 69

รูปที่ ค-6 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์ของขั้นตอนการตรวจหาเส้นตรงยาวในแนวนอน 69

รูปที่ ค-7 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์ของขั้นตอนการหาจั่วหลังคาของอาคาร 70

รูปที่ ค-8 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์การหาเส้นตรงที่ขนานกันในแนวตั้ง 70

รูปที่ ง-1 ส่วนเชื่อมต่อผู้ใช้ของโปรแกรมรู้จำอาคารจากภาพถ่ายข้างทางในชนบท 72



บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ประเทศไทยมีการใช้ประโยชน์จากที่ดินหลายประเภท เช่น เป็นที่อยู่อาศัย เป็นพื้นที่ใช้สอย เป็นพื้นที่เลี้ยงสัตว์ เป็นพื้นที่ทำการเกษตร เป็นต้น หน่วยงานต่างๆที่ต้องการข้อมูลการใช้ประโยชน์ที่ดินมีความต้องการทราบว่าแต่ละพื้นที่ใช้ประโยชน์ในด้านไหนบ้าง เพื่อนำผลสำรวจไปใช้ประโยชน์ต่างๆ เช่น การเก็บภาษี การหาการขยายตัวของชุมชน หาพื้นที่อยู่อาศัยในพื้นที่ห่างไกลจากเขตชุมชน เป็นต้น ในการสำรวจพื้นที่ที่มีขนาดใหญ่ นั้น ถ้าทำโดยภาพถ่ายทางอากาศนั้นจะทำให้ระยะเวลาหลายๆครั้งและมีค่าใช้จ่ายสูง หากไม่ใช้วิธีดังกล่าวก็ต้องใช้คนสำรวจซึ่งใช้แรงงานเป็นจำนวนมากและทำได้ช้า หากสำรวจโดยภาพถ่ายบนรถสำรวจเคลื่อนที่ที่มีการบันทึกภาพและเก็บพิกัด GPS จะมีประสิทธิภาพและประหยัดมากกว่า ในปัจจุบันมีหน่วยงานที่ใช้รถสำรวจติดกล้องถ่ายภาพและบันทึกตำแหน่งไปพร้อมๆกัน หากสามารถตรวจหาและรู้จำประเภทการใช้ประโยชน์บนแปลงที่ดินได้จากภาพที่ได้จากรถสำรวจดังกล่าว ก็จะเป็นแหล่งกำเนิดข้อมูลที่มีประโยชน์ต่อหน่วยงานที่เกี่ยวข้อง

การใช้ประโยชน์จากที่ดินเช่นการอยู่อาศัยหรือพื้นที่มีผู้คนเข้าไปใช้ประโยชน์นั้นส่วนหนึ่งสามารถตรวจดูได้จากการมีอยู่ของอาคาร ซึ่งมีส่วนประกอบสำคัญต่างๆของอาคาร เช่น หลังคา เสา หน้าต่าง ประตูและผนังของอาคาร เป็นต้น หากสามารถตรวจหาส่วนประกอบของตัวอาคารจากภาพถ่ายข้างทางในชนบทได้ ก็จะสามารถชี้ได้ว่าในภาพนั้นมีอาคารตั้งอยู่และใช้เป็นข้อมูลเบื้องต้นในการระบุการใช้ที่ดินต่อไป งานวิจัยนี้ใช้ภาพจากรถสำรวจจากภูเกีสตรีทวิดรูปที่ 1-1

งานวิจัยส่วนใหญ่ในอดีตที่ผ่านมามุ่งเน้นไปที่ภาพถ่ายจากทางอากาศ ส่วนงานวิจัยในปัจจุบันที่ใช้ภาพถ่ายภาคพื้นดินโดยมากจะตรวจหาหน้าต่าง ประตูและระนาบของอาคาร (Plane) ซึ่งงานวิจัยเหล่านี้สามารถนำไปใช้ได้ทั้งในประเทศหรือในตัวเมือง แต่สำหรับอาคารที่ถูกบดบังการตรวจหาลักษณะส่วนใดส่วนหนึ่งของตัวอาคารนั้นอาจจะไม่เพียงพอเพื่อใช้ตรวจหาการมีอยู่ของอาคาร งานวิจัยนี้จึงมุ่งเน้นที่จะตรวจหาอาคารที่ปรากฏอยู่ในภาพโดยจะตรวจหาลักษณะต่างๆของตัวอาคารจากภาพถ่ายภาคพื้นดินเพื่อนำมาเข้าสู่วิธีการทางการเรียนรู้ของเครื่อง (Machine Learning) เพื่อพัฒนาวิธีการการตรวจหาอาคารเพื่อเป็นส่วนหนึ่งของข้อมูลพื้นฐานในการสำรวจการใช้ประโยชน์ที่ดินของหน่วยงานที่เกี่ยวข้องต่อไป



รูปที่ 1-1 ตัวอย่างอาคารที่พบได้ทั่วไปในชนบท

1.2 วัตถุประสงค์

เพื่อพัฒนาวิธีการประมวลผลภาพดิจิทัลและโปรแกรมคอมพิวเตอร์เพื่อตรวจหาอาคารที่มีอยู่ในภาพถ่ายข้างทางในชนบท

1.3 ขอบเขตงานวิจัย

การพัฒนาวิธีการประมวลผลข้อมูลภาพและโปรแกรมคอมพิวเตอร์เพื่อหาอาคาร มีขอบเขตดังนี้

- ใช้ภาพถ่ายภาคพื้นดิน แถนของเลนส์กล้องอยู่ในแนวราบโดยประมาณ
- ใช้ภาพถ่ายที่อยู่ในระยะโฟกัสชัดเจน ถ่ายในที่ที่มีแสงสว่างเพียงพอและไม่มีรอยต่อ
- ภาพที่ใช้ทดสอบเป็นภาพถ่ายด้านข้างของถนน
- มนุษย์เป็นผู้ตรวจสอบความถูกต้องของผลลัพธ์
- อาคารที่ตรวจหาในภาพคือส่วนของอาคารที่ไม่ถูกบดบัง

1.4 ขั้นตอนดำเนินงานวิจัย

- 1) ศึกษางานวิจัยและทฤษฎีต่างๆที่ใช้ในการวิเคราะห์และประมวลผลภาพดิจิทัล
- 2) พัฒนาระบบการทางภาพหรือวิธีการที่เหมาะสมเพื่อใช้ในการทำงานวิจัย
- 3) พัฒนาโปรแกรมต้นแบบ
- 4) ทดสอบโปรแกรมกับภาพตัวอย่างและปรับปรุงแก้ไข
- 5) ทดสอบการทำงานกับภาพถ่าย
- 6) ประเมินผล และสรุปงานวิจัย

1.5 ประโยชน์ที่ได้รับ

- 1) ได้กระบวนการและโปรแกรมทางการประมวลผลภาพดิจิทัลเพื่อหาอาคารที่มีอยู่ในภาพถ่ายข้างทางในชนบท
- 2) สามารถนำโปรแกรมต้นแบบนี้มาต่อยอดเพื่อใช้งานจริงต่อไป เพื่อประหยัดเวลาและค่าใช้จ่ายในการสำรวจประเภทที่อยู่อาศัยของหน่วยงานที่เกี่ยวข้องหรือต้องการข้อมูล

บทที่ 2 ทฤษฎีและหลักการที่เกี่ยวข้อง

2.1 ทบทวนงานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องในที่นี่แบ่งเป็น 2 หัวข้อใหญ่ๆคือ การตรวจหาอาคารจากภาพถ่ายภาคพื้นดิน การตรวจหาอาคารจากภาพถ่ายทางอากาศ ซึ่งแต่ละงานล้วนมีวิธีการที่น่าสนใจแตกต่างกันไป

2.1.1 การตรวจหาอาคารจากภาพถ่ายภาคพื้นดิน

งานวิจัยที่เกี่ยวข้องกับการตรวจหาอาคารจากภาพถ่ายภาคพื้นดินจะแบ่งการวิจัยเป็น 2 กลุ่มคือ การตรวจหาอาคารและการตรวจหาส่วนของอาคาร

2.1.1.1 การตรวจหาอาคาร

ปี 2007 Zhizhong และคณะ (Zhizhong Kang, Sisi Zlatanova et al. 2007) ได้เสนอวิธีการตรวจหาขอบเขตของอาคารโดยการตรวจหาช่วงความแปรปรวน (Range Variance) โดยใช้ฮิสโทแกรมของความแตกต่างระหว่างจุดที่สอดคล้องกัน (Corresponding points) เนื่องจากด้านหน้าของอาคารในต่างประเทศจะสร้างติดกัน แต่ก็จะมีช่องว่างระหว่างอาคาร เช่น ซอย ทางแยก เป็นต้น เมื่อพบซอยหรือทางแยกค่าฮิสโทแกรมของความแตกต่างระหว่างจุดที่สอดคล้องกันจะสูงมากกว่าจุดที่เป็นหน้าระนาบของอาคาร จึงทำให้สามารถหาความกว้างของอาคารนั้นๆได้

ปี 2008 Philip David (David 2008) ได้เสนอวิธีการตรวจหาด้านหน้าของอาคารในสภาพแวดล้อมในเมือง ชั้นแรกหาจุดสุดท้ายตา (Vanishing Point) จากนั้นใช้วิธี RANSAC เพื่อหาเส้นตรงที่พอดีกับแต่ละจุดสุดท้ายตา จากนั้นแบ่งและรวมกลุ่มของจุดตัดที่อยู่ในระนาบเดียวกันแล้วอาคารรอบก็จะได้แต่ละระนาบของอาคาร ผลการทดลองของงานวิจัยนี้คือ สามารถหาแต่ละระนาบของอาคาร(Plane)ได้ แต่ก็มีปัญหากับอาคารที่ซับซ้อนมากๆ เช่น ไปเจอเงาสะท้อนของหน้าต่าง ไปเจอระนาบจากด้านในของอาคารและเจอหลังคาของอาคาร เป็นต้น

ปี 2009 Martin และ Wolfgang (Martin Drauschke and Wolfgang Föorstner 2007) ได้เสนอวิธีการเลือกฟีเจอร์ที่เหมาะสมสำหรับตรวจหาอาคารและส่วนของอาคาร (Building parts) เนื่องจากความหลากหลายของอาคารและสภาพแวดล้อม การเลือกฟีเจอร์ที่เหมาะสมเพื่อมาใช้ในการหาอาคารจึงเป็นเรื่องจำเป็น โดยวิจัยนี้ใช้ชุดของฟีเจอร์จำนวนหนึ่งนำมาเข้าขั้นตอนวิธีเอดาบัส (Adaboost algorithm) แล้วนำผลลัพธ์ที่ได้เทียบกับวิธีที่ใช้ฟีเจอร์เดี่ยวทั้งภาพ การวิจัยนี้ได้ทดลองกับรูปหน้าอาคารทั้งหมด 62 ตัวอย่าง ผลลัพธ์ที่ได้ในส่วนใหญ่มิมีความผิดพลาดอยู่ไม่เกินเฉลี่ยร้อยละ

12 ยกเว้น หลังคากับหน้าต่างซึ่งมีค่าความผิดพลาดอยู่ที่เฉลี่ยร้อยละ 48.1 และเฉลี่ยร้อยละ 41.8 ตามลำดับ

2.1.1.2 การตรวจหาส่วนของอาคาร

ปี 2009 Viraj และคณะ (Viraj Kulkarni, Rohan Nagesh et al. 2009) ได้เสนอวิธีการตรวจหาหน้าต่างจากด้านหน้าของอาคารโดยใช้วิธีฉายโปรเจกชันหาแนวของหน้าต่างทั้งแนวตั้งและแนวนอน ซึ่งจะติกรอบเป็นหน้าต่างได้คร่าวๆ จากนั้นใช้ Snake Algorithm เพื่อปรับกรอบให้เข้ากับรูปร่างของหน้าต่าง งานวิจัยนี้ยังมีปัญหากับหน้าต่างที่เรียงตัวไม่เป็นแนวเดียวหรืออาคารที่เป็นกระจก เนื่องจากระยะห่างระหว่างกระจกแต่ละบานน้อยเกินไปและมักมีแสงสะท้อน

ปี 2011 Vincent และ Baillard (Vincent Brandou and Caroline Baillard 2011) ได้เสนอวิธีหาจั่วหลังคาจากภาพถ่ายภาคพื้นดินโดยใช้วิธีกำจัดสิ่งรบกวนเพื่อหาเฉพาะจุดสนใจ เช่น ใช้การแบ่งส่วนสีต้นไม้ การลบเสาอากาศ เป็นต้น จากนั้นหาขอบของหลังคาโดยใช้ Hough Transform โดยหาเส้นที่องศามีความสัมพันธ์กัน โดยวิธีนี้สามารถหาค่าความถูกต้องได้เฉลี่ยร้อยละ 76.32 โดยใช้ภาพตัวอย่างทั้งหมด 158 ภาพ ข้อจำกัดของงานวิจัยนี้คือหากการตรวจหาขอบด้านข้างของหลังคาทำได้ไม่ดี เช่น มีต้นไม้บังหรือมีอาคารหลังอื่นเป็นพื้นหลังหรือหากตรวจหาขอบด้านล่างของหลังคาได้ไม่ถูกต้องก็จะทำให้การทำงานผิดพลาด

ปี 2012 Milos และคณะ (Milos Miljanovic, Thomas Eiter et al. 2012) ได้เสนอวิธีการหาค่าขีดแบ่งที่เหมาะสมและการแบ่งส่วนภาพหน้าต่างออกจากผนังของอาคาร เรียกวิธีการนี้ว่า Histogram Median Filtering (HMF) โดยสามารถหาหน้าต่างบนระนาบของตึกในเวียนนาและซูริชได้ทั้งแนวนอนและหน้าต่างที่เรียงกันในแนวเฉียงได้

2.1.2 การตรวจหาอาคารจากภาพถ่ายทางอากาศ

ปี 2008 Xiuyun และ Yan (Xiuyun Wu and Yan Li 2012) ได้เสนอวิธีพัฒนาการแยกเส้นขอบของอาคาร (Edge Line) จากภาพถ่ายทางอากาศ เนื่องจากในหลายๆครั้งการหาขอบของอาคารทำได้ยากเนื่องจากรูปร่างที่ซับซ้อนและมีสิ่งรบกวน (Noise) มาก การวิจัยนี้จึงได้พัฒนาการหาเส้นขอบของอาคารโดยใช้วิธีการ marker-controlled watershed ร่วมกับข้อมูลความสูงของ Digital Surface Model (DSM) เพื่อลบถนนและเงาในภาพ นำมาใช้รวมกับการหาขอบแบบแคนนี่ (Canny Edge) เพื่อหาเส้นขอบของอาคาร จากนั้นหาเส้นตรงของเส้นขอบ ก็จะได้เส้นขอบของอาคาร

ปี 2010 Xiao และคณะ (J.Xiao, M.Gerke et al. 2010) ได้เสนอวิธีการตรวจหาอาคารที่มีหลังคาสี่เหลี่ยมแบนจากภาพแนวเฉียงหลายๆมุม โดยใช้ภาพจากหลายๆมุมในการแยกเส้นตรงออกเป็นสามมิติ (3D line Extraction) นอกจากนั้นได้เสนอวิธีตรวจหาหลังคาที่ไม่เป็นสี่เหลี่ยมได้โดย

ใช้ฮิสโทแกรมของการกวาดระนาบของอาคาร (Plane sweeping) ซึ่งโดยไปจะวัดแค่ความสูงแต่ในงานวิจัยนี้ได้เสนอให้มีการใช้มุมด้วย

2.2 ทฤษฎีที่เกี่ยวข้อง

ในหัวข้อนี้จะกล่าวถึงทฤษฎีที่เกี่ยวข้องที่จะเป็นประโยชน์ต่อการทำวิจัย ซึ่งส่วนใหญ่จะเกี่ยวข้องกับการประมวลผลภาพดิจิทัล เช่น การหาวัตถุรูปทรงสามเหลี่ยม การลบส่วนของภาพที่ไม่ต้องการ เป็นต้น

2.2.1 ความหมายของอาคาร

ความหมายของอาคารนั้นได้มีการระบุไว้โดยหลายหน่วยงาน ในที่นี่ได้นำนิยามหรือความหมายที่สำคัญอ้างอิงได้มากกล่าวไว้ ดังนี้

2.2.1.1 พระราชบัญญัติควบคุมอาคาร พ.ศ. ๒๕๒๒ มาตรา ๔ (ศาลปกครอง 2522)

“อาคาร” หมายถึง อาคาร บ้าน เรือน โรง ร้าน แพ คลังสินค้า สำนักงานและสิ่งที่สูงขึ้นอย่างอื่น ซึ่งบุคคลอาจเข้าอยู่หรือเข้าใช้สอยได้
คำว่า “บุคคลอาจเข้าใช้สอยได้” จำกัดความเพียงใด?

ประเด็นนี้ คณะกรรมการกฤษฎีกา เคยพิจารณาให้ความเห็นไว้ใน เรื่องเสร็จที่ ๒๘๓/๒๕๓๕ (สำนักงานคณะกรรมการกฤษฎีกา 2522) ว่า :

ถ้อยคำที่ว่า “บุคคลอาจเข้าใช้สอยได้” นั้น มิได้มีความหมายจำกัดเพียงเฉพาะสิ่งก่อสร้างขึ้น เพื่อให้บุคคลเข้าไปใช้สอยภายในสถานที่ที่ก่อสร้างได้เท่านั้น แต่หากยังหมายถึง สิ่งก่อสร้างในลักษณะที่บุคคลสามารถเข้าใช้สอย ได้รับประโยชน์ในกิจการใดกิจการหนึ่ง โดยที่ตัวผู้ใช้สอยไม่จำเป็นต้องเข้าไปภายในสิ่งก่อสร้างนั้นด้วยแต่อย่างใด

2.2.1.2 ราชบัณฑิตยสถาน

ราชบัณฑิตยสถาน (ราชบัณฑิตยสถาน 2542) ได้ระบุไว้ว่า อาคาร หมายถึง เรือน โรง สิ่งก่อสร้างขึ้นที่มีลักษณะคล้ายคลึงเช่นนั้น อาคาร บ้าน เรือน โรง ร้าน แพ คลังสินค้า สำนักงานและสิ่งก่อสร้างขึ้นอย่างอื่น ซึ่งบุคคลอาจเข้าอยู่หรือเข้าใช้สอยได้และหมายความรวมถึงสิ่งก่อสร้างขึ้น อย่างอื่นตามที่กฎหมายกำหนด เช่น อัฒจันทร์ เขื่อน สะพาน อุโมงค์ บ้าย อุเรือ. (ป. ส. อาคาร อาคาร) .

2.2.1.3 กรมโยธาธิการทหารบก

กรมโยธาธิการทหารบก (กรมโยธาธิการทหารบก 2530) ได้ระบุความหมายของอาคารไว้ว่า อาคาร หมายถึง อาคาร, บ้าน, เรือนโรงและอื่น ในสภาพคล้ายคลึงกันอันมุ่งหมายให้เป็นที่พักอาศัยประจำของทหาร ยานพาหนะ สัตว์พาหนะและเป็นที่พักของเครื่องใช้

ในงานวิจัยนี้จะนิยามว่า อาคาร หมายถึง อาคาร บ้าน เรือน โรง ซึ่งบุคคลอาจเข้าอยู่หรือเข้าใช้สอยได้

2.2.2 การแปลงภาพจากภาพสี (RGB) เป็นภาพขาวดำแบบไล่ระดับ

การแปลงภาพจากภาพสี (RGB) เป็นภาพขาวดำแบบไล่ระดับ (Tarun Kumar and Verma 2010) จะแปลงค่าสีจากสามตัวแปรคือ R, G และ B มาเป็นหนึ่งตัวแปร โดยในตัวแปรนี้จะเก็บค่าความเข้มสีระหว่าง 0 ไปจนถึง 255 ค่า 0 จะเป็นค่าความเข้มสีที่มืดที่สุด ค่า 255 จะเป็นค่าความเข้มสีที่สว่างมากที่สุด ดังรูปที่ 2-1



รูปที่ 2-1 ค่าความเข้มของสีจากค่า 0 ไปจนถึง 255

การแปลงค่าสีจากภาพสี RGB ไปเป็นภาพขาวดำแบบไล่ระดับ สามารถทำได้ใช้สมการ 2.1 ดังนี้

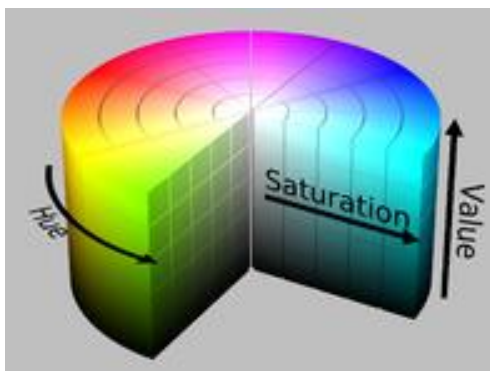
$$Y = 0.21 R + 0.72 G + 0.07 B \quad (2.1)$$

โดยที่ Y เท่ากับค่าความเข้มของสีภาพขาวดำแบบไล่ระดับ

R, G และ B คือค่าสีขาวย สีเขียวและสีน้ำเงินตามลำดับ

2.2.3 การแปลงภาพจากภาพสี (RGB) เป็นแบบจำลองสี HSV

HSV (George H Joblove and Greenberg 1978) คือแบบจำลองสีที่ใช้ลักษณะรูปทรงกระบอกในการเก็บพิกัดสี ซึ่งเป็นที่นิยมใช้มากในการประมวลผลภาพดิจิทัลและการรับภาพของคอมพิวเตอร์ H คือค่าสี โดยค่าสีจะเปลี่ยนไปตามองศาของวงล้อสี (Color wheel) ซึ่งค่าสีจะเริ่มตั้งแต่ 0 องศาหมุนไปจนถึง 360 องศา ที่ 0 องศาจะเป็นค่าสีแดง เมื่อผ่านมาถึง 120 องศาจะเป็นสีเขียว 240 องศาคือสีฟ้า จากที่ 360 องศาจะกลับมาเป็นสีแดงอีกครั้ง S คือค่าความสดของสีโดยมีค่าจาก 0 ไปถึง 1 ค่าความสดยิ่งมากสีสันทันที่ได้จะมีความสดมาก ความสดของสีน้อยก็จะได้สีสันทันน้อยลง ค่าที่ใกล้ 0 ภาพที่ได้จะเป็นภาพขาวดำ (Grayscale) V คือค่าความสว่าง ค่ายิ่งมากยิ่งมีความสว่างมาก ค่าน้อยลงความสว่างก็จะมีลดลง แบบจำลองสี HSV แสดงในภาพที่ 2-2



รูปที่ 2-2 แบบจำลองสี HSV

การแปลงภาพจากแบบจำลองสี RGB เป็น HSV สามารถทำได้โดยแปลงค่า R, G และ B จาก [0,255]ให้อยู่ในช่วง [0,1] คำนวณตามสูตรดังต่อไปนี้

หาค่าสูงสุดโดยใช้สมการ 2.2 และค่าต่ำสุดโดยใช้สมการ 2.3 จากนั้นนำมาหาค่า Chroma โดยใช้สมการ 2.4

$$\max = \max (R,G,B) \quad (2.2)$$

$$\min = \min (R,G,B) \quad (2.3)$$

$$\text{Chroma} = \max - \min \quad (2.4)$$

จากนั้นใช้สมการ 2.5 ในการหาค่าสี ใช้สมการ 2.6 ในการหาค่าความเข้มสีและใช้สมการ 2.7 เพื่อหาค่าความสว่าง

$$H = \begin{cases} 0^\circ, & \text{if Chroma} = 0 \\ 60^\circ \times \frac{G - B}{\text{Chroma}} \bmod 6, & \text{if max} = R \\ 60^\circ \times \frac{B - R}{\text{Chroma}} + 2, & \text{if max} = G \\ 60^\circ \times \frac{R - G}{\text{Chroma}} + 4, & \text{if max} = B \end{cases} \quad (2.5)$$

$$S = \begin{cases} 0, & \text{max} = 0 \\ \frac{\text{Chroma}}{\max}, & \text{max} \neq 0 \end{cases} \quad (2.6)$$

$$V = \max \quad (2.7)$$

2.2.4 การแบ่งส่วนภาพด้วยค่าขีดแบ่ง (Thresholding)

การแบ่งส่วนภาพด้วยค่าขีดแบ่ง (Rafael C. Gonzalez and Richard E. Woods 2009) เป็นวิธีการที่ง่ายที่สุดในการแบ่งส่วนภาพ (Segmentation) การแบ่งส่วนภาพเป็นวิธีการเพื่อแบ่งส่วนใดส่วนหนึ่งของภาพที่สนใจออกมาจากภาพ โดยผลที่ได้จากการแบ่งส่วนภาพด้วยค่าขีดแบ่งจะได้ผลลัพธ์เป็นทวิภาค ค่าที่น้อยกว่าค่าขีดแบ่งจะถูกทำให้กลายเป็น 0 ค่าที่มากกว่าหรือเท่ากับค่าขีดแบ่งจะถูกทำให้กลายเป็น 255 สามารถแสดงเป็นสมการ 2.8 ดังนี้

$$\text{dst}(x,y) = \begin{cases} 0, & \text{if } \text{src}(x,y) < T \\ 255, & \text{otherwise} \end{cases} \quad (2.8)$$

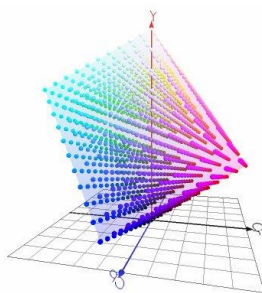
โดย $\text{src}(x,y)$ คือค่าความเข้มของจุดภาพที่พิกัด x และ y ของภาพต้นฉบับ

$\text{dst}(x,y)$ คือค่าความเข้มของจุดภาพที่พิกัด x และ y ของภาพปลายทาง

T คือค่าขีดแบ่ง

2.2.5 การแปลงภาพจากภาพสี (RGB) เป็นแบบจำลองสี $YCbCr$

ภาพถ่ายสีปกติจะเก็บค่าสีแดง สีเขียวและสีฟ้า (RGB) ซึ่งเหมาะแก่การแสดงผลทางจอภาพ แต่ไม่เหมาะที่จะนำมาทำการวิเคราะห์ค่าสีเพราะแบบจำลองสี RGB นั้นจะมีการเก็บค่าความเข้มแสงรวมไว้กับความเข้มสี ในปี 2001 Cheng และคณะ (Cheng, Jiang et al. 2001) ได้ทำการเปรียบเทียบแบบจำลองสีต่างๆ เช่น RGB, YIQ, YUV, normalized RGB, HIS, CIE $L^*a^*b^*$, and CIE $L^*u^*v^*$ เพื่อใช้สำหรับการแยกสีโดยเฉพาะ พบว่าแต่ละแบบจำลองสีมีข้อดีข้อเสียต่างกัน แบบจำลองสี YUV เหมาะแก่การแบ่งส่วนภาพสี เพราะจะมีการแยกระหว่างค่าความเข้มสีและความเข้มแสง ปีค.ศ.1999 Garcia และคณะ(Christophe Garcia and Tziritas 1999) พบว่าค่าความหนาแน่นความเข้มแสง (Y) มีความเบี่ยงเบนน้อยเมื่ออยู่บนแกนความเข้มสี (C_bC_r) ยิ่งพิสูจน์ได้ว่าแบบจำลองสี $YCbCr$ เหมาะแก่การแบ่งส่วนภาพสี แบบจำลองสี $YCbCr$ แสดงในรูปที่ 2-3



รูปที่ 2-3 แบบจำลองสี $YCbCr$

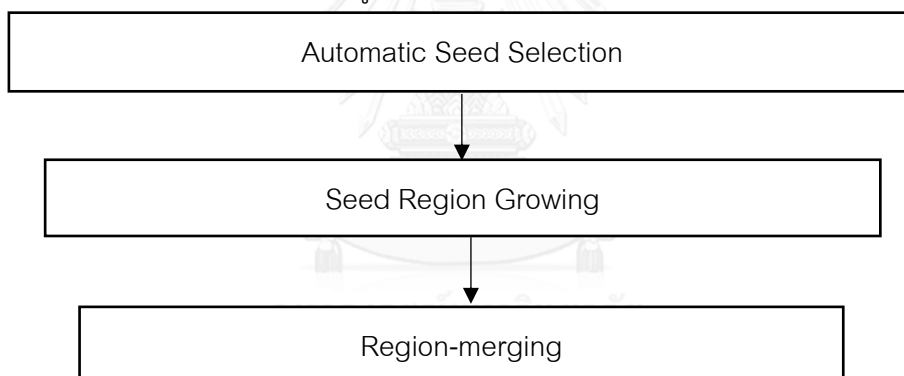
การแปลงจากแบบจำลองสี RGB เป็นแบบจำลองสี YC_bC_r มีสูตรคำนวณ (Colantoni 2012)
 ดังสมการที่ 2.9

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -39.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (2.9)$$

โดย Y คือค่าความสว่าง, C_b คือค่าความสดของสีน้ำเงิน, C_r คือค่าความสดของสีแดง
 R คือ ค่าสีแดง, G คือ ค่าสีเขียว, B คือ ค่าสีน้ำเงิน

2.2.6 Automatic Seeded Region Growing

Seeded region growing ถูกพัฒนาขึ้นครั้งแรกเมื่อปีค.ศ. 1994 โดยอดัมและบิชอป (Adams and Bischof 1994) เพื่อใช้ผสาน (Merge) จุดภาพเข้ากับจุดภาพใกล้เคียง ต่อมาในปีค.ศ. 2001 Fan et al. (Jianping Fan, David K. Y. Yau et al. 2001) ได้พัฒนาวิธีมาประยุกต์ใช้การแบ่งส่วนของสีโดยอัตโนมัติ โดยได้รวมหลักการของการแยกขอบโดยใช้ภาพสีรวมเข้ากับ Seed Region Growing หลักการทำงานดังรูปที่ 2-4



รูปที่ 2-4 แสดงหลักการทำงานของวิธี Automatic Seeded Region Growing

2.2.6.1 การเลือกเมล็ดพันธุ์โดยอัตโนมัติ (The automatic seed selection)

ก่อนจะใช้หลักการ Seeded Region Growing จะต้องมีการวางเมล็ดเริ่มต้น เพื่อเป็นจุดเริ่มต้นในการขยายตัวของบริเวณ (Region) ต่อไป โดยดูจาก 2 เงื่อนไข ได้แก่

จุดภาพเมล็ดพันธุ์ต้องมีความคล้ายคลึงกับจุดภาพข้างเคียงมากกว่าค่าขีดแบ่ง (Threshold) สามารถหาค่าความคล้ายคลึงได้โดยสมการ 2.10

$$\sigma_x = \sqrt{\frac{1}{9} \sum_{i=1}^9 (x_i - \bar{x})^2} \quad (2.10)$$

ที่ x สามารถเป็นได้ทั้งค่าสีจากแกน Y , C_b หรือ C_r
ค่า \bar{x} สามารถหาได้จากสมการ 2.11

$$\bar{x} = \frac{1}{9} \sum_{i=1}^9 x_i. \quad (2.11)$$

นำค่าเบี่ยงเบนมาตรฐานของ Y , C_b และ C_r มารวมกันดังสมการที่ 2.12

$$\sigma = \sigma_Y + \sigma_{C_b} + \sigma_{C_r} \quad (2.12)$$

ทำให้ค่าเบี่ยงเบนมาตรฐานเป็นบรรทัดฐาน $[0,1]$ โดยสมการ 2.13

$$\sigma_N = \sigma / \sigma_{\max} \quad (2.13)$$

ที่ σ_{\max} คือค่าเบี่ยงเบนมาตรฐานที่สูงสุดในภาพ
ความคล้ายคลึงของจุดภาพกับจุดภาพข้างเคียงให้นิยามโดยที่

$$H = 1 - \sigma_N \quad (2.14)$$

จากสมการที่ 2.14 นำ H มาเทียบกับค่าขีดแบ่ง ถ้าพบว่า H ออกมามากกว่าค่าขีดแบ่ง จะได้
ว่าจุดภาพนี้มีโอกาสใช้เป็นเมล็ดพันธุ์เริ่มต้นได้ จากนั้นจะเป็นขั้นตอนการหาระยะทางแบบยุคลิด
จุดภาพเมล็ดเริ่มต้นต้องมีค่าสูงสุดของระยะทางแบบยุคลิดต่อจุดภาพใกล้เคียงทั้ง 8 จุด โดยสมการ
2.15

$$d_i = \frac{\sqrt{(Y - Y_i)^2 + (C_b - C_{b_i})^2 + (C_r - C_{r_i})^2}}{\sqrt{Y^2 + C_b^2 + C_r^2}} \quad (2.15)$$

โดยที่ $i = 1, 2, \dots, 8$

เมื่อระยะทางแบบยุคลิดต่อจุดภาพใกล้เคียงทั้งแปดจุดเรียบร้อยแล้วให้นำมาหาค่าระยะทางที่มาก
ที่สุดสามารถคำนวณได้โดยสมการ 2.16

$$d_{\max} = \max_{i=1}^8 (d_i) \quad (2.16)$$

นำค่าระยะทางแบบยุคลิดที่มากที่สุดมาเปรียบเทียบกับเส้นขีดแบ่ง หากค่าระยะทางมีค่า
น้อยกว่าค่าขีดแบ่ง จะได้ว่าจุดภาพนี้มีโอกาสเป็นเมล็ดพันธุ์ได้ หากค่าความคล้ายคลึงและค่า
ระยะทางแบบยุคลิดที่กล่าวมาข้างต้นตรงตามเงื่อนไขทั้งสองข้อ ให้กำหนดให้จุดภาพนี้คือ เมล็ดพันธุ์
เริ่มต้น ดังรูป 2-5 (b)

2.2.6.2 ทฤษฎีการเติบโตของบริเวณเมล็ดพันธุ์ (Seed region growing)

เมื่อได้ภาพที่มีการวางเมล็ดพันธุ์แล้ว ต่อมาจะเป็นวิธีการทำให้เมล็ดพันธุ์ขยายตัวออก โดยขั้นตอนนี้จะใช้จุดภาพที่ไม่มีการแยกบริเวณและมีจุดภาพข้างเคียงที่มีการกำหนดบริเวณแล้วอย่างน้อย 1 จุด นำมาหาระยะทางแบบยูคลิดระหว่างจุดภาพกับบริเวณที่ติดกัน สามารถคำนวณได้ดังสมการ 2.17

$$d_i = \frac{\sqrt{(Y_i - \bar{Y})^2 + (C_{b_i} - \bar{C}_b)^2 + (C_{r_i} - \bar{C}_r)^2}}{\sqrt{Y_i^2 + C_{b_i}^2 + C_{r_i}^2}} \quad (2.17)$$

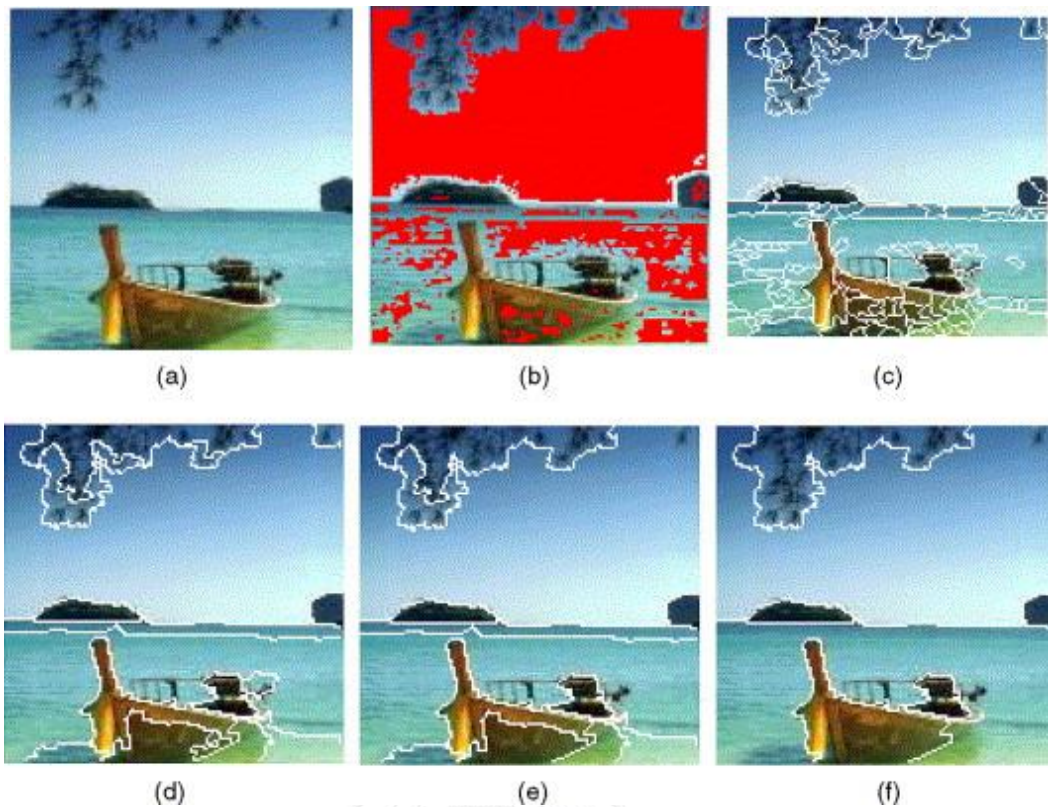
โดยที่ $(\bar{Y}, \bar{C}_b, \bar{C}_r)$ คือค่าเฉลี่ยของแบบจำลอง Y, C_b, C_r ในบริเวณข้างเคียง

จากนั้นนำค่าระยะทางทั้งหมดที่ได้จากระยะทางแบบยูคลิดมาเก็บไว้ในรายการที่เรียงลำดับจากมากไปน้อย มาตรวจสอบจุดภาพข้างเคียง 4 จุด ถ้าจุดภาพข้างเคียงทั้ง 4 จุดอยู่ในบริเวณ (Region) เดียวกัน กำหนดให้จุดภาพมีค่าเหมือนกับจุดภาพข้างเคียง หากจุดภาพข้างเคียงทั้ง 4 จุดอยู่ในบริเวณแตกต่างกัน ให้ใช้วิธีหาระยะทางแบบยูคลิดหาระยะทางระหว่างจุดภาพถึงจุดภาพข้างเคียง จากนั้นกำหนดค่าระยะทางที่น้อยที่สุดให้แก่จุดภาพ สุดท้ายจึงทำการปรับค่าเฉลี่ยของบริเวณนี้ให้เป็นปัจจุบันและเพิ่มบริเวณข้างเคียงที่ยังไม่ได้แยกบริเวณและยังไม่อยู่ในรายการเข้าสู่รายการที่เรียงลำดับจากมากไปน้อย ผลลัพธ์ที่ได้จากขั้นตอน 2.2.5.2 นี้ จะได้บริเวณที่ขยายกว้างขึ้นกว่าขั้นตอน 2.2.5.1 ดังรูป 2-5 (C)

2.2.6.3 ขั้นตอนการผสานบริเวณเข้าด้วยกัน (Region-merging)

จากสองขั้นตอนที่ผ่านมา มีความเป็นไปได้ที่จะเกิดการแบ่งบริเวณมากเกินไปเนื่องจากการหาระยะทางระหว่างจุดภาพถึงจุดภาพบริเวณข้างเคียง ซึ่งขั้นตอนนี้จะทำการเทียบระยะทางระหว่างค่าเฉลี่ยบริเวณถึงค่าเฉลี่ยบริเวณข้างเคียง วิธีนี้จะทำให้ขยายบริเวณออกไปได้อีก สามารถคำนวณได้โดยสมการ 2.18

$$d(R_i, R_j) = \frac{\sqrt{(\bar{Y}_i - \bar{Y}_j)^2 + (\bar{C}_{b_i} - \bar{C}_{b_j})^2 + (\bar{C}_{r_i} - \bar{C}_{r_j})^2}}{\min\left(\sqrt{\bar{Y}_i^2 + \bar{C}_{b_i}^2 + \bar{C}_{r_i}^2}, \sqrt{\bar{Y}_j^2 + \bar{C}_{b_j}^2 + \bar{C}_{r_j}^2}\right)} \quad (2.18)$$



รูปที่ 2-5 ตัวอย่างผลลัพธ์แต่ละขั้นตอนของวิธีการ *Automatic Seeded Region Growing*
 (a) ภาพต้นฉบับ (b) ผลลัพธ์การวางเมล็ดเริ่มต้น (*Initial Seeds*) (c) ผลลัพธ์ของ *Seeded Region Growing* (d) ผลลัพธ์ของบริเวณข้างเคียงโดยความสัมพันธ์ของ *Euclidean* ระยะทางน้อยกว่า 0.1 (e) ผลลัพธ์ของการรวมบริเวณที่ขนาดน้อยกว่า $1/150$ ของภาพ (f) ผลลัพธ์หลังประมวลผล

2.2.7 การตรวจหาเส้นตรงด้วยวิธีการแปลงฮัฟ (Line Hough Transform)

วิธีการแปลงฮัฟ (Gonzalez R.C. and R.E) คือวิธีการทางการประมวลผลภาพดิจิทัลเพื่อใช้หารูปทรงใดๆในรูปภาพ วิธีการหลักของทฤษฎีนี้ จะใช้การลงคะแนน (Voting) ในที่นี้จะกล่าวถึงทฤษฎีการหาเส้นตรงจากรูปภาพด้วยวิธีการแปลงฮัฟ (Duda and Hart 1972) สมการที่ใช้ในการหาเส้นตรงปกติตั้งสมการ 2.19

$$y = mx+c \quad (2.19)$$

โดยที่ m คือ ความชัน

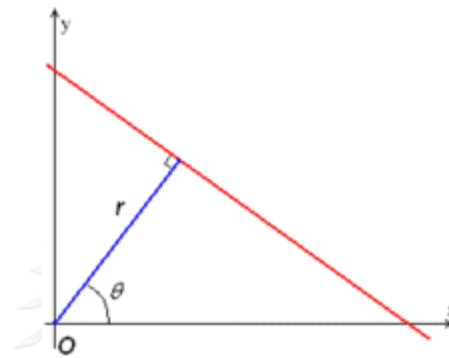
c คือ จุดตัดแกนแนวราบ (y -intercept)

ในที่นี้พบว่าสมการนี้ เมื่อนำมาใช้กับการประมวลผลภาพดิจิทัลจะยังมีประสิทธิภาพไม่เพียงพอเนื่องจากมีปัญหาเมื่อหาเส้นแนวตั้งจะทำให้ค่า m เป็นอนันต์ ภายหลังจาก Shapiro และ Stockman จึงได้นำเสนอสมการใหม่ดังสมการ 2.20

$$x \cos \theta + y \sin \theta = r \quad (2.20)$$

โดยที่ r คือ ระยะห่างระหว่างเส้นและจุดศูนย์กลาง (origin)

θ คือ มุมที่ทำกับแกนแนวนอน



รูปที่ 2-6 รูปแสดงตัวแปรสำหรับสมการ 2.20

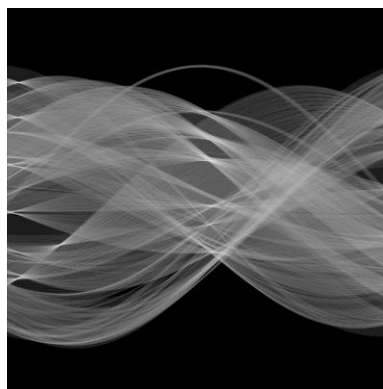
วิธีการนำมาใช้

1. สร้างตารางความสัมพันธ์ระหว่าง r และ θ โดย θ จะมีค่าอยู่ในช่วง -90 ถึง 89 และ ρ จะมีค่าเท่ากับระยะทางมากที่สุดเท่าที่เป็นไปได้

ตารางที่ 2.1 ตารางแสดงความสัมพันธ์ระหว่าง r และ θ

r, θ	-90	-89	88	89
-Dmax					
...					
Dmax					

2. เมื่อพบจุดใดๆบนรูปภาพ (x,y) นำสมการที่ 2.20 มาใช้โดยแทนค่า θ ตั้งแต่ 0 องศาถึง 359 จากนั้นนำมาลงคะแนนในตาราง 2.1
3. เมื่อทำทั้งรูปจนสำเร็จแล้ว หาจุดที่ค่า r และ θ มีค่ามากที่สุดก็จะสามารถหาเส้นตรงได้ เมื่อนำค่าในตารางแสดงความสัมพันธ์ระหว่าง r และ θ มาวาดเป็นภาพจะได้ดังรูปตัวอย่างที่ 2-7



รูปที่ 2-7 ตัวอย่างผลลัพธ์เมื่อใช้สมการ 2.20 ลงคะแนนลงในตารางความสัมพันธ์ระหว่าง r และ θ

(R. Haralick and L. Shapiro 1992)

2.2.8 การเพิ่มจำนวนของภาพในชุดของการฝึก

การเพิ่มจำนวนของภาพในชุดของการฝึก (Lecun) เพื่อเพิ่มความหลากหลายในการรู้จำ เนื่องจากชุดของการฝึกที่น้อยหรือมีความซ้ำซ้อนมากเกินไปจะทำให้เกิดปัญหาโครงข่ายประสาทเทียมจดจำชุดการฝึกมากเกินไป (Overfitting) ซึ่งส่งผลทำให้เกิดความผิดพลาดในการนำโครงข่ายประสาทเทียมไปทดสอบกับข้อมูลชุดอื่น ซึ่งการเพิ่มจำนวนชุดการฝึกสามารถทำได้หลายวิธี เช่น การหมุนภาพ (rotation) , การกลับภาพ (flip) , การใส่สิ่งรบกวนลงไปในภาพ (noise) เป็นต้น ซึ่งภาพผลลัพธ์ที่ได้จากการกลับภาพในแนวนอนจะช่วยเพิ่มความหลากหลายให้กับชุดการฝึกซึ่งมีส่วนลดโอกาสในการเกิดปัญหาโครงข่ายประสาทเทียมจดจำชุดการฝึกมากเกินไป

2.2.9 การทำให้ข้อมูลในแต่ละแกนแนลเป็นบรรทัดฐานเดียวกัน (Data normalization)

ข้อมูลที่ได้จากภาพในธรรมชาติมีความหลากหลาย จึงจำเป็นต้องทำการปรับข้อมูลให้อยู่ในบรรทัดฐานเดียวกัน (Rafael C. González and Woods 2007) การปรับข้อมูลให้อยู่ในบรรทัดฐานเดียวกันสามารถทำได้โดยสมการ 2.21

$$Intensity_i(x, y) = \frac{Intensity_i(x, y) - \mu}{std} \quad (2.21)$$

โดยที่ i คือ ชุดของข้อมูลแกนแนล Y, U และ V

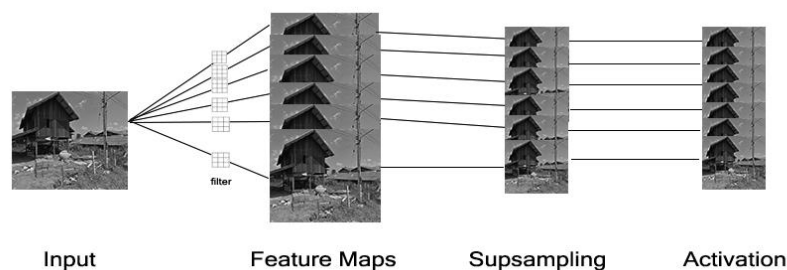
Intensity คือ ค่าสี

μ คือ ค่าเฉลี่ย

std คือ ค่าเบี่ยงเบนมาตรฐาน

2.2.10 โครงข่ายประสาทเทียมแบบคอนโวลูชันนัล

โครงข่ายประสาทเทียมแบบคอนโวลูชันนัล (Alex Krizhevsky, Ilya Sutskever et al. 2012, LISALab 2015) เป็นโครงข่ายประสาทเทียมชนิดหนึ่งซึ่งเป็นที่นิยมมากในการรู้จำรูปภาพและวีดิทัศน์ โครงข่ายประสาทเทียมนี้จำลองมาจากการทำงานของสมองมนุษย์ โครงข่ายประสาทเทียมแบบคอนโวลูชันนัลเป็นโครงข่ายประสาทเทียมที่มีการรู้จำแบบต้องมีการสอน (Supervised Learning) กล่าวคือจะต้องมีการสอนโครงข่ายประสาทเทียมในขั้นตอนการฝึกก่อน โครงข่ายประสาทเทียมจึงจะรู้จำลักษณะเฉพาะของวัตถุต่างๆได้ โครงข่ายประสาทเทียมแบบคอนโวลูชันนัลมีส่วนประกอบหลายชั้นทำงานร่วมกันเป็นโครงข่ายใหญ่ ข้อมูลนำเข้าจะใช้ค่าพิกเซลโดยตรงโดยไม่ต้องผ่านกระบวนการหาพีเจอร์ เนื่องจากสามารถฝึกโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลให้หาพีเจอร์แบบอัตโนมัติได้ ในแต่ละชั้น (layer) ของโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลจะประกอบด้วยชั้นตอนย่อย ดังรูปที่ 2-8



รูปที่ 2-8 ชั้นตอนย่อยในแต่ละชั้นของโครงข่ายประสาทเทียมแบบคอนโวลูชันนัล

ชั้นตอนย่อยในแต่ละชั้น(layer)ของโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลสามารถอธิบายได้เป็น 3 ชั้นตอน ดังนี้

2.2.10.1 ชั้นคอนโวลูชัน

ชั้นคอนโวลูชันมีหน้าที่หลักคือการแยกพีเจอร์ของวัตถุ โดยค่าคอนโวลูชันของแต่ละเคอร์เนลได้รับมาจากการฝึก โดยแต่ละเคอร์เนลจะมีหน้าที่แยกพีเจอร์ต่างกัน คอนโวลูชันชั้นแรกจะได้รับพีเจอร์ระดับต่ำ เช่น ขอบ เส้นและมุม เป็นต้น คอนโวลูชันชั้นถัดไปจะได้รับพีเจอร์ระดับสูงขึ้น เช่น การรวมกันของเส้นหลายเส้น เป็นต้น

2.2.10.2 ชั้น ReLU

ReLU (Rectified linear units) คือชั้นฟังก์ชันกระตุ้น (Activation function) มีหน้าที่ตัดสินใจเพื่อส่งข้อมูลไปสู่เซลล์ประสาทเทียมอื่นๆ ดังสมการ 2.22

$$f(x) = \max(0, x) \quad (2.22)$$

2.2.10.3 ชั้น Pooling

ชั้นตอนนี้มีวัตถุประสงค์เพื่อลดจำนวนข้อมูลให้น้อยลง (subsampling) เพื่อลดความแปรปรวนของข้อมูลและลดจำนวนของโนดในชั้นซ่อน (Hidden layer) โดยดำเนินการหาค่ามากที่สุดหรือค่าเฉลี่ยในแต่ละพื้นที่โดยไม่ทับซ้อนกัน หลังจากดำเนินการชั้นตอนนี้ พีเจอร์ภาพถ่ายขนาดใหญ่จะถูกทำให้เล็กลงซึ่งจำเป็นมากในการตรวจหาวัตถุและการรู้จำ

2.2.10.4 ชั้น Dropout

ชั้นตอนนี้มีวัตถุประสงค์เพื่อเพิ่มความยากในขั้นตอนการฝึกเพื่อลดโอกาสเกิดโมเดลจดจำรูปแบบการฝึกมากเกินไป โดยวิธีการนี้จะสุ่มค่าความน่าจะเป็นระหว่าง 0.5-1 ในนิวรอนแต่ละตัว นิวรอนที่ถูกคัดออกจะไม่ส่งต่อข้อมูลไปสู่ชั้นถัดไป วิธีนี้สามารถเพิ่มความเร็วให้กับขั้นตอนการฝึกอีกด้วย



บทที่ 3

แนวคิดและกระบวนการทำงาน

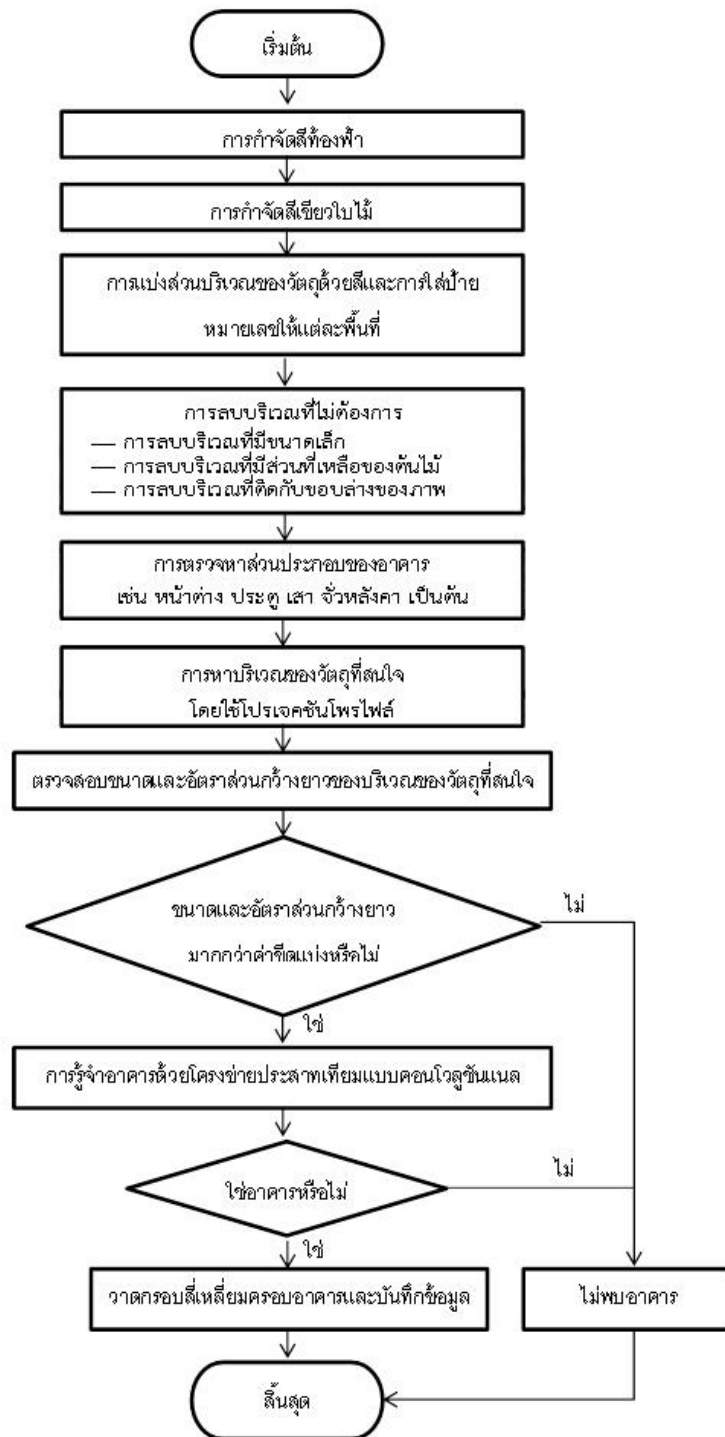
งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการประมวลผลภาพดิจิทัลและโปรแกรมคอมพิวเตอร์ เพื่อตรวจหาอาคารที่มีอยู่ในภาพถ่ายข้างทางในชนบทด้วยอัลกอริทึมทางการประมวลผลภาพดิจิทัล โดยใช้ภาพถ่ายจากกูเกิลสตรีทวิวซึ่งใช้กล้อง Ladybug[®] 2 ในการถ่ายภาพ ภาพที่ได้มีความละเอียด 640 พิกเซล x 600 พิกเซล

เนื่องจากภาพที่ใช้ในงานวิจัยนี้ใช้ภาพถ่ายจากพื้นที่ชนบทของประเทศไทย ซึ่งส่วนมากจะมีสิ่งกีดขวาง เช่น ต้นไม้ เป็นต้น ซึ่งอาจส่งผลให้การประมวลผลภาพผิดพลาดได้ จึงจำเป็นต้องมีกระบวนการทำงานเพื่อกำจัดสิ่งกีดขวางหรือสิ่งที่ไม่จำเป็นในการประมวลผลการตรวจหาอาคารออกจากภาพถ่ายก่อน จึงจะเข้าสู่ขั้นตอนการสกัดชิ้นส่วนของอาคาร เช่น หน้าต่าง ประตู เสา จั่ว หลังคา คาน เป็นต้น ขั้นตอนสุดท้ายคือการรู้จำอาคารด้วยโครงข่ายประสาทเทียมเพื่อระบุอาคารตัวอย่างภาพถ่ายจากกูเกิลสตรีทวิวที่ใช้ในงานวิจัยนี้ ดังรูปที่ 3-1



รูปที่ 3-1 ภาพถ่ายจากกูเกิลสตรีทวิว

หลักการการทำงานของระบบตรวจหาอาคารจากภาพถ่ายด้วยการประมวลผลภาพดิจิทัล สามารถแสดงได้ในแผนภูมิสายงาน (Flow chart) ดังรูปที่ 3-2



รูปที่ 3-2 แผนภูมิสายงานของการตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท

จากแผนภูมิสายงานรูปที่ 3-2 สามารถอธิบายการทำงานขั้นตอนต่างๆได้ดังนี้

3.1 การกำจัดสีท้องฟ้า

เมื่อแปลงภาพสี RGB เป็นภาพขาวดำแบบไล่ระดับตามทฤษฎีในหัวข้อ 2.2.2 แล้วส่วนของภาพบริเวณท้องฟ้าซึ่งมีค่าสีสว่างมากกว่าวัตถุอื่นๆ เช่น ต้นไม้ อาคาร เป็นต้น เพราะฉะนั้นการกำจัดบริเวณท้องฟ้าในภาพถ่ายสามารถทำได้โดยกำหนดค่าขีดแบ่ง เพื่อแบ่งระหว่างวัตถุอื่นๆกับท้องฟ้าออกจากกัน โดยวิธีนี้จะหาขอบเขตของท้องฟ้าโดยวาดฮิสโทแกรมจากบนลงล่าง จากนั้นหาค่าของฮิสโทแกรมของท้องฟ้าน้อยกว่าค่าขีดแบ่งจะถือเป็นจุดสิ้นสุดของท้องฟ้า

3.2 การกำจัดสีเขียวใบไม้

การกำจัดสีเขียวใบไม้มีความหลากหลายของสีเขียวตั้งแต่สีเขียวอ่อนไปจนถึงสีเขียวแก่ ในงานวิจัยนี้ได้แปลงภาพสี RGB เป็นแบบจำลองสี HSV ตามทฤษฎีในหัวข้อ 2.2.3 ซึ่งข้อดีของแบบจำลองสี HSV คือ แบบจำลองสีนี้ได้แยกค่าสี ค่าความสดของสีและค่าความสว่าง ทำให้ง่ายต่อการจำกัดช่วงของสีเขียวของพืช

3.3 การแบ่งส่วนภาพบริเวณของวัตถุด้วยสีและการใส่ป้ายหมายเลขให้แก่ละบริเวณ

ขั้นตอนนี้เป็นการแยกบริเวณต่างๆในภาพโดยใช้สีเป็นเกณฑ์ในการแยกโดยใช้วิธี Automatic seeded region growing ตามทฤษฎีในหัวข้อ 2.2.6 สีที่มีค่าสีใกล้เคียงกันจะถูกรวมไว้ด้วยกัน บริเวณที่เล็กกว่าค่าขีดแบ่งจะถูกรวมเข้ากับบริเวณที่ค่าเฉลี่ยของสีใกล้เคียงที่สุด จากนั้นใส่ป้ายหมายเลขให้แก่ละบริเวณเพื่อให้ง่ายในการใช้งานต่อไป

3.4 การลบบริเวณที่ไม่ต้องการ

จากขั้นตอนข้างต้น อาทิ การลบสีท้องฟ้าและการลบสีเขียวใบไม้ ยังคงเหลือบริเวณที่ไม่ต้องการในภาพ เช่น บริเวณที่มีขนาดเล็ก เศษชิ้นส่วนของต้นไม้ที่เหลือจากขั้นตอนการลบสีเขียวใบไม้ พื้นดิน ถนน เป็นต้น ในขั้นตอนนี้จะใช้บริเวณที่แบ่งส่วนภาพแล้วจากขั้นตอน 3.3 มาใช้ประโยชน์ต่อ ซึ่งวิธีการแบ่งเป็น 3 ข้อดังนี้

3.4.1 การลบบริเวณที่มีขนาดเล็ก

เนื่องจากบริเวณที่มีขนาดเล็กส่วนมากจะมีขนาดเล็กเกินไปที่จะนำมาตรวจหาอาคารและมีโอกาสที่จะสร้างความผิดพลาดให้เกิดขึ้นขั้นตอนการตรวจหาอาคารได้ ขั้นตอนนี้จะทำการกำจัดบริเวณที่มีขนาดเล็กกว่าค่าขีดแบ่ง โดยจะใช้ทฤษฎีการแบ่งส่วนภาพในหัวข้อ 2.2.4

3.4.2 การลบบริเวณที่มีส่วนที่เหลือของต้นไม้

จากขั้นตอนการลบต้นไม้ในหัวข้อ 3.2 จะยังคงมีบางส่วนของใบไม้ที่หลงเหลืออยู่เนื่องมาจากค่าสีเขียวที่อยู่นอกเหนือช่วง HSV ที่ตั้งไว้ แนวคิดของขั้นตอนนี้คือหาบริเวณที่ได้จากการแบ่งส่วนภาพในหัวข้อ 3.3 ที่สีของต้นไม้ส่วนมากถูกลบไปแล้วแต่ยังคงเหลือสีเขียวเหลือเป็นเศษเล็กเศษน้อยอยู่บางส่วน กล่าวคือบริเวณที่มีพิกเซลว่างเปล่ามากกว่าเฉลี่ยร้อยละ 30

3.4.3 การลบบริเวณที่ติดกับขอบล่างของภาพ

ในงานวิจัยนี้ใช้ภาพถ่ายข้างทางในชนบทจากรถสำรวจ บริเวณด้านล่างของภาพถ่ายส่วนมากจะเป็นขอบถนน พื้นดินและพื้นหญ้า ขั้นตอนนี้จะทำการลบบริเวณที่ได้จากการแบ่งส่วนภาพในหัวข้อ 3.3 ที่ติดกับขอบล่างของภาพเพื่อลบขอบถนน พื้นดินและพื้นหญ้าหญ้า

3.5 การตรวจสอบหาส่วนประกอบของอาคาร

ขั้นตอนนี้เป็นการเป็นการตรวจสอบหาส่วนประกอบของอาคารจากภาพถ่าย โดยอาศัยลักษณะทางเรขาคณิตของโครงสร้างอาคารและสีของอาคาร แบ่งเป็น 2 ขั้นตอน ดังนี้

3.5.1 การตรวจสอบหาเส้นตรง

โครงสร้างของอาคารส่วนใหญ่จะเป็นรูปทรงสี่เหลี่ยม ในภาพถ่ายภาคพื้นดินส่วนมากจะมีสิ่งกีดขวางตามธรรมชาติ เช่น ต้นไม้ มาบดบังโครงสร้างของอาคาร ทำให้เป็นไปได้ยากที่จะหาโครงสร้างที่สมบูรณ์ของอาคาร ขั้นตอนนี้ใช้ประโยชน์จากฮัฟฟทรานฟอร์มในทฤษฎีหัวข้อ 2.2.7 เพื่อตรวจสอบหาเส้นตรงในภาพถ่าย การตรวจสอบหาเส้นตรงเพื่อหาส่วนของอาคาร แบ่งเป็น 4 ข้อดังนี้

3.5.1.1 การหาเส้นตรงยาวในแนวราบ

โครงสร้างอาคารมักจะมีคานหรือส่วนของอาคารซึ่งมีขอบเป็นเส้นตรงในแนวราบ งานวิจัยนี้ใช้ประโยชน์ของลักษณะดังกล่าวของอาคารที่ขนานกับพื้นซึ่งมักมีส่วนของขอบที่เป็นเส้นตรงโดยการหาเส้นตรงด้วยฮัฟฟทรานฟอร์ม

3.5.1.2 การหาเส้นตรงที่ขนานกันในแนวตั้ง

โครงสร้างของอาคารมักจะมีเสา ซึ่งอยู่บริเวณมุมของอาคาร ภาพที่ได้จากด้านหน้าของอาคาร จะเห็นเป็นเส้นตรงแนวตั้งที่ขอบทั้งสองฝั่งของอาคาร งานวิจัยนี้จึงใช้ประโยชน์จากเสาซึ่งมีลักษณะตั้งฉากกับพื้นโดยฮัฟฟทรานฟอร์มหาเส้นตรงที่อยู่ในแนวตั้ง จากนั้นจับคู่เส้นตรงโดยทั้งสองเส้นจะต้องขนานกันและอยู่ในแนวระดับเดียวกัน

3.5.1.3 การหาสีเหลี่ยม

ด้านหน้าของอาคารในชนบทส่วนใหญ่จะมีหน้าต่างและประตูซึ่งมีลักษณะเป็นสีเหลี่ยม งานวิจัยนี้จึงใช้ประโยชน์จากสีเหลี่ยมที่ได้จากประตูและหน้าต่างเหล่านี้ในการระบุการมีอยู่ของอาคาร

3.5.1.4 การหาจั่วหลังคา

หลังคาของอาคารหากเป็นลักษณะจั่ว ก็สามารถใช้เป็นพีเจอร์หนึ่งในการตรวจหาอาคารได้ จั่วของหลังคาสามารถตรวจหาด้วยเทคนิคประมวลผลภาพดิจิทัลได้โดยใช้ฮัฟทรานสฟอร์ม แนวคิดเริ่มจากการตรวจหาเส้นตรง จากนั้นหาเส้นตรงที่ทำมุมแหลมชี้ขึ้นข้างบน

3.5.2 การตรวจหาสีอาคาร

สีของอาคารหากเป็นสีที่เด่นชัดมาจากสีทั่วไปในธรรมชาติ เช่น แดง น้ำเงิน น้ำตาล เหลืองและขาว ก็สามารถช่วยในการตรวจหาอาคารได้ สีแดงและน้ำเงินส่วนมากจะเป็นสีของหลังคา สีเหลือง น้ำตาลและขาวส่วนมากจะเป็นสีของตัวอาคาร ขั้นตอนนี้ใช้แบบจำลองสี HSV ตามทฤษฎีในหัวข้อ 2.2.3 ในการตรวจหาสีของอาคาร

3.6 การหาบริเวณของวัตถุที่สนใจ

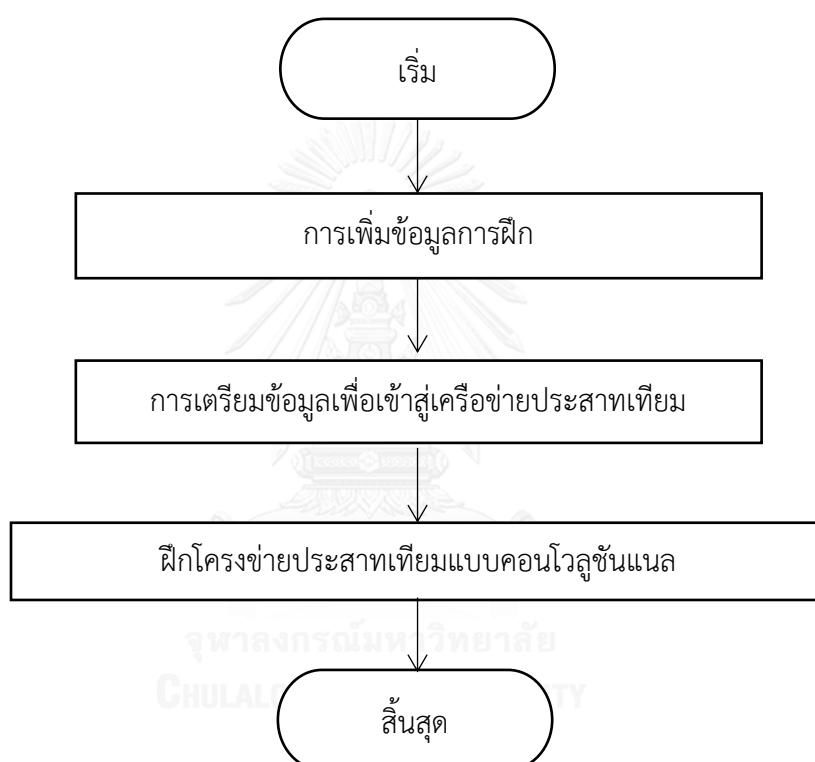
หลังจากขั้นตอนหาชิ้นส่วนของอาคารในหัวข้อ 3.5 แล้ว นำผลที่ได้จากขั้นตอนการหาเส้นตรงในแนวราบในหัวข้อ 3.5.1.1 การหาเส้นตรงที่ขนานกันในแนวตั้งในหัวข้อ 3.5.1.2 การหาสีเหลี่ยมในหัวข้อ 3.5.1.3 การหาจั่วหลังคาในหัวข้อ 3.5.1.4 และการตรวจหาสีอาคารในหัวข้อ 3.5.2 มาซ้อนทับกันเป็นภาพเดียว จากนั้นใช้วิธีการโปรเจกชัน (Projection profile) ในการหาบริเวณที่มีชิ้นส่วนของอาคารที่ตรวจหาพบในขั้นตอน 3.5 อยู่รวมกันหนาแน่น โดยการโปรเจกชันจะทำทั้งแนวตั้งและแนวนอน วิธีการนี้สามารถหาผลลัพธ์ได้มากกว่าหนึ่งบริเวณของวัตถุที่สนใจ

3.7 การตรวจสอบขนาดและอัตราส่วนกว้างยาว

จากขั้นตอนการหาวัตถุที่สนใจในขั้นตอน 3.6 อาจทำให้มีสิ่งแปลกปลอมที่ไม่ใช่อาคารปะปนมาได้ การตรวจสอบเบื้องต้นสามารถทำได้โดยวัดขนาดบริเวณ หากเล็กกว่าค่าขีดแบ่งจะถือว่าเป็นสิ่งรบกวน (noise) อีกวิธีหนึ่งได้แก่การหาอัตราส่วนของบริเวณวัตถุ (Aspect ratio) หากรูปแบบอัตราส่วนบริเวณของวัตถุไม่สมเหตุสมผลต่อการเป็นอาคาร เช่น ความสูงมากกว่าความกว้างมากเกินไป ความกว้างกว้างกว่าความสูงมากเกินไป เป็นต้น

3.8 การรู้จำอาคารด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันแนล

ขั้นตอนนี้มีจุดประสงค์เพื่อยืนยันว่าบริเวณที่ได้จากขั้นตอนการหาอาคารด้วยวิธีการประมวลผลภาพดิจิทัลเป็นอาคารจริงหรือไม่ โดยจะใช้ระบบโครงข่ายประสาทเทียมแบบคอนโวลูชันแนลในทฤษฎีหัวข้อ 2.2.10 ช่วยในการยืนยันความถูกต้อง โดยข้อมูลของบริเวณจะถูกนำมาทำให้ข้อมูลในแต่ละแชนแนลเป็นบรรทัดฐานเดียวกัน (Data normalization) จากนั้นนำข้อมูลที่ได้มาเข้าโครงข่ายประสาทเทียมเพื่อตรวจสอบว่าภาพที่ได้ใช่อาคารจริงหรือไม่ โดยขั้นตอนการฝึกแสดงในแผนภูมิสายงาน ดังนี้



รูปที่ 3-3 แผนภูมิสายงานขั้นตอนการฝึกการรู้จำอาคาร

3.8.1 การเพิ่มข้อมูลการฝึก

การฝึกโครงข่ายประสาทเทียมที่มีประสิทธิภาพควรฝึกกับข้อมูลที่หลากหลายเพื่อลดโอกาสเกิดปัญหาโครงข่ายประสาทเทียมจดจำชุดการฝึกมากเกินไป (Overfitting) ซึ่งส่งผลทำให้เกิดความผิดพลาดในการนำโครงข่ายประสาทเทียมไปทดสอบกับข้อมูลชุดอื่น ซึ่งการเพิ่มจำนวนชุดการฝึกในงานวิจัยนี้ใช้วิธีการการกลับภาพในแนวนอน (Horizontal Flip)

3.8.2 การเตรียมข้อมูลเพื่อเข้าสู่โครงข่ายประสาทเทียม

การเตรียมข้อมูลเพื่อเข้าสู่โครงข่ายประสาทเทียมจะดำเนินการปรับข้อมูลให้อยู่ในบรรทัดฐานเดียวกัน โดยใช้การปรับข้อมูลให้อยู่ในบรรทัดฐานเดียวกันโดยค่าเฉลี่ย (Mean normalization) จะใช้เพื่อการปรับข้อมูลทั้งหมดในชุดการฝึกอยู่ในช่วงเดียวกันและการทำให้ค่าความสว่างอยู่ในบรรทัดฐานเดียวกันจะถูกใช้เพื่อทำให้ค่าความสว่างของภาพเสมอกัน

3.8.3 การฝึกโครงข่ายประสาทเทียมโครงข่ายประสาทเทียมแบบคอนโวลูชันนัล

โครงข่ายประสาทเทียมแบบคอนโวลูชันนัลเป็นโครงข่ายประสาทเทียมแบบต้องใช้ในการสอนเพื่อรู้จำ ในขั้นตอนนี้นี้จึงได้นำชุดฝึกมาเพื่อสอนให้โครงข่ายประสาทเทียมรู้จำอาคารในภาพถ่ายข้างทางในชนบท

หลังจากฝึกโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลแล้ว โครงข่ายประสาทเทียมนี้จะมีประโยชน์อย่างมากในการรู้จำอาคารเพื่อระบุว่าผลลัพธ์ที่ได้ใช่อาคารจริงหรือไม่ ภาพตัวอย่างของภาพที่จะนำเข้าสู่โครงข่ายประสาทเทียมจะเป็นภาพที่ครอบตัดส่วนภาพ (Crop) เฉพาะในสี่เหลี่ยมสีแดงเท่านั้น ดังแสดงในรูปที่ 3-4



รูปที่ 3-4 ตัวอย่างภาพผลลัพธ์จากขั้นตอนตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท

บทที่ 4

การทดลองและผลการทดลอง

4.1 ภาพที่ใช้ในการทดลอง

ภาพที่ใช้ในการทดลองเป็นภาพที่ได้จากกูเกิลสตรีทวิว โดยใช้กูเกิลสตรีทวิวเอพีไอเป็นเครื่องมือในการดึงภาพออกมา ความละเอียดของภาพ 640 x 600 พิกเซล รูปแบบของไฟล์ภาพคือเจเพ็ก (JPEG) มุมของกล้องในการถ่ายภาพเป็นมุม 0 องศาจากแนวระดับ (มุมขนานกับแนวระดับ) โดยความสูงของกล้องที่ถูกติดตั้งมีระยะห่างจากพื้นดินประมาณ 2.5 เมตร ภาพที่ถ่ายเป็นภาพด้านข้างของถนน ใช้ภาพที่ถ่ายในประเทศไทย 10 จังหวัด ได้แก่ อ่างทอง พระนครศรีอยุธยา ฉะเชิงเทรา ชัยนาท เชียงราย ชลบุรี นครปฐม ปทุมธานี สุพรรณบุรี และ อุทัยธานี

4.2 สภาพแวดล้อมด้านซอฟต์แวร์และฮาร์ดแวร์

4.2.1 ฮาร์ดแวร์

หน่วยประมวลผล Intel Core i7-360QM / 2.3 GHz

การ์ดแสดงผลภาพ Nvidia Geforce GT 630M/ 2 GB / 96 Cuda cores

หน่วยความจำ 12.0 GB

4.2.2 ซอฟต์แวร์

Ubuntu 14.04 LTS

OpenCV 2.4.11 (Image processing library)

Torch7 (Machine Learning library)

Google Street View API

4.3 การหาค่าพารามิเตอร์ที่เหมาะสมสำหรับการประมวลผลภาพดิจิทัล

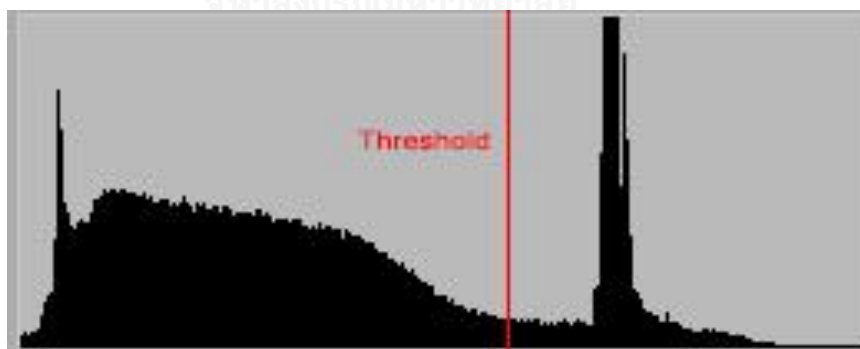
การประมวลผลภาพดิจิทัล เช่น ฮัฟทรานสฟอร์ม การตรวจหาสีของอาคาร จำเป็นต้องมีการปรับพารามิเตอร์เพื่อให้ได้ผลที่เหมาะสมแก่การตรวจหาอาคาร ขั้นตอนนี้เป็น การทดลองหาค่าพารามิเตอร์ที่เหมาะสมให้กับกระบวนการทำงานในหัวข้อ 3.1 – 3.7 โดยทำการเลือกภาพที่น่าสนใจ 50 ภาพเพื่อมาทดลองปรับค่าพารามิเตอร์ที่เหมาะสมในการตรวจหาอาคาร เป็นภาพอาคารทั้งหมด 20 ภาพและเป็นภาพที่ไม่มีอาคารทั้งหมด 30 ภาพ ภาพที่ไม่มีอาคารจะเป็นภาพถ่ายวิวธรรมชาติที่เห็นได้ตามข้างถนน เช่น ทุ่งนา ไร่ ภูเขา ลำคลอง แม่น้ำ เป็นต้น

4.3.1 การลบบริเวณท้องฟ้า

โดยปกติแล้วเมื่อแปลงภาพสี RGB เป็นสีขาวดำแบบไล่ระดับตามทฤษฎีที่ 2.2.2 แล้วจะได้ผลออกมาดังรูปที่ 4-1 บริเวณบริเวณท้องฟ้าจะมีความสว่างมากกว่าบริเวณพื้นดิน หากได้ค่าขีดแบ่งที่เหมาะสมก็จะสามารถแยกท้องฟ้าออกจากวัตถุที่อยู่บนพื้นดินได้



รูปที่ 4-1 ภาพตัวอย่างเมื่อแปลงภาพสี RGB เป็นขาวดำแบบไล่ระดับ



รูปที่ 4-2 ฮิสโตแกรมของภาพตัวอย่าง

จากฮิสโตแกรมรูปที่ 4-2 จะเห็นว่าค่าขีดแบ่งที่เหมาะสมจะอยู่ที่ช่วง 110-160 ค่าที่น้อยกว่า 110 จะเป็นส่วนของวัตถุบนพื้นดิน ค่าที่มากกว่า 160 จะเป็นส่วนของท้องฟ้า ขั้นตอนผู้วิจัยได้ทดลองใช้วิธีการหาค่าขีดแบ่งอัตโนมัติเพื่อหาค่าขีดแบ่ง แต่เมื่อนำมาทดสอบกับภาพจริงแล้ว กลับพบว่าไม่สามารถแยกท้องฟ้าออกจากวัตถุได้เท่าที่ควร ผู้วิจัยจึงได้หาค่าที่เหมาะสมโดยนำภาพทดสอบตั้งต้น

ทั้ง 50 ภาพมาหาค่าที่เหมาะสมที่สุดในการขีดแบ่งระหว่างท้องฟ้ากับวัตถุซึ่งค่าที่แยกได้ดีที่สุดคือ 147 การใช้ค่าขีดแบ่งจะใช้กับครึ่งบนของภาพเท่านั้น ซึ่งผลลัพธ์สามารถแยกท้องฟ้าได้ดียกเว้นในท้องฟ้าที่มีเมฆสีดำ ตัวอย่างผลลัพธ์ของขั้นตอนการทดลองแยกท้องฟ้าออกจากภาพถ่ายแสดงในรูปแบบที่ 4-3



รูปที่ 4-3 ตัวอย่างผลของขั้นตอนการทดลองแยกท้องฟ้าออกจากภาพถ่าย

4.3.2 การลบสีเขียวใบไม้

สีเขียวของพืชมีตั้งแต่เขียวอ่อนไปจนถึงเขียวเข้ม การทดลองนี้ใช้แบบจำลองสี HSV เพื่อตรวจสอบสีเขียวในภาพถ่าย จึงจำเป็นต้องปรับหาค่าสี ค่าความสดของสีและค่าความสว่างที่เหมาะสมเพื่อให้ได้ค่าสีที่ดีที่สุดในการตรวจสอบสีเขียวของพืชโดยไม่กระทบถึงสีของอาคาร จากการทดลองใน 50 ภาพ พบว่าค่าสีที่เหมาะสมอยู่ในช่วง 40 ถึง 120 ค่าความสดของสีอยู่ในช่วง 19 ถึง 100 ค่าความสว่างอยู่ในช่วง 0 ถึง 66 ตัวอย่างผลที่ได้จากการทดลองการกำจัดใบไม้สีเขียวในภาพถ่ายแสดงในรูปแบบที่ 4-4 (ข)



(ก)

(ข)

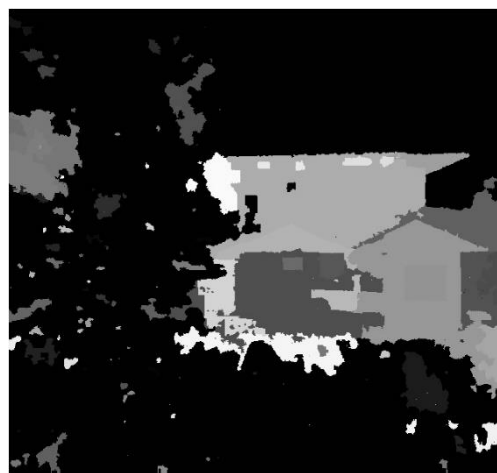
รูปที่ 4-4 ภาพนำเข้า (ก) และผลที่ได้จากการทดลองการกำจัดใบไม้สีเขียวในภาพถ่าย (ข)

4.3.3 การแบ่งส่วนภาพบริเวณของวัตถุด้วยสีและการติดป้ายหมายเลขให้แก่ละบริเวณ

การแบ่งส่วนภาพบริเวณของวัตถุด้วยสีใช้วิธีการ Automatic seeded region growing ซึ่งจำเป็นต้องปรับค่าให้เหมาะสมเพื่อให้สามารถแบ่งส่วนภาพได้ถูกต้อง หากตั้งค่าขีดแบ่งความเหมือนของค่าสีสูงเกินไป จะทำให้ค่าสีที่แตกต่างกันเล็กน้อยถูกมองว่าเป็นคนละบริเวณ หากตั้งค่าขีดแบ่งความเหมือนน้อยเกินไป จะทำให้ค่าสีที่แตกต่างกันถูกมองว่าเป็นบริเวณเดียวกัน จากการทดลองค่าที่เหมาะสมในขั้นตอนการเลือกเมล็ดโดยอัตโนมัติ ค่าความเหมือนต้องมากกว่า 0.8 ระยะทางแบบยุคลิดต่อจุดภาพต้องน้อยกว่า 0.08 ในขั้นตอนการขยายเมล็ดพันธุ์ ให้นำพิกเซลที่ยังไม่กำหนดบริเวณมาหาระยะทางแบบยุคลิดต่อบริเวณข้างเคียงที่น้อยที่สุด ในขั้นตอนผสานบริเวณเข้าด้วยกัน บริเวณนั้นต้องมีระยะทางแบบยุคลิดต่อบริเวณข้างเคียงที่น้อยที่สุดและมีระยะทางน้อยกว่า 0.04 ขั้นตอนสุดท้ายคือการรวมบริเวณเล็กเกินค่าขีดแบ่งเข้ากับบริเวณใหญ่ โดยบริเวณที่มีขนาดน้อยกว่า 2400 พิกเซล จะถูกนำมาหาระยะทางแบบยุคลิดต่อบริเวณข้างเคียงที่น้อยที่สุดและมีระยะทางน้อยกว่า 0.02 ตัวอย่างผลที่ได้จากการทดลองแบ่งส่วนภาพบริเวณของวัตถุด้วยสีแสดงในรูปที่ 4-5 (ข)



(ก)



(ข)

รูปที่ 4-5 ภาพนำเข้า (ก) และผลที่ได้จากการทดลองแบ่งส่วนภาพบริเวณของวัตถุด้วยสี (ข)

4.3.4 การลบบริเวณที่ไม่ต้องการ

การทดลองนี้เกิดขึ้นมีจุดประสงค์เพื่อลบบริเวณที่ไม่ต้องการที่ยังหลงเหลือมาจากขั้นตอนการแยกท้องฟ้าและการลบสีเขียวต้นไม้ แบ่งเป็น 3 ขั้นตอน ได้แก่

1. การกำจัดบริเวณที่เล็กเกินค่าขีดแบ่ง

หลังจากขั้นตอนการแบ่งส่วนภาพบริเวณของวัตถุด้วยสี ยังคงมีบริเวณที่เล็กเกินค่าขีดแบ่งที่ไม่สามารถรวมกับบริเวณอื่นได้ เนื่องจากไม่มีบริเวณข้างเคียงหรือค่าความเหมือนระหว่างบริเวณกับบริเวณข้างเคียงน้อยเกินค่าขีดแบ่ง ทำให้เหลือเป็นบริเวณขนาดเล็กซึ่งยากแก่การประมวลผล วิธีนี้จึงทำการกำจัดบริเวณที่น้อยกว่าค่าขีดแบ่งออก ค่าขีดแบ่งที่ใช้ในงานวิจัยนี้คือ 500 พิกเซล

2. การลบบริเวณที่มีสีขาวมากเกินไป

หลังจากการประมวลผลขั้นตอนการกำจัดสีเขียวใบไม้ พบว่ายังคงมีส่วนของใบไม้หลงเหลืออยู่ในภาพ ซึ่งเกิดขึ้นจากค่าสีของพิกเซลเหล่านั้นอยู่นอกเหนือช่วงที่ตั้งไว้ จึงจำเป็นต้องมีกระบวนการเพื่อลบพิกเซลเหล่านั้น ส่วนของใบไม้ที่ถูกกำจัดไปแล้วจะถูกตั้งค่าให้มีค่าของพิกเซลเท่ากับ 255 เพราะฉะนั้นหากพบว่าบริเวณใดมีค่าพิกเซลเท่ากับ 255 มากกว่าร้อยละ 30 ของบริเวณ บริเวณนั้นจะถูกลบออก

3. การลบบริเวณที่ติดกับขอบล่างของภาพ

ภาพที่นำมาใช้ในงานวิจัยนี้มาจากรถสำรวจ เพราะฉะนั้นภาพบริเวณขอบล่าง ส่วนมากจะเป็นถนนหรือพื้นดินบริเวณหน้าอาคาร การลบบริเวณที่ติดกับขอบล่างของภาพจึงสามารถลบบริเวณที่

เป็นถนนหรือพื้นดินหน้าอาคารได้ ในงานวิจัยนี้จะลบทุกๆบริเวณที่ติดกับขอบล่างสูงขึ้นมา 100 พิกเซล ตัวอย่างภาพผลที่ได้จากการทดลองกำจัดบริเวณที่ไม่ต้องการแสดงในรูปที่ 4-6 (ข)



(ก)



(ข)

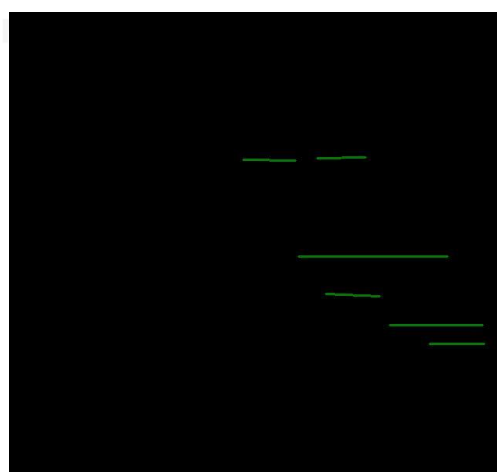
รูปที่ 4-6 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองกำจัดบริเวณที่ไม่ต้องการ (ข)

4.3.5 การทดลองหาเส้นตรงยาวในแนวนอน

การทดลองนี้เพื่อหาค่าที่เหมาะสมที่จะตรวจหาคานของอาคาร หากสั้นเกินไปจะตรวจพบสิ่งรบกวน (noise) มาก ค่าที่เหมาะสมคือเส้นตรงที่ได้ตัวยาวมากกว่า 100 พิกเซลและองศาของเส้นตรงอยู่ระหว่าง ± 5 องศา ตัวอย่างภาพผลที่ได้จากการทดลองหาเส้นตรงยาวในแนวนอนรูปที่ 4-7 (ข)



(ก)

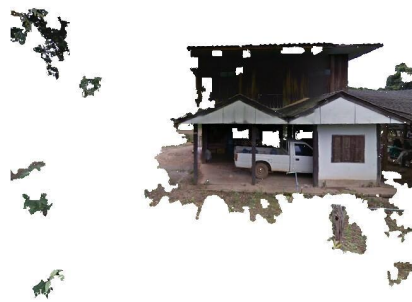


(ข)

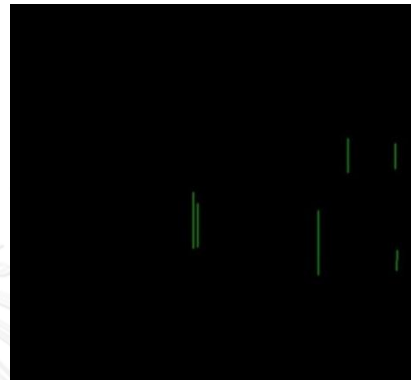
รูปที่ 4-7 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองหาเส้นตรงยาวในแนวนอน (ข)

4.3.6 การทดลองหาเส้นตรงที่ขนานกันในแนวตั้ง

การทดลองนี้เพื่อหาค่าที่เหมาะสมที่จะตรวจหาเสาของอาคาร โดยหาเส้นตรงที่อยู่ในแนวตั้ง จากนั้นจับคู่เส้นตรงโดยทั้งสองเส้นจะต้องขนานกันและอยู่ในแนวระดับเดียวกัน หากสั้นเกินไปจะตรวจพบสิ่งรบกวน (noise) มาก โดยมีองศาต่างกันไม่เกิน ± 5 องศา ระยะห่างระหว่างเส้นทั้งสองต้องมากกว่า 100 พิกเซลและระหว่างเส้นสองเส้นห้ามเป็นบริเวณที่ถูกกำจัดไปแล้ว ตัวอย่างภาพผลที่ได้จากการทดลองหาเส้นตรงที่ขนานกันในแนวตั้งแสดงในรูปที่ 4-8 (ข)



(ก)

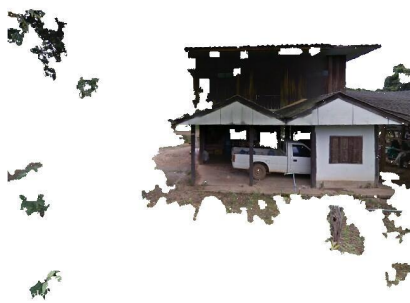


(ข)

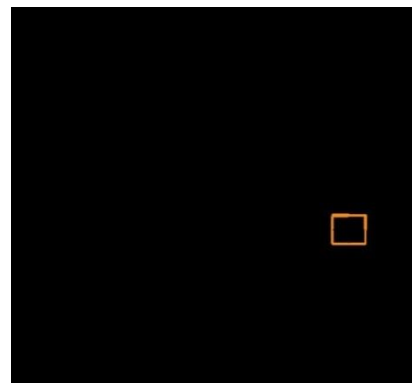
รูปที่ 4-8 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองหาเส้นตรงที่ขนานกันในแนวตั้ง (ข)

4.3.7 การหาสี่เหลี่ยมในภาพถ่าย

การทดลองนี้เพื่อหาค่าที่เหมาะสมที่จะตรวจหาประตูและหน้าต่าง ผลที่ได้จากการทดลองบริเวณเหมาะสมของรูปสี่เหลี่ยมที่ได้ควรมีขนาดมากกว่า 300 พิกเซล หากขนาดของสี่เหลี่ยมเล็กกว่ามีโอกาสสูงที่จะเป็นสิ่งรบกวน (noise) ตัวอย่างภาพผลที่ได้จากการทดลองหาสี่เหลี่ยมในภาพถ่ายแสดงในรูปที่ 4-9 (ข)



(ก)



(ข)

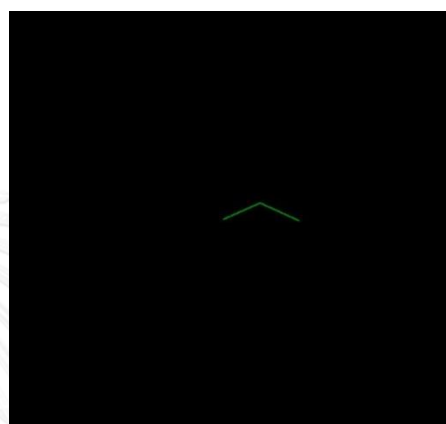
รูปที่ 4-9 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองหาสี่เหลี่ยมในภาพถ่าย (ข)

4.3.8 การหาจั่วหลังคา

การทดลองนี้เพื่อหาค่าที่เหมาะสมในขั้นตอนการตรวจหาหลังคาของอาคาร จั่วของหลังคาสามารถตรวจหาด้วยเทคนิคประมวลผลภาพดิจิทัลได้โดยใช้ Hough transform ตรวจหาเส้นตรง จากนั้นหาเส้นตรงสองเส้นที่ทำมุมตรงข้ามกัน โดยองศาของทั้งคู่แตกต่างกันได้ไม่เกิน ± 5 องศา เช่น 45 องศา กับ 135 องศา และเนื่องจากมีหลายภาพที่ Hough transform ตรวจพบส่วนยอดของจั่วไม่ติดกัน จึงขจัดเซตค่าที่ส่วนยอดของทั้งสองเส้นให้อยู่ห่างกันได้ไม่เกิน ± 7 พิกเซลทั้งในแนวตั้งและแนวนอน ตัวอย่างผลที่ได้จากการทดลองหาจั่วหลังคาแสดงในรูปที่ 4-10 (ข)



(ก)



(ข)

รูปที่ 4-10 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองหาจั่วหลังคา (ข)

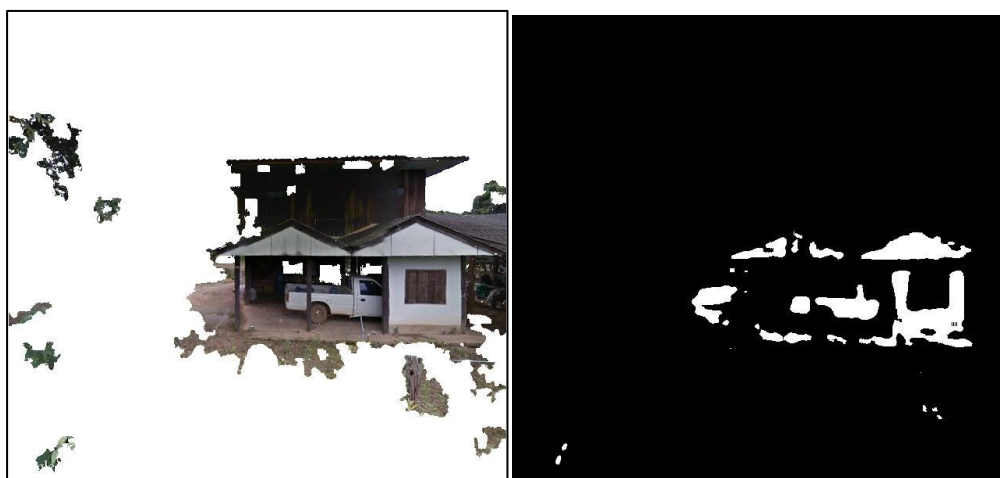
4.3.9 การตรวจหาสีอาคาร

การทดลองนี้เพื่อหาค่าที่เหมาะสมในการตรวจหาสีที่นำมาใช้ทาอาคาร ขั้นตอนนี้ใช้แบบจำลองสี HSV ในการตรวจหาสีแดง น้ำเงิน น้ำตาล เหลืองและขาว หากตั้งค่าของช่วงสีไว้กว้างมากเกินไปจะทำให้ไปตรวจพบวัตถุอื่น ๆ ที่มีทั่วไปในธรรมชาติ หากตั้งค่าของช่วงสีไว้มากเกินไปจะทำให้หาอาคารได้น้อยลง ค่าสีที่เหมาะสมที่ได้จากการทดลองแสดงในตารางที่ 4.1

ตารางที่ 4.1 ค่า HSV ที่เหมาะสมจากการทดลองเพื่อตรวจหาอาคาร

สี	ค่าสี	ค่าความเข้มของสี	ค่าความสว่างของสี
สีแดง	0-34, 340-360	35-100	33-86
สีน้ำเงิน	150-200	35-94	19-78
สีน้ำตาล	216-244	1-44	66-80
สีเหลือง	16-50	10-17	11-35
สีขาว	30-60	19-78	39-58

เมื่อตรวจพบสีที่อยู่ในช่วงที่ตั้งไว้ จะทำการแยกออกมาดังรูปที่ 4-11 (ข)



(ก)

(ข)

รูปที่ 4-11 ภาพนำเข้า (ก) และภาพผลที่ได้จากการทดลองตรวจหาสีอาคาร (ข)

4.3.10 การหาขอบเขตของบริเวณของวัตถุที่สนใจ

การทดลองนี้เพื่อหาค่าที่เหมาะสมสำหรับจุดเริ่มต้นและจุดสิ้นสุดของวิธีการโปรเจกชัน (Projection profile) โดยจะทดลองวิธีการโปรเจกชันทั้งแนวตั้งและแนวนอน ค่าขีดแบ่งในการเป็นจุดเริ่มต้นนั้นหากตั้งไว้น้อยเกินไป มีโอกาสที่จะเจอสิ่งแปลกปลอมได้มาก หากตั้งไว้สูงเกินไปอาจทำให้ได้บริเวณผิดพลาด โปรเจกชันตามแนวตั้งจุดเริ่มต้นคือตำแหน่งที่วิธีการโปรเจกชันนับได้เกิน 7 พิกเซล จุดสิ้นสุดคือตำแหน่งที่วิธีการโปรเจกชันนับได้น้อยกว่า 7 พิกเซล หลังจากจุดสิ้นสุดไปแล้วหาพบจุดเริ่มต้นใหม่โดยที่ห่างจากจุดสิ้นสุดเดิมน้อยกว่า 100 พิกเซลจะถือว่าจุดเริ่มต้นใหม่อยู่ในบริเวณเดียวกับบริเวณเดิม หากจุดสิ้นสุดใหม่อยู่ห่างจากจุดสิ้นสุดเดิมเกิน 100 พิกเซลจะถือว่าจุดเริ่มต้นใหม่คือบริเวณใหม่ ต่อจากนั้นดำเนินการโปรเจกชันในแนวนอนในแต่ละบริเวณเพื่อหาความสูงของบริเวณนั้นๆ โดยจุดเริ่มต้นเริ่มตั้งแต่พบค่าโปรเจกชันมากกว่า 15 พิกเซลไปจนถึงจุดแรกจากด้านล่างที่ที่พบค่าโปรเจกชันมากกว่า 15 พิกเซลเพื่อกรองสิ่งรบกวน (noise)

4.3.11 การตรวจสอบขนาดและอัตราส่วนของบริเวณ

การทดลองนี้เพื่อหาค่าที่เหมาะสมเพื่อตรวจสอบบริเวณที่ได้ว่าใช่อาคารหรือไม่ ขั้นตอนแรกตรวจสอบขนาดของบริเวณ หากบริเวณเล็กเกินไปมีโอกาสสูงที่จะเป็นสิ่งรบกวนดังรูปที่ 4-12 บริเวณที่ได้ต้องมีความกว้างอย่างน้อย 100 พิกเซลและความสูงอย่างน้อย 100 พิกเซล อัตราส่วนกว้างถึงยาวของบริเวณต้องไม่ต่างกันมากเกินไป



รูปที่ 4-12 ความผิดพลาดที่เกิดจากไม่ได้ตรวจสอบขนาดและอัตราส่วนของบริเวณ

4.5 การนำพารามิเตอร์มาทดสอบกับภาพถ่าย

ขั้นตอนนี้ได้ทำการนำพารามิเตอร์ที่ได้รับจากหัวข้อ 4.4 มาทดสอบกับภาพถ่ายตัวอย่าง ซึ่งยังไม่มี การใช้โครงข่ายประสาทเทียมในการทดสอบนี้ โดยสรุปผลการทดสอบได้ดังนี้

4.5.1 การลบบริเวณท้องฟ้าและการลบสีเขียวของใบไม้

จากการทดสอบขั้นตอนการลบบริเวณท้องฟ้าพบว่าสามารถดำเนินการลบบริเวณท้องฟ้าได้ดี ในสถานะที่ท้องฟ้าโปร่งหรือมีเมฆขาว แต่ยังคงพบปัญหาไม่สามารถทำงานได้อย่างมีประสิทธิภาพใน เมฆดำ อันเนื่องมาจากวิธีการนี้ใช้ค่าขีดแบ่ง (Threshold) ในการแยกบริเวณท้องฟ้ากับวัตถุบนพื้นดิน โดยปกติแล้วบริเวณท้องฟ้าจะมีค่าความสว่างมากกว่าวัตถุบนพื้นราบ แต่ในกรณีนี้ความเข้มของเมฆสีดำให้ค่าความสว่างใกล้เคียงกับวัตถุบนพื้นราบทำให้ไม่สามารถแยกเมฆสีดำออกจากภาพถ่ายได้ ซึ่ง บางครั้งเมฆดำเหล่านี้จะสร้างความผิดพลาดในขั้นตอนการตรวจหาอาคารได้ เนื่องจากใกล้เคียงกับ ช่วงค่าสีขาวยที่ใช้ในการตรวจหาอาคาร

ขั้นตอนการลบสีเขียวใบไม้สามารถลบสีเขียวในภาพได้แต่จะมีบางบริเวณซึ่งมีเงามาบดบังทำให้ค่าที่ได้อยู่นอกเหนือจากช่วงที่กำหนดไว้ สีเขียวของต้นไม้ในส่วนนี้จึงไม่ถูกลบไปด้วย แต่ส่วน บริเวณที่ไม่ถูกลบเหล่านี้ก็เป็นเพียงส่วนเล็กน้อยจากภาพใหญ่เท่านั้น ซึ่งไม่เป็นปัญหาแก่การตรวจหา อาคารแต่อย่างใด

4.5.2 การทดสอบหาอาคารที่มีสิ่งบดบังและไม่มีสิ่งบดบัง

ขั้นตอนนี้ได้ทดสอบขั้นตอนการตรวจหาอาคารจากภาพถ่ายกับภาพอาคารโดยที่ไม่มีสิ่งกีดขวาง 100 ภาพและภาพอาคารที่มีสิ่งกีดขวาง 100 ภาพ ได้ความถูกต้องดังตารางที่ 4.2

ตารางที่ 4.2 ความถูกต้องที่ได้จากการทดลองในแต่ละโมเดล

รายละเอียด	ความถูกต้อง (เฉลี่ยร้อยละ)
ภาพอาคารที่ไม่มีสิ่งกีดขวาง	92
ภาพอาคารที่มีสิ่งกีดขวาง	83

ผลการทดสอบแสดงความถูกต้องที่น่าพอใจ โดยสามารถหาอาคารได้หลากหลายรูปแบบ ความผิดพลาดที่เกิดขึ้นส่วนใหญ่เกิดขึ้นมาจาก 2 สาเหตุหลัก ได้แก่ รอยต่อของภาพที่เกิดขึ้นจากการอัลกอริทึมการต่อภาพของกูเกิล ดังรูปที่ 4-13 ซึ่งจะส่งผลทำให้ขั้นตอนการตรวจหาเส้นตรงในภาพถ่าย เช่น การหาเส้นตรงแนวราบ การหาจั่วหลังคา การหาสี่เหลี่ยมและการหาเส้นตรงขนานในแนวตั้งผิดพลาด เนื่องจากเส้นตรงแต่ละเส้นจะถูกกำหนดความยาวขั้นต่ำไว้ ซึ่งรอยต่อเหล่านี้จะทำให้เส้นตรงไม่ต่อเนื่อง ทำให้ความยาวของเส้นตรงไม่ถึงค่าขีดแบ่งที่ตั้งไว้



รูปที่ 4-13 รอยต่อของภาพที่เกิดขึ้นจากการอัลกอริทึมการต่อภาพของกูเกิล

สาเหตุที่สองคือรายละเอียดของอาคารมีน้อยเกินไปทำให้ถูกคัดออกไปในขั้นตอนการโพรเจกชัน โดยสามารถเกิดขึ้นได้หลายกรณี เช่น ค่าสีอยู่นอกเหนือช่วงที่ตั้งไว้ อาคารถูกบดบัง รายละเอียดมากเกินไป เป็นต้น ซึ่งเมื่อนำมาเข้าสู่ขั้นตอนการโพรเจกชันแล้วจะไม่สามารถสรุปได้ว่าชิ้นส่วนนั้นคืออาคารหรือสิ่งรบกวน



(ก)



(ข)

รูปที่ 4-14 (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์จากการดำเนินการตรวจหาอาคารจากภาพถ่าย

จากรูปที่ 4-14 จะเห็นได้ว่าภาพผลลัพธ์จากการดำเนินการตรวจหาอาคารจากภาพถ่าย ไม่จะเป็นการหาเส้นตรงของอาคารและการหาสีของอาคารได้ผลออกมาเพียงเล็กน้อย ซึ่งอาจเกิดได้จากหลายปัจจัย เช่น ค่าความสว่างในภาพทำให้สีมืดหรือสว่างมากกว่าค่าของขั้นตอนการหาสีของอาคารตั้งไว้ มีวัตถุ ยกตัวอย่างเช่น ต้นไม้มาบดบังรายละเอียดของตัวบ้านมากเกินไป จนทำให้ไม่สามารถตรวจหาโครงสร้างของอาคารได้ เป็นต้น ในกรณีของรูปที่ 4-14 นี้มีต้นไม้มาบดบังโครงสร้างของบ้านและสีของบ้านไม่สม่ำเสมอจึงทำให้เกิดความผิดพลาดขึ้น

4.5.3 การทดสอบกับภาพด้านข้างของรถสำรวจภูเกิลสตรีทวิว

ขั้นตอนนี้ทดสอบกับภาพที่มีอาคาร 332 ภาพและภาพที่ไม่มีอาคาร 604 ภาพ ความถูกต้องที่ได้สำหรับภาพที่มีอาคารร้อยละ 82.5 และความถูกต้องสำหรับภาพที่ไม่มีอาคารร้อยละ 96.58 ค่าความผิดพลาดเชิงบวก เกิดขึ้นจาก 3 สาเหตุหลักคือ ต้นไม้ที่ไม่สามารถลบออกได้ในขั้นตอนการลบต้นไม้เนื่องจากต้นไม้เมื่อนำมาเข้าขั้นตอนหาเส้นตรงโดยฮัฟฟทรานฟอร์มแล้วจะทำให้หาเส้นตรงได้ไม่ถูกต้องและเกิดสิ่งรบกวน (noise) จำนวนมาก บริเวณการเกษตรที่ถูกเผาจะทำให้ขั้นตอนการตรวจหาสีของอาคารตรวจพบเป็นสีน้ำตาลและการสะท้อนของน้ำจะทำให้ขั้นตอนการตรวจหาสีของอาคารตรวจพบเป็นสีขาว ความผิดพลาดเชิงบวกจากการทดลองแสดงในตารางที่ 4.3

ตารางที่ 4.3 ความผิดพลาดเชิงบวกจากการทดลอง

สาเหตุของความผิดพลาด	จำนวนภาพ
ต้นไม้ที่ไม่สามารถลบออกได้ในขั้นตอนการลบต้นไม้	21
บริเวณการเกษตรที่ถูกเผา	5
การสะท้อนของน้ำ	2

ค่าผิดพลาดเชิงลบมี 3 สาเหตุหลัก คือ อาคารที่อยู่ไกล มีวัตถุมาบดบังมากเกินไปและตรวจหาส่วนอาคารได้น้อยเกินไป ซึ่งจากสามข้อที่กล่าวมานี้ล้วนทำให้ขั้นตอนการการโปรเจกชันไม่สามารถสรุปได้ว่าชิ้นส่วนนั้นคืออาคารหรือสิ่งรบกวน ความผิดพลาดเชิงลบจากการทดลอง แสดงในตารางที่ 4.4

ตารางที่ 4.4 ความผิดพลาดเชิงลบจากการทดลอง

สาเหตุของความผิดพลาด	จำนวนภาพ
อาคารอยู่ไกล	11
มีวัตถุมาบดบัง	39
ตรวจการรายละเอียดของอาคารได้น้อย	8

4.6 การรู้จำอาคารด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันนัล

การทดลองนี้เพื่อหาโมเดลโครงข่ายประสาทเทียมที่เหมาะสมในการใช้ในการรู้จำอาคารจากภาพถ่าย โดยใช้ภาพเพื่อมาฝึกโครงข่ายประสาทเทียมทั้งหมด 7,343 ภาพ แบ่งเป็นภาพที่มีอาคาร 3,995 ภาพและภาพที่ไม่มีอาคาร 3,348 ภาพ ซึ่งใช้ภาพถ่ายจากจังหวัดอ่างทอง ออยุธยา ฉะเชิงเทรา ลำปาง ชลบุรี ชัยนาทและเชียงราย ภาพเพื่อนำมาทดสอบทั้งหมด 3,014 ภาพ แบ่งเป็นภาพที่มีอาคาร 1,514 ภาพและภาพที่ไม่มีอาคารทั้งหมด 1,500 ภาพ ซึ่งใช้ภาพถ่ายจากกาญจนบุรี ลำพูน และนครปฐม โดยทดลอง 2 การทดลองดังนี้

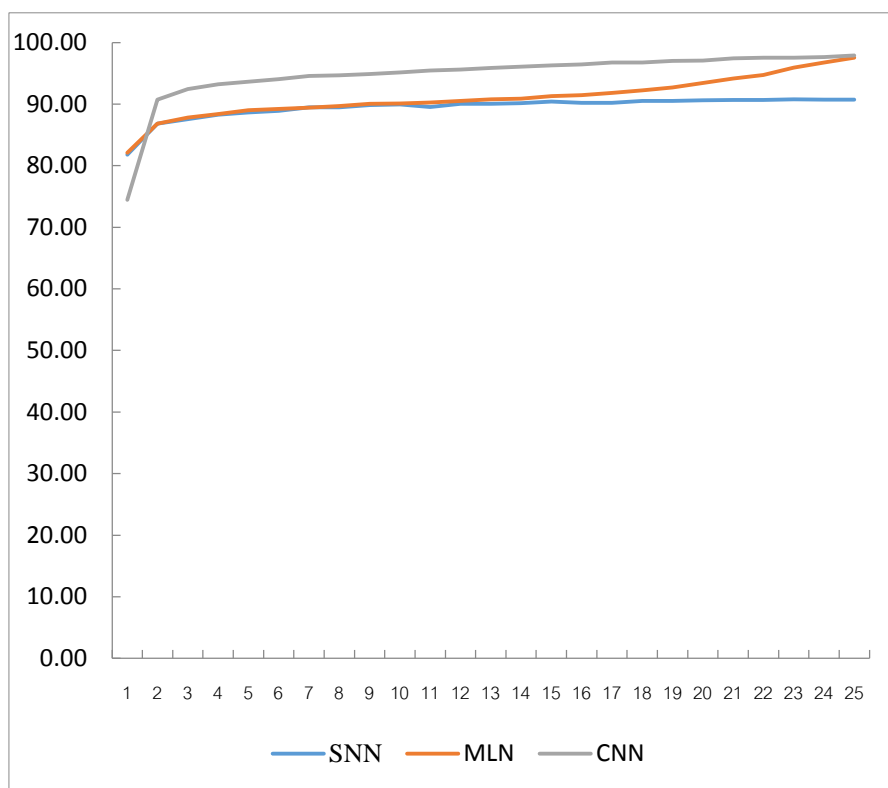
4.6.1 การทดลองเปรียบเทียบโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลกับโครงข่ายประสาทเทียมอื่น

การทดลองนี้มีเป้าหมายเพื่อเปรียบเทียบผลลัพธ์ระหว่างโครงข่ายประสาทเทียมแบบคอนโวลูชันนัล (CNN) กับโครงข่ายประสาทเทียมอื่นๆ เช่น โครงข่ายประสาทเทียมแบบชั้นเดียว (SNN) และโครงข่ายประสาทเทียมแบบหลายชั้น (MLN) ข้อมูลเข้าใช้ขนาด 32 พิกเซล x 32 พิกเซล โครงข่ายประสาทเทียมแบบหลายชั้นจะใช้ชั้นซ่อนทั้งหมดสองชั้นและทุกโมเดลจะทำการฝึกโมเดลละ 25 รอบ ผลลัพธ์ที่ได้แสดงในตารางที่ 4.5

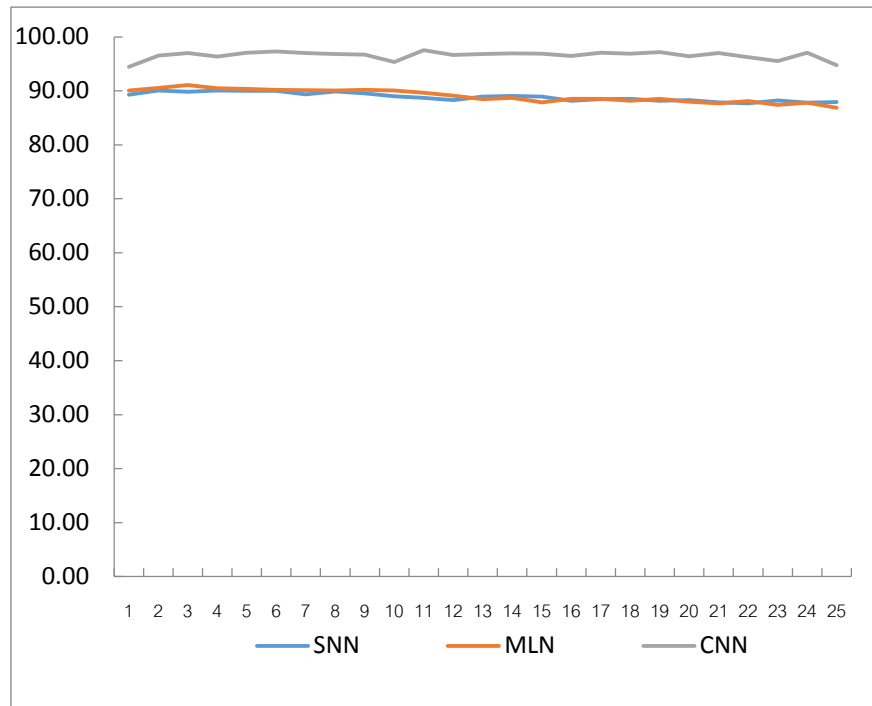
ตารางที่ 4.5 ผลการทดลองเปรียบเทียบโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลกับโครงข่ายประสาทเทียมอื่นๆ

โมเดล	ความถูกต้องสูงสุด	ความถูกต้องต่ำสุด
โครงข่ายประสาทเทียมแบบชั้นเดียว (SNN)	86.20%	87.78%
โครงข่ายประสาทเทียมแบบหลายชั้น (MLN)	91.10%	86.89%
โครงข่ายประสาทเทียมแบบคอนโวลูชันนัล (CNN)	97.51%	94.46%

โครงข่ายประสาทเทียมแบบคอนโวลูชันนัลให้ความถูกต้องสูงสุดร้อยละ 97.51 โครงข่ายประสาทเทียมแบบชั้นเดียวใช้เวลาในการฝึกที่ดีที่สุดแต่วิธีนี้สามารถรู้จำอาคารได้ถูกต้องสูงสุดร้อยละ 86.20 และมีปัญหาจำการฝึกมากเกินไปในขั้นตอนทดสอบกับภาพทดสอบ โครงข่ายประสาทเทียมแบบหลายชั้นสามารถรู้จำได้ถูกต้องร้อยละ 91.10 วิธีนี้ใช้เวลาในขั้นตอนการฝึกมากที่สุด อีกทั้งยังมีปัญหาจำชุดการฝึกมากเกินไปอีกด้วย ความถูกต้องของการฝึกและการทดสอบแสดงในรูปที่ 4-15 และ 4-16



รูปที่ 4-15 แผนภูมิแสดงความถูกต้องที่เพิ่มขึ้นในแต่ละรอบของการฝึก



รูปที่ 4-16 แผนภูมิแสดงค่าความถูกต้องของการทดลองกับชุดทดสอบ

4.6.2 การทดลองหาโมเดลที่ดีที่สุดในการรู้จำอาคาร

การทดลองนี้เพื่อหาโมเดลโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลที่ดีที่สุดในการรู้จำอาคาร โดยทดลองจากข้อมูลนำเข้าขนาด 32 พิกเซล x 32 พิกเซล, 96 พิกเซล x 96 พิกเซล และ 192 พิกเซล x 192 พิกเซล นำมาทดลองกับโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลแบบ 2 ชั้น, 3 ชั้น และ 5 ชั้น โดยสรุปจะได้เป็น 9 โมเดลดังตารางที่ 4.6

ตารางที่ 4.6 โมเดลที่ใช้ในการทดลองหาโมเดลโครงข่ายประสาทเทียมแบบคอนโวลูชันแนลที่ดีที่สุดในการรู้จำอาคาร

Model no.	Input (resolution)	C1	P1	C 2	P2	C3	P3	C4	P4	C5	P5
1	32x32	32x5x5	2x2	64x5x5	2x2	-	-	-	-	-	-
2	32x32	32x5x5	2x2	64x5x5	2x2	128x3x3	-	-	-	-	-
3	32x32	32x5x5	2x2	64x5x5	-	128x3x3	-	256x3x3	-	256x3x3	-
4	96x96	32x7x7	5x5	64x7x7	3x3	-	-	-	-	-	-
5	96x96	32x7x7	5x5	64x7x7	-	128x7x7	-	-	-	-	-
6	96x96	32x7x7	2x2	64x7x7	3x3	128x7x7	-	256x3x3	-	256x3x3	-
7	192x192	32x11x11	5x5	64x11x11	5x5	-	-	-	-	-	-
8	192x192	32x11x11	5x5	64x11x11	5x5	-	-	-	-	-	-
9	192x192	32x11x11	2x2	64x11x11	2x2	128x11x11	2x2	256x7x7	-	256x7x7	-

โดยชั้นคอนกรีตแทนที่ด้วยตัวอักษร Ci โดยที่ i คือเลขที่ชั้น ซึ่งขนาดของชั้นคอนกรีตคือ $K \times m \times m$ โดยที่ K คือจำนวนแผ่นที่พีเจอร์ m และ n คือขนาดของเคอร์เนล , ชั้น पुलिंगแทนที่ด้วยตัวอักษร Pi โดยที่ i คือเลขที่ชั้น ขนาดของชั้น पुलिंगแทนที่ด้วย $m \times n$ ทุกโมเดลทำการฝึกโมเดล 25 รอบ ผลความถูกต้องแสดงในตารางที่ 4.7

ตารางที่ 4.7 ความถูกต้องที่ได้จากการทดลองในแต่ละโมเดล

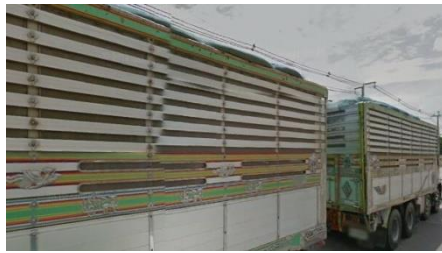
หมายเลขโมเดล	ภาพที่มีอาคาร	ภาพไม่มีอาคาร	ผลลัพธ์เฉลี่ย
1	97.95%	97.06%	97.50%
2	96.49%	97.40%	96.90%
3	98.08%	96.60%	97.34%
4	98.21%	97.40%	97.80%
5	98.67%	97.33%	98.00%
6	98.87%	96.93%	97.90%
7	97.81%	98.20%	98.00%
8	97.95%	97.60%	97.77%
9	98.94%	96.73%	97.84%

จากการทดลองทุกโมเดลให้ความถูกต้องเฉลี่ยร้อยละ 96.49 ถึงเฉลี่ยร้อยละ 98.94 ผลลัพธ์ที่ได้แสดงให้เห็นว่าโมเดลที่ 5 กับโมเดล 7 ให้ความถูกต้องที่ดีที่สุดถึงร้อยละ 98.00

4.7 ผลการทดลอง

4.7.1 ทดสอบทั้งระบบกับภาพถ่ายในชนบท

ภาพเพื่อนำมาทดสอบทั้งหมด 3,014 ภาพ แบ่งเป็นภาพที่มีอาคาร 1,514 ภาพและภาพที่ไม่มีอาคารทั้งหมด 1,500 ภาพ ซึ่งใช้ภาพถ่ายจากอ่างทอง นนทบุรี ปทุมธานี อุทัยธานี กาญจนบุรี ลำพูน และนครปฐม ความถูกต้องที่ได้สำหรับภาพที่มีอาคารเฉลี่ยร้อยละ 88 ความถูกต้องที่ได้สำหรับภาพที่ไม่มีอาคารเฉลี่ยร้อยละ 97.6 ค่าความผิดพลาดเชิงลบเกิดขึ้นทั้งหมด 182 ภาพ เกิดจากการตรวจหาที่ผิดพลาดของโครงข่ายประสาทเทียมคอนกรีตชั้นนัล 67 ภาพ เกิดจากการตรวจหารายละเอียดของอาคารได้น้อยเกินไป 118 ภาพ ค่าความผิดพลาดเชิงบวกเกิดขึ้น 37 ภาพ สาเหตุจากการตรวจหาที่ผิดพลาดของโครงข่ายประสาทเทียมคอนกรีตชั้นนัล ตัวอย่างของภาพที่ตรวจหาผิดพลาดของโครงข่ายประสาทเทียมคอนกรีตชั้นนัลแสดงในรูปที่ 4-17 และตัวอย่างตัวอย่างภาพที่ตรวจหาไม่ได้เนื่องจากรายละเอียดของอาคารได้น้อยเกินไปแสดงในรูปที่ 4-18



(ก)



(ข)



(ค)



(ง)

รูปที่ 4-17 ตัวอย่างภาพที่ตรวจหาผิดพลาดของโครงข่ายประสาทเทียมคอนโวลูชันนัล



(ก)



(ข)

รูปที่ 4-18 ตัวอย่างภาพที่ตรวจหาไม่ได้เนื่องจากรายละเอียดของอาคารได้น้อยเกินไป

4.7.2 ทดสอบกับภาพในหัวข้อ 4.5.3

จากหัวข้อที่ 4.5.3 การทดสอบกับภาพด้านข้างของรถสำรวจ ได้ทดสอบกับภาพที่มีอาคาร 332 ภาพและภาพที่ไม่มีอาคาร 604 ภาพ ความถูกต้องที่ได้สำหรับภาพที่มีอาคารเฉลี่ยร้อยละ 82.5 และความถูกต้องสำหรับภาพที่ไม่มีอาคารเฉลี่ยร้อยละ 96.58 ค่าความผิดพลาดเชิงบวก เกิดขึ้นจาก 3 สาเหตุหลักคือ ต้นไม้ที่ไม่สามารถลบออกได้ในขั้นตอนการลบต้นไม้เนื่องจากต้นไม้เมื่อนำมาเข้า

ขั้นตอนหาเส้นตรงโดยฮัพทรานฟอร์มแล้วจะทำให้หาเส้นตรงได้ไม่ถูกต้องและเกิดสิ่งรบกวน (noise) จำนวนมาก บริเวณการเกษตรที่ถูกเผาจะทำให้ขั้นตอนการตรวจหาสีของอาคารตรวจพบเป็นสีน้ำตาล และการสะท้อนของน้ำจะทำให้ขั้นตอนการตรวจหาสีของอาคารตรวจพบเป็นสีขาว เมื่อได้นำมาประมวลผลในโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลแล้วสามารถลดค่าความผิดพลาดเชิงบวกได้ เนื่องจากโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลสามารถรู้จำได้ถูกต้องสูงสุดร้อยละ 98 ซึ่งสามารถรู้จำวัตถุที่ไม่ใช่อาคารได้ดี ทำให้ความถูกต้องสำหรับภาพที่มีอาคารเพิ่มขึ้นเป็นเฉลี่ยร้อยละ 87.5 และความถูกต้องสำหรับภาพที่ไม่มีอาคารเพิ่มขึ้นเป็นเฉลี่ยร้อยละ 100

4.7.3 การทดลองเปรียบเทียบความถูกต้องของโครงข่ายประสาทเทียมโดยใช้ภาพที่ผ่านกระบวนการทำงานในงานวิจัยนี้เทียบกับการนำภาพต้นฉบับเข้าสู่โครงข่ายประสาทเทียมโดยตรง

การทดลองนี้เป็นการเปรียบเทียบผลลัพธ์ความถูกต้องของโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลโดยใช้ภาพที่ผ่านกระบวนการทำงานในงานวิจัยนี้เทียบกับการนำภาพต้นฉบับเข้าสู่โครงข่ายประสาทเทียมโดยตรง โดยข้อมูลนำเข้าจะแบ่ง 2 ชุด คือ ชุดภาพที่ผ่านการประมวลผลภาพในงานวิจัยนี้และชุดภาพที่ไม่ผ่านการประมวลผลภาพในงานวิจัยนี้ ซึ่งแต่ละชุดจะใช้ภาพเพื่อมาฝึกโครงข่ายประสาทเทียมทั้งหมด 7,343 ภาพ แบ่งเป็นภาพที่มีอาคาร 3,995 ภาพและภาพที่ไม่มีอาคาร 3,348 ภาพเพื่อนำมาทดสอบทั้งหมด 3,014 ภาพ แบ่งเป็นภาพที่มีอาคาร 1,514 ภาพและภาพที่ไม่มีอาคารทั้งหมด 1,500 ภาพ ผลความถูกต้องแสดงในตารางที่ 4.8

ตารางที่ 4.8 ความถูกต้องของโครงข่ายประสาทเทียมที่ได้จากการการผ่านกระบวนการทำงานในงานวิจัยนี้เทียบกับการนำภาพต้นฉบับเข้าสู่โครงข่ายประสาทเทียมโดยตรง

รายละเอียด	ภาพที่มีอาคาร	ภาพที่ไม่มีอาคาร
ความถูกต้องเฉลี่ยของโครงข่ายประสาทเทียมเมื่อใช้ภาพที่ได้จากการการผ่านกระบวนการทำงานในงานวิจัยนี้เป็นภาพนำเข้า	87.50%	97.60%
ความถูกต้องเฉลี่ยของโครงข่ายประสาทเทียมที่ได้จากการนำภาพต้นฉบับนำเข้าสู่โครงข่ายประสาทเทียมโดยตรง	71.89%	97.53%

จากการทดลองขั้นต้นพบว่าความถูกต้องเฉลี่ยของโครงข่ายประสาทเทียมเมื่อใช้ภาพที่ได้จากการการผ่านกระบวนการทำงานในงานวิจัยนี้มีความถูกต้องเฉลี่ยของภาพที่มีอาคารเฉลี่ยร้อยละ 87.50 ความถูกต้องเฉลี่ยของภาพที่ไม่มีอาคารเฉลี่ยร้อยละ 97.60 ส่วนการนำภาพต้นฉบับเข้าสู่โครงข่ายประสาทเทียมโดยตรงให้ความถูกต้องเฉลี่ยน้อยกว่าการนำภาพมาผ่านกระบวนการตรวจหาอาคารในงานวิจัยนี้ก่อน คือ ความถูกต้องเฉลี่ยของภาพที่มีอาคารเฉลี่ยร้อยละ 71.89 ความถูกต้องเฉลี่ยของภาพที่ไม่มีอาคารเฉลี่ยร้อยละ 97.53 เมื่อนำมาเปรียบเทียบกันแล้วจะเห็นความแตกต่างของความถูกต้องเฉลี่ยอย่างชัดเจนสำหรับภาพที่มีอาคาร อีกทั้งการนำภาพต้นฉบับเข้าสู่ประสาทเทียมโดยตรงยังพบปัญหาโครงข่ายประสาทเทียมรู้จำการฝึกมากเกินไป (Overfitting) อีกด้วย

4.8 การวิเคราะห์ผลการทดลอง

เนื่องจากงานวิจัยนี้เป็นการตรวจหาอาคารจากภาพถ่ายในพื้นที่ชนบทของประเทศไทย ซึ่งภาพที่นำมาประมวลผลไม่มีการควบคุมสภาวะแวดล้อมอาจทำให้เกิดความผิดพลาดในการประมวลผลได้มาก ทางผู้วิจัยได้พยายามทำภาพถ่ายให้อยู่ในสภาวะที่ควบคุมได้โดยทำการลบสิ่งที่จะเป็นปัญหาออกไปจากภาพถ่าย เช่น ต้นไม้ ท้องฟ้าและพื้นถนน ซึ่งความถูกต้องที่ได้ออกมาค่อนข้างดีตามที่ทดลองในหัวข้อ 4.5.1 เมื่อลบสิ่งเหล่านี้ออกไปแล้วก็จะเหลือส่วนที่เข้าข่ายว่าจะเป็นส่วนของอาคาร อย่างไรก็ตามก็ยังคงมีหลายปัจจัยที่จะทำให้เกิดความผิดพลาดได้ เช่น ต้นไม้ที่บดบังอาคารไว้แม้จะถูกลบไปแล้วในขั้นตอนก่อนหน้าแต่ก็ทำให้ส่วนที่เหลือของอาคารในภาพไม่ต่อเนื่องซึ่งยากแก่การวิเคราะห์ ผู้วิจัยได้ใช้การตรวจหาสีที่นิยมในการใช้เป็นสีของอาคารหรือส่วนของ เช่น ขาว แดง น้ำเงิน น้ำตาล เหลือง และเทาโครงสร้างของอาคาร ซึ่งจากการสังเกตบ้านส่วนใหญ่ในชนบทจะมีลักษณะเป็นทรงสี่เหลี่ยม แต่การจะตรวจหาสี่เหลี่ยมที่สมบูรณ์นั้นทำได้ยากเนื่องจากมีสิ่งบดบังอยู่มากมาย ผู้วิจัยจึงได้หาเส้นตรงจากภาพถ่ายแทนซึ่งทำได้ง่ายกว่า โดยตรวจหาเส้นตรงยาวในแนวราบเพื่อหาส่วนของอาคาร การตรวจหาเส้นตรงในแนวตั้งเพื่อหาเสาหรือขอบของอาคาร การหาจั่วของอาคาร ตรวจหาสีเหลี่ยมเพื่อหาประตูและหน้าต่างของอาคาร แนวคิดของผู้วิจัยคือหากพบส่วนเหล่านี้อยู่ในบริเวณเดียวกันก็มีโอกาสเป็นไปได้สูงที่บริเวณนั้นจะเป็นอาคาร โดยการกำหนดค่าพารามิเตอร์ต้องเป็นไปอย่างรอบคอบ หากตั้งค่าไว้รัดกุมน้อยเกินไปจะทำให้เกิดสิ่งแปลกปลอม (noise) ได้มาก หากตั้งรัดกุมมากเกินไปก็จะทำให้ไม่สามารถหาอาคารบางหลังได้ แนวคิดดังกล่าวได้ทดลองไว้ในหัวข้อ 4.5.3 โดยทดสอบกับภาพที่มีอาคาร 332 ภาพและภาพที่ไม่มีอาคาร 604 ภาพ ความถูกต้องที่ได้สำหรับภาพที่มีอาคารเฉลี่ยร้อยละ 82.5 และความถูกต้องสำหรับภาพที่ไม่มีอาคารเฉลี่ยร้อยละ 96.58 ซึ่งเป็นความถูกต้องที่สูงมาก แต่ก็ยังมีความผิดพลาดเกิดขึ้นตามที่อธิบายไว้ในหัวข้อ 4.5.3 ค่าความผิดพลาดเชิงบวก เกิดขึ้นจาก 3 สาเหตุใหญ่ๆ คือ ต้นไม้ที่ไม่สามารถลบออกได้ในขั้นตอนการลบต้นไม้เนื่องจากต้นไม้เมื่อนำมาเข้าขั้นตอนหาเส้นตรงโดยฮัฟทรานฟอรั่มแล้วจะทำให้หาเส้นตรงได้ไม่ถูกต้องและเกิดสิ่งรบกวน (noise) จำนวนมาก พื้นที่การเกษตรที่ถูกเผาจะทำให้ขั้นตอนการตรวจหาสี

ของอาคารตรวจพบเป็นสีน้ำตาลและการสะท้อนของน้ำจะทำให้ขั้นตอนการตรวจหาสีของอาคารตรวจพบเป็นสีขาว ค่าผิดพลาดเชิงลบมี 3 สาเหตุ คือ อาคารที่อยู่ไกล มีวัตถุมาบังมากเกินไปและตรวจหาส่วนอาคารได้น้อยเกินไป ซึ่งจากสามข้อที่กล่าวมานี้ล้วนทำให้ขั้นตอนการการโปรเจคชันไม่สามารถสรุปได้ว่าชิ้นส่วนนั้นคืออาคารหรือสิ่งรบกวนเนื่องจากมีรายละเอียดน้อยเกินไป ส่วนค่าความผิดพลาดเชิงบวกนั้น ผู้วิจัยมีความเห็นว่าสามารถลดค่าความผิดพลาดนี้ได้โดยเพิ่มการเรียนรู้ของเครื่องจักรเข้าไป โดยได้นำผลที่ได้จากขั้นตอนการประมวลผลภาพดิจิทัลมาวิเคราะห์รู้จำผ่านโครงข่ายประสาทเทียมแบบคอนโวลูชันนัลเพื่อระบุว่าสิ่งที่ตรวจพบนั้นใช่อาคารจริงหรือไม่ โดยได้ออกแบบโมเดลไว้ทั้งหมด 9 โมเดล รายละเอียดได้อธิบายไว้ในหัวข้อ 4.6 ซึ่งโมเดล 5 และ 7 มีความถูกต้องในการรู้จำเฉลี่ยร้อยละ 98.0 ซึ่งสามารถกำจัดความผิดพลาดเชิงบวกในขั้นตอน 4.5.3 ได้และยังช่วยยืนยันความถูกต้องให้กับการตรวจหาอาคารอีกด้วยดังที่แสดงในหัวข้อ 4.7



บทที่ 5

สรุปผลและข้อเสนอแนะ

5.1 สรุปผล

ในงานวิจัยนี้ได้เสนอวิธีการตรวจหาและรู้จำอาคารจากภาพถ่ายข้างทางในชนบท ภาพที่นำมาทดสอบเป็นภาพที่มาจากพื้นที่ชนบทของประเทศไทยซึ่งถ่ายภาพโดยกูเกิลสตรีทวิว มีขนาด 640 พิกเซล x 600 พิกเซล โดยการประมวลผลภาพดิจิทัลประกอบด้วย การกำจัดสีของต้นไม้อกจากภาพถ่าย การกำจัดสีท้องฟ้าออกจากภาพถ่าย การแบ่งส่วนภาพสีด้วยวิธี Automatic seeded region growing การหาเส้นตรงในภาพถ่ายด้วยฮัฟฟทรานสฟอร์มและการหาบริเวณที่สนใจด้วยวิธีการโพรเจกชัน นอกจากนี้ยังได้มีการใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันนัลเพื่อใช้ในการรู้จำบริเวณที่สนใจเพื่อยืนยันว่าภาพนั้นมีอาคารอยู่หรือไม่

ภาพสำหรับการฝึกโครงข่ายประสาทเทียมใช้ทั้งหมด 7,343 ภาพ แบ่งเป็นภาพที่มีอาคาร 3,995 ภาพและภาพที่ไม่มีอาคาร 3,348 ภาพ ภาพสำหรับทดสอบทั้งหมด 3,936 ภาพ แบ่งเป็นภาพที่มีอาคาร 1,832 ภาพและเป็นภาพที่ไม่มีอาคาร 2,104 ภาพ ผลการทดลองได้ค่าความถูกต้องเฉลี่ยร้อยละ 87.5 สำหรับภาพที่มีอาคารและเฉลี่ยร้อยละ 97.6 สำหรับภาพที่ไม่มีอาคาร

วิธีการที่เสนอนี้ใช้ได้ดีกับอาคารในมุมมองจากถนนที่มีมุมมองของเลนส์และความละเอียดของภาพเช่นในกูเกิลสตรีทวิว แต่ยังคงมีความผิดพลาดในการตรวจหาอาคารที่อยู่ไกลและอาคารที่ถูกบดบังด้วยต้นไม้ เนื่องจากรายละเอียดของอาคารที่ได้รับน้อยเกินไปทำให้โปรแกรมไม่สามารถตรวจหาอาคารได้ นอกจากนี้ยังมีความผิดพลาดจากรอยต่อของภาพต้นฉบับจากกูเกิลสตรีทวิวทำให้ขั้นตอนการตรวจหาเส้นตรงของอาคารไม่ถูกต้อง

5.2 ข้อเสนอแนะ

จากการทดลองที่เสนอในงานวิจัยนี้พบว่ายังมีสิ่งที่สามารถปรับปรุงเพื่อให้การทำงานดีขึ้นได้ ดังนี้

1. การประมวลผลโดยรวมค่อนข้างช้าอันเนื่องมาจากการแบ่งส่วนภาพสีโดยใช้ Automatic seeded region growing ซึ่งใช้ในการแบ่งส่วนภาพบริเวณของสีและปรับค่าความใกล้เคียงของสี ถ้าสามารถลดเวลาในส่วนนี้ได้ ก็จะสามารถทำให้ประมวลผลโดยรวมเร็วขึ้น
2. ในส่วนของการตรวจหาอาคารด้วยวิธีการประมวลผลภาพดิจิทัล เช่น การหาเส้นตรงยาวในภาพถ่าย การหาเส้นตรงแนวตั้งซึ่งขนานกัน การตรวจหาสีในภาพถ่ายและการหาจั่วหลังคาในภาพถ่าย มีการกำหนดค่าต่างๆไว้ตายตัว ซึ่งมีผลทำให้ไม่สามารถตรวจหาภาพอาคารบางภาพที่ค่าอยู่นอกเหนือตามที่ตั้งไว้ได้ หากมีการปรับค่าให้เหมาะสมมากขึ้นหรือเพิ่มวิธีให้

ตรวจหาอาคารได้เหมาะสมมากขึ้นก็จะทำให้ได้ผลการทดลองถูกต้องมากขึ้น อีกวิธีการหนึ่งคือการกำหนดค่าให้ช่วงพารามิเตอร์กว้างขึ้นเพื่อตรวจหารายละเอียดส่วนของอาคารได้มากขึ้น ซึ่งวิธีการนี้จะส่งผลให้มีโอกาสพบสิ่งแปลกปลอม (noise) มากขึ้นเช่นกัน แต่ก็อาจใช้โครงข่ายประสาทเทียมเพื่อระบุส่วนที่ตรวจหาว่าใช่อาคารหรือไม่

3. ภาพที่ได้จากภูเกิลสตรีทวิว บางภาพจะมีรอยต่อของภาพที่เกิดจากภูเกิลซึ่งเป็นในงานวิจัยนี้หากมีวิธีการที่สามารถทำให้รอยต่อเหล่านั้นถูกต้องก็จะทำให้ได้ความถูกต้องเพิ่มมากขึ้นได้
4. การนำวิธีการในงานวิจัยนี้ไปใช้กับภาพที่ได้จากกล้องอื่น ควรเป็นภาพจากกล้องที่มีมุมมอง (Viewing Angle) เท่ากับหรือใกล้เคียงกับกล้องถ่ายภาพ Ladybug[®] 2 และมีความละเอียด 640 พิกเซล x 600 พิกเซล หากใช้กล้องที่มีมุมมองกว้างกว่า Ladybug[®] 2 ก็จำเป็นต้องปรับภาพเพื่อให้มุมมองเท่ากับกล้อง Ladybug[®] 2 พร้อมทั้งลดความละเอียดลงให้เป็น 640 พิกเซล x 600 พิกเซล นอกจากนี้หากตำแหน่งกล้องถ่ายภาพสูงจากพื้นดิน ไม่เท่ากับ 2.50 ม. การลบส่วนล่างของภาพและการลบท้องฟ้า อาจต้องมีการกำหนดค่าพารามิเตอร์ใหม่ในหัวข้อ 4.3 เพื่อความเหมาะสม

รายการอ้างอิง

ภาษาไทย

กรมโยธาธิการทหารบก (2530) การขึ้นทะเบียนประวัติอาคารและสิ่งก่อสร้าง.

ราชบัณฑิตยสถาน (2542). ความหมายของคำว่าอาคาร. พจนานุกรม ฉบับราชบัณฑิตยสถาน.

ศาลปกครอง (2522). คดีพิพาทเกี่ยวกับการที่เจ้าหน้าที่ของรัฐกระทำการโดยไม่ชอบด้วยกฎหมาย. ศาลปกครองสูงสุด.

สำนักงานคณะกรรมการกฤษฎีกา (2522). พระราชบัญญัติควบคุมอาคาร. สภานิติบัญญัติแห่งชาติ.

ภาษาอังกฤษ

Adams, R. and L. Bischof (1994). "Seeded Region Growing." IEEE Transactions on Pattern Analysis and Machine Intelligence 16(6): 641–647.

Alex Krizhevsky, et al. (2012). "Imagenet Classification With Deep Convolutional Neural Networks." Advances in neural information processing systems: 1097-1105.

Cheng, H. D., et al. (2001). "Color image segmentation: advance and prospects." Pattern Recognition 34: 2259-2281.

Christophe Garcia and G. Tziritas (1999). "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis." IEEE Transaction on Multimedia 1(3): 264 - 277.

Colantoni, P. (2012). "YCbCr transformations." from <http://www.couleur.org/index.php?page=transformations#YCbCrf>.

David, P. (2008). "Detection of Building Facades in Urban Environments." Proc. of SPIE 6978.

Duda, R. O. and P. E. Hart (1972). Use of the Hough Transformation to Detect Lines and Curves in Pictures. Communications of the ACM. NY,USA, ACM New York. 15: 11-15.

George H Joblove and D. Greenberg (1978). "Color Spaces for Computer Graphics." ACM siggraph computer graphics 12(3): 20-25.

Gonzalez R.C. and W. R.E Line Hough Transform, Prentice Hall.

J.Xiao, et al. (2010). Automatic Detection of Buildings with Rectangular Flat Roofs from Multi-View Oblique Imagery. ISPRS Archives. Saint-Mandé, France. 38.

Jianping Fan, et al. (2001). "Automatic Image Segmentation by Integrating Color-Edge Extraction And Seeded Region Growing." IEEE Trans. On Image Processing 10: 1454-1466.

Lecun, Y. "Invariance." 2015, from <http://yann.lecun.com/exdb/lenet/>.

LISALab (2015). "Convolutional Neural Networks (LeNet)." from <http://deeplearning.net/tutorial/lenet.html>.

Martin Drauschke and Wolfgang Förstner (2007). Selecting Appropriate Features For Detecting Buildings And Building Parts. ISPRS Archives. Beijing, China. 38.

Milos Miljanovic, et al. (2012). "Detection of Windows in Facades using Images Processing Algorithms." Indian Journal of Computer Science and Engineering 3(4): 539-547.

R. Haralick and L. Shapiro (1992). Computer and Robot Vision, Addison-Wesley

Rafael C. Gonzalez and Richard E. Woods (2009). Thresholding. Digital image processing, Pearson Education.

Rafael C. González and R. E. Woods (2007). Normalization. Digital Image Processing, Prentice Hall: 85.

Tarun Kumar and K. Verma (2010). "A Theory Based on Conversion of RGB image to Gray image." International Journal of Computer Applications 7(2): 7-10.

Vincent Brandou and Caroline Baillard (2011). Gable Roof Detection in Terrestrial Images. ISPRS Conference, PIA 2011: 97-108.

Viraj Kulkarni, et al. (2009) Window Detection in Frontal Facades.

Xiuyun Wu and Yan Li (2012). Integrated Method of Building Edge Line Feature Extraction from Aerial Imagery and Digital Surface Model. 18th International Conference Geoinformatics. Beijing, China.

Zhizhong Kang, et al. (2007). "Automatic Detection of Range Variance of Facades from Vehicle Based Image Sequence." ISPRS XXXVI-5/C55.





ภาคผนวก

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ภาคผนวก ก
ภาพตัวอย่างผลลัพธ์ที่ได้จากการประมวลผลภาพดิจิทัล

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ผลลัพธ์ที่ได้จากการทดลองในสภาวะแวดล้อมจริงในพื้นที่ชนบทของประเทศไทยโดยใช้ภาพจากด้านข้างรถสำรวจของกูเกิ้ลสตรีทวิว อาคารที่ตรวจพบจะถูกวาดรูปสี่เหลี่ยมล้อมรอบ ซึ่งวิธีการในงานวิจัยนี้สามารถตรวจหาได้มากกว่าหนึ่งอาคารในหนึ่งภาพ ภาพตัวอย่างผลลัพธ์แสดงในรูปแบบที่ ก-1



(1)



(2)



(3)



(4)



(5)



(6)

รูปที่ ก-1 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง



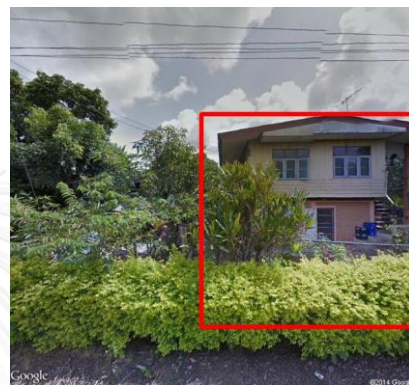
(7)



(8)



(9)



(10)



(11)



(12)

รูปที่ ก-2 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ)



(13)



(14)



(15)



(16)



(17)



(18)

รูปที่ ก-3 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ)



(19)



(20)



(21)



(22)



(23)



(24)

รูปที่ ก-4 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ)



(15)



(26)



(27)



(28)



(29)



(30)

รูปที่ ก-5 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ)



(31)



(32)



(33)



(34)



(35)



(36)

รูปที่ ก-6 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ)



(37)



(38)



(39)



(40)



(41)



(42)

รูปที่ ก-7 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ)



(43)



(44)



(45)



(46)



(47)



(48)

รูปที่ ก-8 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ)



(49)



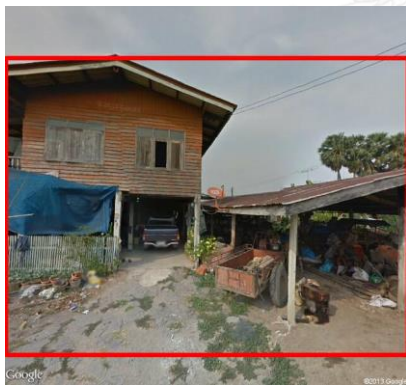
(50)



(51)



(52)



(53)



(54)

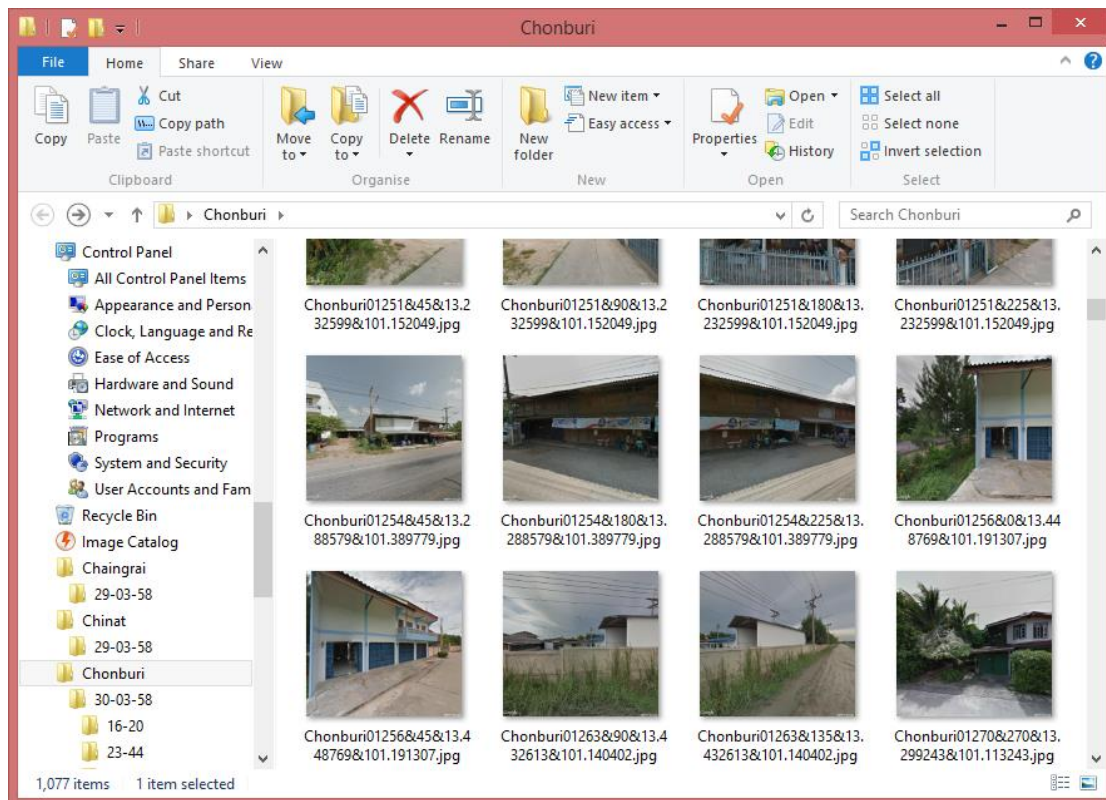
รูปที่ ก-9 ตัวอย่างผลลัพธ์ที่ได้จากการทดลอง(ต่อ)

ภาคผนวก ข

ตัวอย่างการใช้ประโยชน์จากผลลัพธ์ที่ได้จากการประมวลผลภาพดิจิทัล

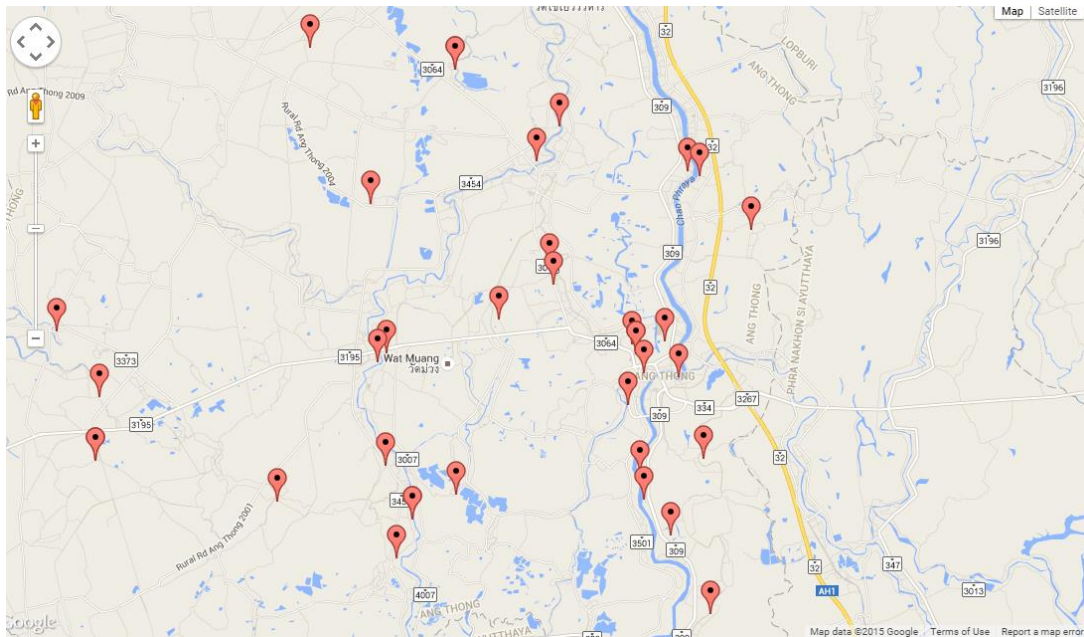
จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ภาพถ่ายข้างทางในชนบทในงานวิจัยนี้ นำมาจาก Google Street View ซึ่งมีการเก็บค่าพิกัดจีพีเอสไว้ในชื่อของภาพ เช่น Chonburi01123&270&13.445759&101.065782.jpg เป็นต้น โดยข้อมูลที่ได้รับจะมีชื่อจังหวัดที่ถ่ายภาพ ลำดับของภาพ องศาของกล้องที่ถ่าย ค่าละติจูดและลองจิจูด โดยแต่ละค่าจะถูกคั่นด้วยเครื่องหมาย “&” ดังรูปที่ ข-1



รูปที่ ข-1 ตัวอย่างชื่อภาพจากกูเกิลสตรีทวิวที่นำมาใช้ในการทดลอง

ภาพที่ผ่านการประมวลผลแล้ว หากพบว่ามีการอยู่จะถูกแยกไฟล์เตอร์ออกจากภาพที่ไม่มีอาคาร ซึ่งสามารถนำภาพทั้งหมดในไฟล์เดออร์นี้มาวาดพิกัดลงในแผนที่ก็จะสามารถเป็นประโยชน์ต่อหน่วยงานเกี่ยวข้องที่ต้องการข้อมูลได้ ดังแสดงในรูปที่ ข-2



รูปที่ ข-2 ตัวอย่างการใช้ประโยชน์หลังการประมวลผลภาพถ่ายดิจิทัลโดยการวาดพิกัดจีพีเอสลงในแผนที่ของกูเกิล

ภาคผนวก ค

ส่วนเชื่อมต่อผู้ใช้ของโปรแกรมตรวจหาบริเวณที่สนใจในภาพถ่ายข้างทางในชนบท

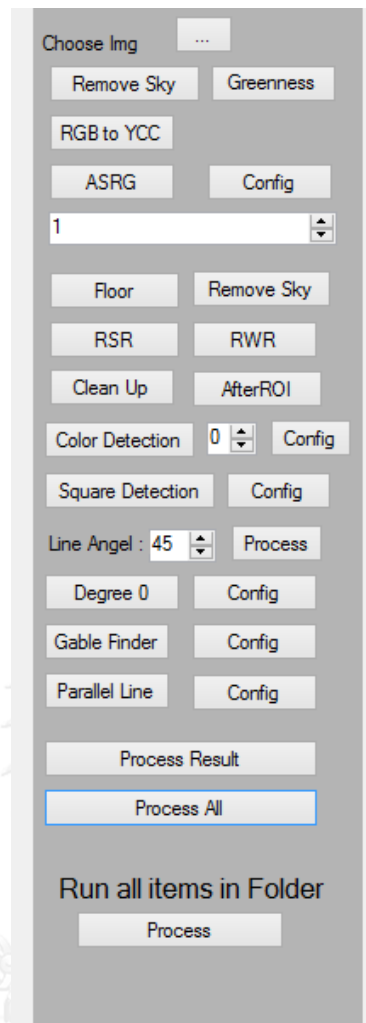
จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

โปรแกรมตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท พัฒนาขึ้นโดยใช้ Visual Studio 2013 โดยใช้ภาษา C# ในการเขียน ซึ่งนำไลบรารี EMGUCV ซึ่งเป็นไลบรารีเกี่ยวกับการประมวลผลภาพดิจิทัลมาประยุกต์ใช้ โดยโปรแกรมตรวจหาอาคารจากภาพถ่ายข้างทางในชนบทนี้สามารถดำเนินการลบท้องฟ้า ลบสีของต้นไม้ แบ่งส่วนบริเวณภาพด้วยสี ลบส่วนที่ไม่ต้องการในภาพออก ตรวจสอบเส้นตรง ตรวจสอบสีของอาคารและการโปรเจกชันได้ หน้าเชื่อมต่อผู้ใช้ของโปรแกรมมีลักษณะดังรูปที่ ค-1

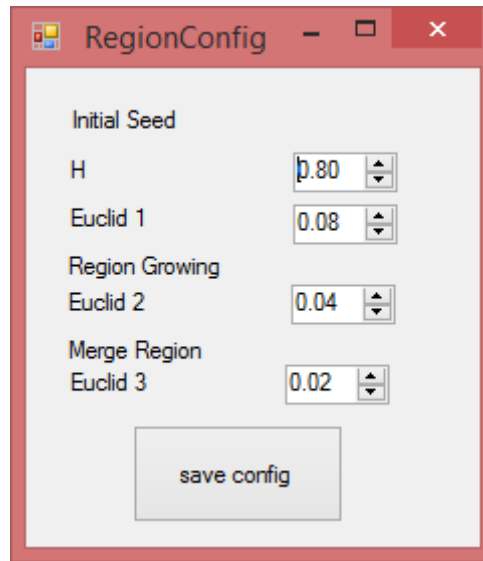


รูปที่ ค-1 หน้าเชื่อมต่อผู้ใช้ของโปรแกรมตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท

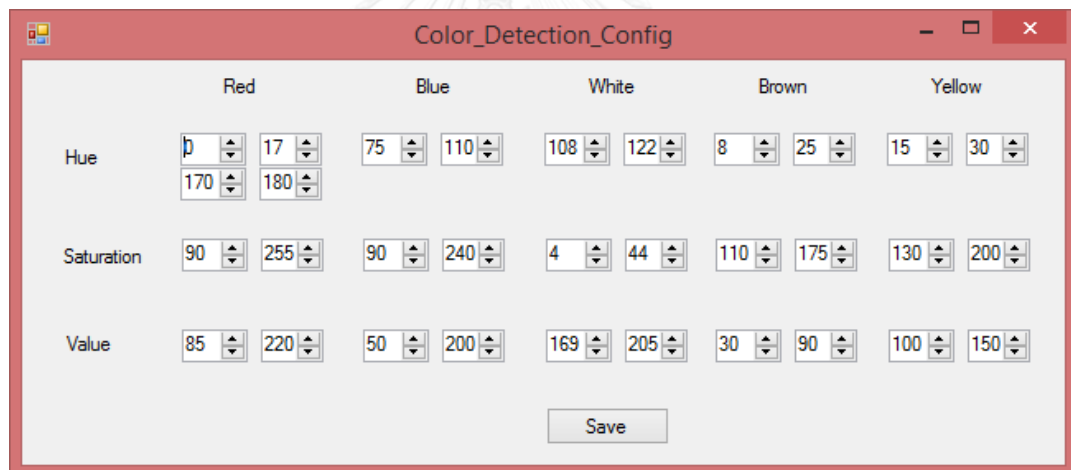
ในส่วนของแผงควบคุมโปรแกรมจะสามารถเลือกประมวลผลทั้งแบบทีละขั้นตอนหรือแบบทุกขั้นตอนได้ โดยยังสามารถเลือกได้อีกว่าจะประมวลผลหนึ่งภาพหรือทั้งโฟลเดอร์



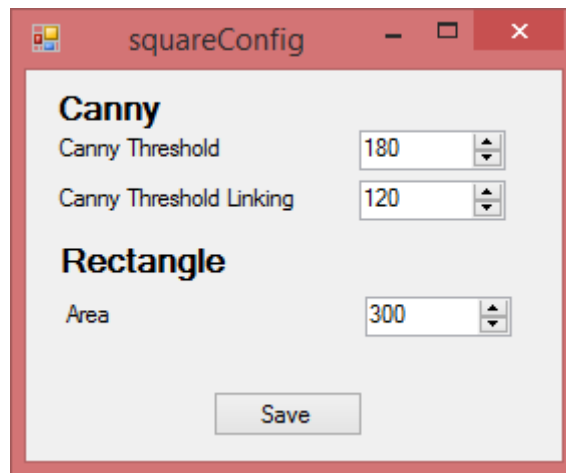
รูปที่ ค-2 แผงควบคุมโปรแกรมตรวจหาอาคารจากภาพถ่ายข้างทางในชนบท
 จากรูปที่ ค-2 จะเห็นได้ว่าในรูปภาพมีปุ่ม “Config” เพื่อตั้งค่าพารามิเตอร์ในแต่ละขั้นตอน โดย
 สามารถปรับได้ทั้งค่า *Automatic Seeded Region Growing*, การตรวจหาสีของอาคาร, การ
 ตรวจหาสี่เหลี่ยม, การตรวจหาเส้นตรงยาวในแนวนอน, การตรวจหาจั่วหลังคาและการตรวจหา
 เส้นตรงที่ขนานกันในแนวตั้ง หน้าเชื่อมต่อนักใช้ของแต่ละขั้นตอนมีแสดงในรูปที่ ค-3 ถึง ค-8



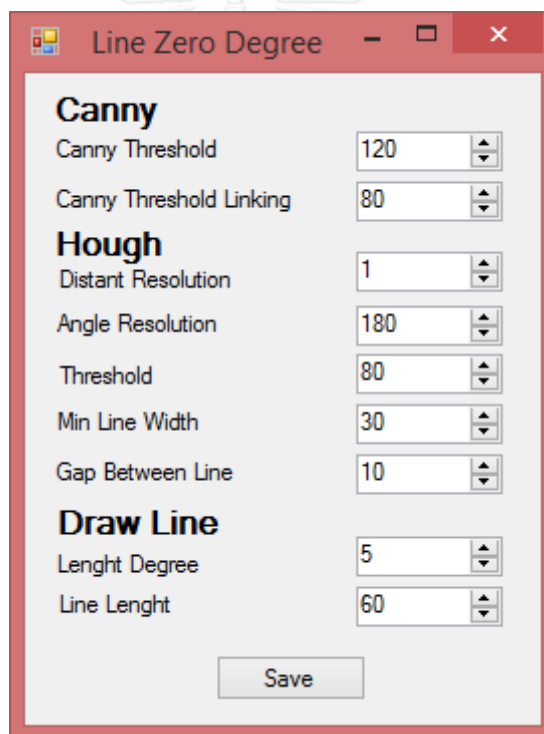
รูปที่ ค-3 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์ของขั้นตอน Automatic Seeded Region Growing



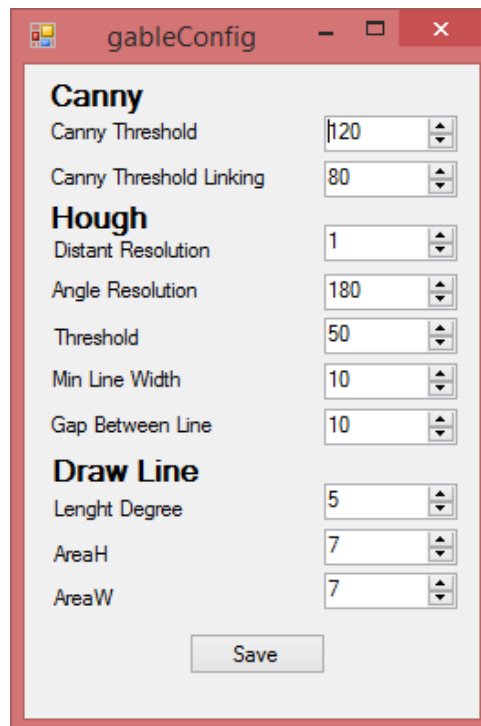
รูปที่ ค-4 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์ของขั้นตอนการตรวจหาสีของอาคาร



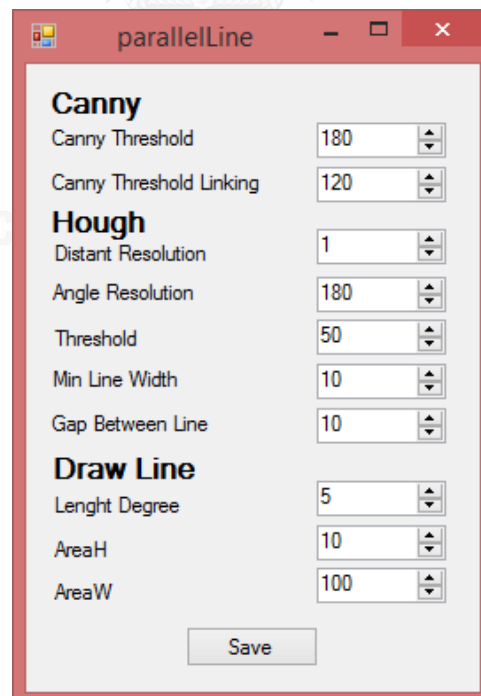
รูปที่ ค-5 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์ของขั้นตอนการตรวจหาสี่เหลี่ยมในภาพถ่าย



รูปที่ ค-6 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์ของขั้นตอนการตรวจหาเส้นตรงยาวในแนวราบ



รูปที่ ค-7 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์ของขั้นตอนการหาจั่วหลังคาของอาคาร



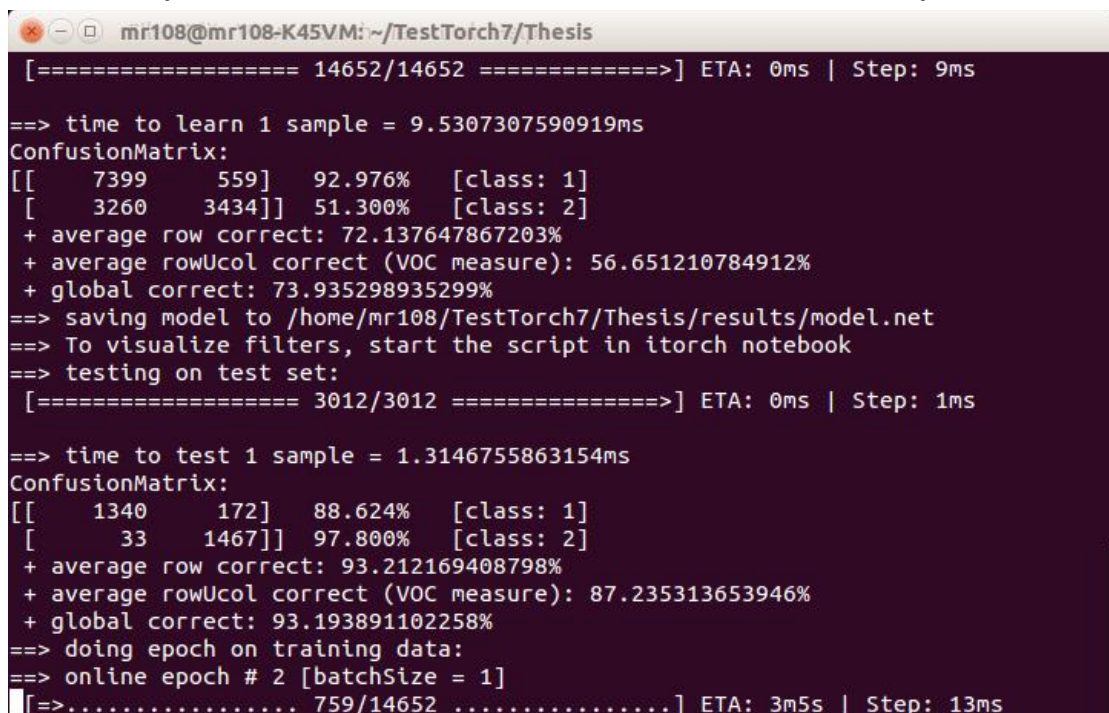
รูปที่ ค-8 หน้าเชื่อมต่อผู้ใช้เพื่อปรับพารามิเตอร์การหาเส้นตรงที่ขนานกันในแนวตั้ง

ภาคผนวก ง

ส่วนเชื่อมต่อผู้ใช้ของโปรแกรมรู้จำอาคารจากภาพถ่ายข้างทางในชนบท

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

หลังจากที่ได้ภาพผลลัพธ์จากขั้นตอนการตรวจหาบริเวณที่สนใจจากภาพถ่ายข้างทางในชนบทแล้วต่อไปจะเข้าสู่กระบวนการรู้จำอาคาร ซึ่งขั้นตอนนี้ได้ทำการทดลองใน Ubuntu 10.4.14 โดยใช้ภาษา LuaJit และใช้ไลบรารี Torch7 ซึ่งเป็นไลบรารีเกี่ยวกับการเรียนรู้ของเครื่องจักร (Machine Learning) ที่มีประสิทธิภาพสูงในการประมวลผลภาพดิจิทัล โดยในงานวิจัยนี้ได้ใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันนัลในการประมวลผลภาพดิจิทัล ตัวอย่างส่วนเชื่อมต่อผู้ใช้ของโปรแกรมรู้จำอาคารจากภาพถ่ายข้างทางในชนบทจะเป็น console based ดังรูปที่ ง-1



```

mr108@mr108-K45VM:~/TestTorch7/Thesis
[===== 14652/14652 =====>] ETA: 0ms | Step: 9ms
==> time to learn 1 sample = 9.5307307590919ms
ConfusionMatrix:
[[ 7399  559]  92.976%  [class: 1]
 [ 3260 3434]]  51.300%  [class: 2]
+ average row correct: 72.137647867203%
+ average rowUcol correct (VOC measure): 56.651210784912%
+ global correct: 73.935298935299%
==> saving model to /home/mr108/TestTorch7/Thesis/results/model.net
==> To visualize filters, start the script in itorch notebook
==> testing on test set:
[===== 3012/3012 =====>] ETA: 0ms | Step: 1ms

==> time to test 1 sample = 1.3146755863154ms
ConfusionMatrix:
[[ 1340  172]  88.624%  [class: 1]
 [   33 1467]]  97.800%  [class: 2]
+ average row correct: 93.212169408798%
+ average rowUcol correct (VOC measure): 87.235313653946%
+ global correct: 93.193891102258%
==> doing epoch on training data:
==> online epoch # 2 [batchSize = 1]
[=>..... 759/14652 .....] ETA: 3m5s | Step: 13ms

```

รูปที่ ง-1 ส่วนเชื่อมต่อผู้ใช้ของโปรแกรมรู้จำอาคารจากภาพถ่ายข้างทางในชนบท



บทความเรื่อง “Building Detection from Terrestrial Images” เสนอในงานประชุมวิชาการ 2015 The 7th International Conference on Future Computer and Communication (ICFCC 2015) ที่ Nanyang Executive Centre, Nanyang Technological University (NTU) ประเทศสิงคโปร์ ในช่วงวันที่ 20-22 พฤษภาคม ปี พ.ศ. 2558 บทความนี้ได้รับการตอบรับให้ตีพิมพ์ไว้ใน Journal of Image and Graphics



BUILDING DETECTION FROM TERRESTRIAL IMAGES

Teerapat Chaloeivoot and Suebskul Phiphobmongkol
 Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand
 Email: Teerapat.Ch@student.chula.ac.th, suebskul.p@chula.ac.th

Abstract— This paper presents a method to detect buildings in terrestrial images. High resolution terrestrial images are normally taken from land survey vehicles. These images and other surveyed data along roads are needed by many agencies that require new data as time passes by. Land use in rural area is an example that needs information about buildings and can benefit from terrestrial images. The proposed method was aimed to detect buildings in terrestrial images to benefit the above needs. The method consists of two stages. The first stage removes unwanted objects, performs image segmentation, and finds regions of interest. Image processing techniques such as greenness extraction, sky detection, color segmentation, color detection, shape detection are used. The second stage performs building detection. It includes the possible building parts detection, projection profiles finding, and the building determination. The method can identify a partial building if the whole building is not shown in an image. The proposed method was tested on 936 images (332 images with buildings and 604 images without buildings). The images were from Google Street View. The accuracy was determined by human inspection. The method gave promising results with an average accuracy of 82.5%. Positive faults were 4.7% average.

Index Terms—Building Detection, Terrestrial image, Image Processing, Image Segmentation

I. INTRODUCTION

Land use may include living, grazing, agriculture, industries. Many agencies need information about land use for many purposes, such as, taxation, community growth measurement, new dweller area finding. Building identification is a feature that helps specifying living or industrial land use. Building identification from aerial survey photos is a widely used method. However, in Thailand, this method is not used frequently due to the high cost of survey. An option is to use human surveyors to survey the area. This option requires many surveyors, and takes a long time. Another option is to use a survey vehicle that can take high-resolution photos with GPS coordinates and look for buildings later on computer. This method is considered more efficient and save time. At present, there are agencies that operate such vehicles. If a computer can detect buildings in the surveyed images there would be a new source of land use information for the relevant agencies.

The proposed method was aimed to find buildings or building parts in real environment images. Most buildings have specific elements in common, such as, roof, gables, doors, windows, and pillars. Difficulties include the detection of an incomplete structures or shapes of the buildings, building obstructed by other objects. Algorithms based on color detection, Line Hough transform, automatic seeded region growing were used to find potential features that can help specify a building. In the building determination process, it considers detected building parts such as gables, long horizontal lines, parallel vertical lines, rectangles, unnatural colors by using rules described later in the paper.

II. RELATED WORK

Most researches of building detection from terrestrial images detected façade, windows, and plane. Zhizhong, Sisi, and Ben (2005) [1] introduced an automatic detection of range variance of facades by projective difference of corresponding points on a common projective plane. Therefore, the algorithm could separate building façade from cross road. Haider et al. (2007) [2] presented windows detection in the urban environment based on Adaboost to optimize cascaded classifier for detection. Martin and Wolfgang (2009) [3] applied Adaboost classification framework to classify feature of building parts. This research focused on façade, roof, windows and windowpane. Viraj, Rohan and Hong (2009) [4] described a method to identify windows from a building frontal façade by projection profiles and snake algorithm. Vincent and Caroline (2011) [5] presented a method for gable roof detection from terrestrial images. Milos and Thomas (2012) [6] used image-processing to detect windows in facades at several different orientations and scales.

An approach that detect buildings from aerial images used LIDAR data and image processing techniques [7]. Others used image-processing such as Antonis and Hichem (2008) [8], Xiuyun and Yan(2008) [9], Qiongchen and Zhiguo (2009) [10], Masoud and Parvaneh (2009) [11] and Melissa and Parvaneh (2012) [12]

III. PROPOSED METHOD

The proposed method has two stages. The first stage removes unwanted objects, performs image segmentation

using automatic seeded region growing and performs region analysis. The second stage determines whether the image contains a building or not. It uses the result from building-part extraction and projection profiles. The system overview of the methodology is shown in Fig. 1.

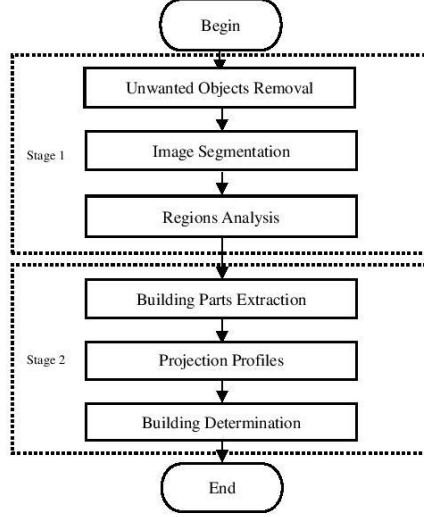


Figure 1. System Overview

A. Unwanted Object Removal

This process was intended to remove the objects that are not part of a building, e.g. trees, sky. The result will then be appropriate for image segmentation. Two steps are described as follows:

Tree removal: This step was used to remove trees by detecting green color of trees by using HSV color space [13][14]. The transformation of RGB to HSV performed using the following equation (1) to (6), where R, G, and B are in the range [0, 1], H is [0°, 360°], and S and V are [0, 100]:

$$\max = \max(R, G, B) \quad (1)$$

$$\min = \min(R, G, B) \quad (2)$$

$$\text{Chroma} = \max - \min \quad (3)$$

Equation (1) to (3) calculated Chroma value. Then, compute HSV value in (4), (5) and (6).

$$H = \begin{cases} 0^\circ, & \text{if Chroma} = 0 \\ 60^\circ \times \frac{G - B}{\text{Chroma}} \bmod 6, & \text{if max} = R \\ 60^\circ \times \frac{B - R}{\text{Chroma}} + 2, & \text{if max} = G \\ 60^\circ \times \frac{R - G}{\text{Chroma}} + 4, & \text{if max} = R \end{cases} \quad (4)$$

$$S = \begin{cases} 0, & \text{max} = 0 \\ \frac{\text{Chroma}}{\text{max}}, & \text{max} \neq 0 \end{cases} \quad (5)$$

$$V = \max \quad (6)$$

The proper range of Hue to detect green color of tree was 40-120. Saturation was 19-100. Value was 0-66.

Sky Removal: In landscape images, sky is mostly brighter than objects on the ground. Automatic Threshold method [15][16] was used to separate the sky in the images. From the experiment, the proper range of threshold values was 130-150. The sky removal method can be performed using equation (7) where T is threshold value, src(x,y) is horizontal and vertical addresses of source image and dst(x,y) is horizontal and vertical addresses of destination image.

$$\text{dst}(x, y) = \begin{cases} 0, & \text{if } y > \text{Image height} \div 2 \\ 0, & \text{if } \text{src}(x, y) > T \\ \text{src}(x, y), & \text{otherwise} \end{cases} \quad (7)$$

B. Image Segmentation

This process segmented each area of an image according to colors. The algorithm was automatic seeded region growing as used by Shih and Cheng (2005) [17]. This algorithm was developed from Seeded Region Growing which was presented earlier by Adams and Bischof (1994) [18]. $Y C_b C_r$ color space was used in Shih and Cheng method. The RGB to $Y C_b C_r$ conversion performed using (8), where R, G, and B are in the range [0, 1], Y is [16, 235], and C_b and C_r are [16, 240]:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 65.481 & 128.553 & 24.996 \\ -39.797 & -74.203 & 112 \\ 112 & -39.786 & -14.214 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix} \times \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (8)$$

The process began with initial seeds. The seed selection determined 2 conditions. First condition is a similarity seed pixel checking. It was considering a 3x3 neighborhood. It started with calculates a standard deviation in (9):

$$\sigma_x = \sqrt{\frac{1}{9} \sum_{i=1}^9 (x_i - \bar{x})^2} \quad (9)$$

where x are Y, C_b , C_r and \bar{x} is mean value of 3x3 neighborhood. The similarity of pixel to its neighbor is defined as

$$H = 1 - \left(\frac{\sigma_Y + \sigma_{C_b} + \sigma_{C_r}}{\sigma_{\max}} \right) \quad (10)$$

where σ_{\max} is maximum standard deviation in the image. A similarity value (10) must be higher than a threshold value. From our experiment, a proper similarity threshold value is 0.80.

Second condition is a relative Euclidean distance to its neighbor checking. The formula is as follow:

$$d_i = \frac{\sqrt{(Y-Y_i)^2 + (C_b - C_{b_i})^2 + (C_r - C_{r_i})^2}}{\sqrt{Y^2 + C_b^2 + C_r^2}} \quad (11)$$

where $i=1,2,3,\dots,8$. The relative Euclidean distance in (11) is used to find maximum distance in (12):

$$d_{max} = \max_{i=1}^8(d_i) \quad (12)$$

The maximum relative Euclidean distance to its neighbor must be less than a threshold value. A proper threshold value from our experiment is 0.08

A pixel that met two conditions above is classified as seed pixel. Next, the growing method started with labeling seeded pixel. Then, unclassified pixels were calculated a relative Euclidean distance to its adjacent regions in (13):

$$d_i = \frac{\sqrt{(\bar{Y}_i - \bar{Y})^2 + (C_{b_i} - C_b)^2 + (C_{r_i} - C_r)^2}}{\sqrt{\bar{Y}_i^2 + C_{b_i}^2 + C_{r_i}^2}} \quad (13)$$

where $(\bar{Y}, \bar{C}_b, \bar{C}_r)$ are average value of Y, C_b, C_r information in that region. The pixel merged into adjacent minimum relative Euclidean distance region. This step performed until no unclassified pixel left. Next, the region similarity checking performed using (14):

$$d(R_i, R_j) = \frac{\sqrt{(\bar{Y}_i - \bar{Y}_j)^2 + (C_{b_i} - C_{b_j})^2 + (C_{r_i} - C_{r_j})^2}}{\min(\sqrt{\bar{Y}_i^2 + C_{b_i}^2 + C_{r_i}^2}, \sqrt{\bar{Y}_j^2 + C_{b_j}^2 + C_{r_j}^2})} \quad (14)$$

The region merged into adjacent minimum relative Euclidean distance region. The minimum relative Euclidean distance must be less than threshold value. We used 0.04 as threshold value. Final step is merged small region size. Our threshold value described below:

$$\text{Threshold} = \frac{\text{image height} \times \text{image width}}{150} \quad (15)$$

where 150 was obtained from experiment. The region that size smaller than (15) merged into adjacent minimal distance region. The threshold of minimal distance Euclidean distance of adjacent region from our experiment must be higher than 0.02.

C. Regions Analysis

This process culled out regions that are unlikely to contain buildings. It involves three steps: noise removal, left-over tree removal, and low-region removal.

Noise removal: The tiny regions would be removed because they were too small for the detection and can be burden or cause error in determination process. In our experiment, if the region size was less than 500 pixels, it will be eliminated.

Left-over tree removal: From trees removing step, it cannot remove all green color because of several reasons such as light, various color of leaves. Removal occurred if the number of white pixels is greater than 30% (from the experiment). This method will wipe out the left-over trees regions.

Low-region removal: Since the source images were taken from a survey vehicle, the bottom parts were mostly road, grass, or wall. Eliminating such region helps reduce error and burden to the process.

After all, unwanted objects in original image (Fig. 2a) are mostly eliminated. The result presents in Fig. 2b.



(a) Original image (b) Pre-processing results

Figure 2. Pre-processing results before building parts extraction process

D. Building Parts Extraction

Color Detection: The colors: red, blue, brown, yellow and white are widely used as building paints which are clearly distinct from nature. Each color will be evaluated by size and position. Red and blue detected only the top half of images for checking rooftop. Brown, yellow and white are detected in the middle of images for checking body of building. The proper HSV color ranges for building detection are shown in Table I:

TABLE I. THE PROPER HSV COLOR RANGES FOR BUILDING DETECTION

Color	H	S	V
Red	0-34, 340-360	35-100	33-86
Blue	150-200	35-94	19-78
Brown	216-244	1-44	66-80
Yellow	16-50	10-17	11-35
White	30-60	19-78	39-58

Line Detection: Structure of most buildings contains rectangles. In the real environment, it is hard to find perfect shapes because of obstacles such as trees. This step uses the benefits of Line Hough Transform [19] to help finding the straight edges of buildings. The angle finding formula (16) is as follows:

$$\text{Angle (in degree)} = -\left(\tan^{-1}\left(\frac{y_2 - y_1}{x_2 - x_1}\right) \times \left(\frac{180}{\pi}\right)\right) \quad (16)$$

where y_1, y_2, x_1, x_2 are pixel coordinates at the tips of a line.

Line detection can be categorized into four cases below:

Horizontal Line: Detection of long horizontal line(s). The line length must be at least 100 pixels.

Parallel Vertical lines: Detection of a pair of parallel vertical lines whose centers are almost of the same vertical level. Each line is far from the each other more than a threshold value. Our test used a threshold value at least 100 pixels.

Rectangle Detection [20]: Detection of quadrilateral shape with four corners. This step aims to extract windows and doors. The size of quadrilateral shape must be larger than a threshold value. Our threshold value is 300 pixels.

Gable Detection: Detection of two lines with opposite angle in the same horizontal area and their top points are near each other (less than a threshold value). From our

experiment, threshold value used ± 7 pixels from top point of line.

E. Group all building parts

After performing building parts extraction process, this step superimposes all results (Fig. 3a) and use projection profile to find the height and width of each group (Fig 3b).

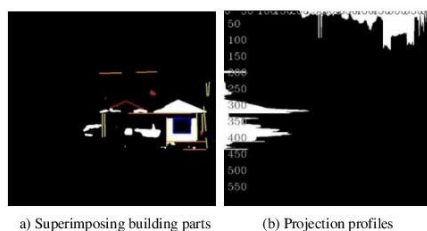


Figure 3. Superimposing and Projection profiles

F. Building Determination

To consider these elements, decision-making rules are needed. The rules use size and aspect ratio values as described below. Though these values were chosen empirically from 20 trial images at the beginning of our experiment, the values were applied to all tested images. Further investigation for more appropriate values should be done for real use.

Size: The group elements size must be at least 100 pixels x 100 pixels.

Aspect ratio: The width to height ratio of each group must be between 4:3 to 7:5.

IV. EXPERIMENTS

A. Data test sets

The location that the test images were taken was in the rural area of Thailand. The buildings had various structures and many of them were blocked out by obstacles, e.g. trees. These images were from Google Street View survey vehicles. The resolution was 640 x 600 pixels. Many kinds of building textures such as wood and cement were found. Some buildings had no facade, some were covered by trees.

B. Unwanted object removal and image segmentation

Sky, vegetation, noise, and left-over trees were removed after automatic seeded region growing method. Building parts would be left in the resulting images. The results are shown in Table II:

TABLE II. OBJECT REMOVAL AND IMAGE SEGMENTATION RESULTS

Method	Accuracy
Remove sky	94%
Remove vegetation	97%
Regions Analysis	100%

Errors were caused by dark cloud, trees with green colors which were out of the defined range. The problems can mostly be solved by the left-over tree removal.

C. Experiments on images with buildings.

Testing was performed on 100 images of no obstacles in front of buildings and 100 images of building with obstacles. The results are shown in Table III:

TABLE III. RESULTS ON IMAGES WITH BUILDING(WITH/WITHOUT OBSTACLES)

Descriptions	Accuracy
Building without obstacles	92%
Building with obstacles in front	83%

The results showed satisfied accuracy in both cases. The errors were caused by the edge joint of source images (Fig. 4a) and by no-wall building (Fig. 4b), which made buildings undetected in projection profiles.



Figure 4. Error cases

D. Experiments on mixed data set

Experiment was done on 936 images (332 images with buildings and 604 images without buildings) in the rural area environment. The results are shown in Table IV:

TABLE IV. RESULTS FROM MIXED DATASET

Descriptions	Accuracy
Images with buildings	82.5%
Images without building	95.3%

Some false positives occurred due to trees, burned agricultural area, and large water surface. Some trees were not totally removed by the prior process and created faults. The reason was the confusing Hough line detection in many angles. The burned agricultural area was detected as large brown color and water reflection detected as large white color in color detection. Thus, the situation led to mistaking the area as buildings. The results are shown in Table V:

TABLE V. FALSE POSITIVE CASES

Reasons	Number of images
Trees that were not removed in early process	21
Burned agricultural area	5

Water reflection	2
------------------	---

False negatives also occurred mostly due to objects between the cameras and the buildings. Far-away buildings were also too small to be detected. Too few building parts detected can also cause false negatives. The false negatives are shown in Table VI:

TABLE VI. FALSE NEGATIVE CASES

Reasons	Number of images
Small Building (far away)	11
Obstacle objects (trees)	39
Few building parts were detected	8

V. CONCLUSION

This proposed method can detect buildings in terrestrial images that contain whole buildings or part of them. The test images were taken in the rural area of Thailand by Google. Building detection was tested on 936 images. An average accuracy of 82.5% was achieved with average positive faults of 4.7%. Errors were from objects, such as trees that were over the buildings in images. Too small buildings and unclear building parts also caused errors. Other causes were from out-of-range tree colors, water reflection, and burned agriculture. The program, sometimes, did not detect some walls and may detect more than one building. The difficulties were from the handling of objects over buildings and the analysis of existed buildings from the detected components. However, the result showed that the proposed method could be for real use and is satisfied.

REFERENCES

- [1] Z. Kang, S. Zlatanova and B. Gorte, "Automatic detection of range variance of facades from vehicle based image sequence" in Proc. 5th International Symposium on Mobile Mapping Technology, Italy, 2007.
- [2] H. Ali, C. Seifert, N. Jindal, L. Paletta, and G. Paar, "Window detection in facades," in Proc. 14th International Conf. on Image Analysis and Processing (ICIAP), pp. 837-842, 2007.
- [3] M. Drauschke, and W. Förstner, "Selecting appropriate features for detecting buildings and building parts," in Proc. 21st Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS), Beijing, China, vol. 16, pp.447-452, 2008
- [4] V. Kulkarni, R. Nagesh, and H. Wu, "Window detection in frontal façades." Project work at CS294-69 Image Manipulation and Computational Photography, University of Barkley, 2011.
- [5] V. Brandou, and C. Baillard, "Gable roof detection in terrestrial images," in Proc. Photogrammetric Image Analysis, pp. 97-108, 2011.
- [6] M. Miljanovic, T. Eiter, and U. Egly, "Detection of windows in facades using image processing algorithms," IJCVSE Vol. 3, 2012
- [7] J. Xiao, "Automatic building outlining from multi-view oblique images," in Proc. 1st ISPRS Annu. of the Photogrammetry,

- Remote Sensing and Spatial Information Sciences, no. 3, pp. 323-328, 2012.
- [8] A. Katartzis, and H. Sahlbi, "A stochastic framework for the identification of building rooftops using a single remote sensing image," IEEE Trans. on Geoscience and Remote Sensing, vol. 46, no. 1, pp. 259-271, 2008.
- [9] X. Wu, and Y. Li, "Integrated method of building edge line feature extraction from aerial imagery and digital surface model," In Proc. 18th International Conf. on Geoinformatics, pp. 1-5, 2010.
- [10] Q. Wang, Z. Jiang, J. Yang, D. Zhao, and Z. Shi, "A hierarchical connection graph algorithm for gable-roof detection in aerial image," Geoscience and Remote Sensing Letters, vol. 8, no. 1, pp. 177-181, 2001.
- [11] M. S. Nosrati, and P. Saeedi, "A novel approach for polygonal rooftop detection in satellite/aerial imageries," in Proc. 16th IEEE International Conf. on Image Processing (ICIP), pp. 1709-1712, 2009.
- [12] M. Cote, and P. Saeedi, "Automatic rooftop extraction in nadir aerial imagery of suburban regions using corners and variational level set evolution," IEEE Trans. on Geoscience and Remote Sensing, vol. 51, no. 1, pp.313-328, 2013.
- [13] M. K. Agoston, Computer Graphics and Geometric Modeling: Implementation and Algorithms, London: Springer, pp. 300-306.
- [14] R G Kuehni, Color space and its divisions: color order from antiquity to the present, UK: Wiley, 2003.
- [15] R. C. Gonzalez, and R. E. Woods, Thresholding in Digital Image Processing, Pearson Education, 2002, pp. 595-611.
- [16] M. Sezgin, "Survey over image thresholding techniques and quantitative performance evaluation," Journal of Electronic imaging, vol. 13, no. 1, pp. 146-168, 2004.
- [17] F. Y. Shih, and S. Cheng, "Automatic seeded region growing for color image segmentation," Image and vision computing, vol.23, no. 10, pp. 877-886, September 2005.
- [18] R. Adams and L. Bischof, "Seeded region growing," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 16, no. 6, pp. 641-647, 1994.
- [19] N. Aggarwal, and W. Clem Karl, "Line detection in images through regularized Hough transform," IEEE Trans. on Image Processing, vol. 15, no. 3, pp. 582-591, 2006.
- [20] C. R. Jung, and R. Schramm, "Rectangle detection based on a windowed Hough transform," In Proc. 17th Brazilian Symposium on Computer Graphics and Image Processing, 2004, pp. 113-120.

Teerapat Chaloeivoot received his B.Eng. degree in computer engineering from Department of Engineering, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Thailand in 2011. He is currently pursuing master degree in Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand. His current research interests are image processing, vision and machine learning.

Suebskul Phiphompongkol received his B.Eng. degree in electrical engineering in 1980 and M.Sci. degree in computer science in 1982 from Chulalongkorn University, Bangkok, Thailand. He received his Ph.D. in computer science from Auburn University in 1991. His current research interests are digital image processing, intelligent transportation systems, Artificial intelligence, computer vision, real-time simulation and computer hardware.

ประวัติผู้เขียนวิทยานิพนธ์

นายธีรพัฒน์ เฉลยวุฒิ เกิดวันที่ 23 พฤษภาคม พ.ศ. 2532 ที่จังหวัดสระบุรี สำเร็จ การศึกษาระดับปริญญาตรีวิศวกรรมศาสตรบัณฑิต ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะ วิศวกรรมศาสตร์ ที่มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี หลังจากนั้นได้เข้าศึกษาต่อใน หลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2554

