

การประมาณจากสามมิติโดยการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุ



นายณัฏพล ฟุตระกูล

ศูนย์วิทยุทรัพยากร  
วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต  
สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์  
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย  
ปีการศึกษา 2551  
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

THREE DIMENSIONAL SCENE APPROXIMATION USING OBJECT RELATION ANALYSIS



Mr. Natchapon Futragoon

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

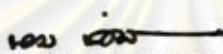
Academic Year 2008

Copyright of Chulalongkorn University


หัวข้อวิทยานิพนธ์ การประมาณจากสามมิติโดยการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุ  
โดย นายณัฏพล พุฒระกูล  
สาขาวิชา วิศวกรรมคอมพิวเตอร์  
อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. พิษณุ คนองชัยยศ

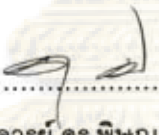
---


คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้แนบวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง  
ของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

  
..... คณบดีคณะวิศวกรรมศาสตร์  
(รองศาสตราจารย์ ดร.บุญสม เลิศธีรวงศ์)

คณะกรรมการสอบวิทยานิพนธ์

  
..... ประธานกรรมการ  
(รองศาสตราจารย์ ดร.สมชาย ประสิทธิ์จุฑาระกูล)

  
..... อาจารย์ที่ปรึกษาวิทยานิพนธ์  
(ผู้ช่วยศาสตราจารย์ ดร.พิษณุ คนองชัยยศ)

  
..... กรรมการภายนอกมหาวิทยาลัย  
(รองศาสตราจารย์ ภาวดี สมภักดี)

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ณัชพล พุตระกูล : การประมาณฉากสามมิติโดยการวิเคราะห์ความสัมพันธ์ระหว่าง  
วัตถุ. (THREE DIMENSIONAL SCENE APPROXIMATION USING OBJECT  
RELATION ANALYSIS) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ผู้ช่วยศาสตราจารย์ ดร.  
พิชญ์ คนองชัยยศ , 77 หน้า

การสร้างฉากแบบจำลองสามมิติจากรูปภาพแบบอัตโนมัติจะสามารถลดภาระและ  
ต้นทุนในการสร้างแบบจำลองสามมิติเป็นอย่างมาก แต่การประมาณตำแหน่งสามมิติจากภาพที่  
มีเพียงสองมิติ นั้นไม่ใช่งานที่ง่าย แม้จะมีนักวิจัยมากมายได้ทำการค้นคว้าแต่ก็ยังไม่มียุทธวิธีที่ดี  
ที่สุด การนำการวิเคราะห์การวางตัวของวัตถุเป็นวิธีหนึ่งที่สามารถนำมาประมาณตำแหน่งสาม  
มิติได้ โดยใช้การประมาณตำแหน่งโดยอ้างอิงจากพื้น อย่างไรก็ตามยังมีข้อจำกัดทั้งในด้าน  
ความแม่นยำ และความสวยงามของแบบจำลองสามมิติที่ได้เมื่อรูปภาพที่เป็นต้นแบบนั้นมี  
ความซับซ้อน งานวิจัยนี้จึงได้นำเสนอการระบุความสัมพันธ์ระหว่างวัตถุ และนำมาใช้ในการ  
ประมาณฉากสามมิติ เพราะความสัมพันธ์ระหว่างวัตถุจะสามารถวิเคราะห์ได้ง่ายกว่าการระบุว่า  
วัตถุใดอยู่ที่ใดในภาพ โดยการระบุความสัมพันธ์ระหว่างวัตถุนั้นทำได้โดยการวิเคราะห์ลักษณะ  
จากส่วนย่อยของรูปภาพ จากนั้นนำลักษณะส่วนย่อยนี้มาผ่านการจำแนกประเภทเป็นหก  
ประเภทโดยอาศัยวิธีการเรียนรู้โดยเครื่องจักร จากนั้นอาศัยความสัมพันธ์ทั้งหกประเภทนี้ใน  
การประมาณตำแหน่งในพิกัดสามมิติและสร้างเป็นแบบจำลองขึ้นมา ผลการทดสอบงานวิจัยนี้  
โดยการสร้างซอฟต์แวร์มาเพื่อทดสอบแนวคิดวิจัยนี้ โดยทดสอบกับรูปทั้งหมด 94 รูปซึ่งเป็น  
ทั้งรูปที่เป็นภาพถ่าย และรูปที่ถูกสังเคราะห์ขึ้นจากโปรแกรมสามมิติ ผลการทดลองแสดงให้เห็น  
เห็นว่าวิธีการที่นำเสนอสามารถนำไปประยุกต์ได้กับรูปภาพที่มีสภาพแวดล้อมต่าง ๆ กันได้  
และมีความคลาดเคลื่อนจากประมาณฉากสามมิติอยู่ที่เฉลี่ยร้อยละ 38.99

## ศูนย์วิทยทรัพยากร

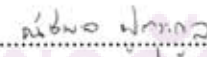
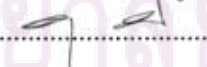
ภาควิชา ..... วิศวกรรมคอมพิวเตอร์ ..... ลายมือชื่อนิสิต ..... อนันต์ ฟูตระกูล  
สาขาวิชา ..... วิศวกรรมคอมพิวเตอร์ ..... ลายมือชื่ออ.ที่ปรึกษาวิทยานิพนธ์หลัก .....  
ปีการศึกษา ..... 2551 .....

## 5070268821 : MAJOR COMPUTER ENGINEERING

KEY WORD : OBJECT RELATION, SCENE ANALYSIS, SINGLE IMAGE, AUTOMATIC PHOTO RECONSTRUCTION, IMAGE-BASED MODELING

NATCHAPON FUTRAGOON : THREE DIMENSIONAL SCENE APPROXIMATION USING OBJECT RELATION ANALYSIS. ADVISOR : Asst. Prof. PIZZANU KANONGCHAIYOS, Dr., 77 pp

Automatic Image-based modeling method greatly reduces construction time in 3D modeling. Somehow, 3D model reconstruction from a single image is not a trivial task. Many researchers have proposed a lot of algorithms but there are no generic solution. Object placement is one of methods for approximating 3D position by using ground as a reference. However, there are some limitations in both accuracy and appearance of model when applying to complicate image scenes. Therefore, This thesis proposes a novel method by analyzing relations between objects, then uses these relations to infer 3D position. The main idea is identifying a relation is easier than identifying object in the scene. Relations are defined in 6 classes. Classification of relation is performed by extracting features from images then these features are used as input for K-nearest neighbor classifier. 3D position estimation and appearance of 3D model are determined according to relation classes. We implement software and test our system with 96 image scenes. The experimental results show that our method can reconstruct model from various types of scene with average error of 38.99 %

Department ..... Computer Engineering ..... Student's Signature..... .....  
 Field of Study..... Computer Engineering ..... Advisor's Signature ..... .....  
 Academic Year ..... 2008 .....

## กิตติกรรมประกาศ

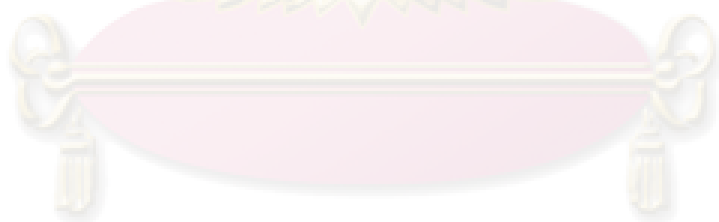
ขอขอบคุณ ผศ.ดร. พิษณุ คนองชัยยศ ที่ปรึกษาของวิทยาลัยนิพนธ์นี้ วิทยาลัยนิพนธ์เล่มนี้ คงจะเกิดขึ้นไม่ได้ถ้าขาดอาจารย์ ขอบคุณอาจารย์ที่ช่วยชี้แนะ สั่งสอน ตลอดมา

ขอขอบคุณคณะกรรมการสอบวิทยาลัยนิพนธ์ฉบับนี้ ที่ให้ข้อคิดและข้อเสนอแนะต่าง ๆ อันเป็นประโยชน์อย่างยิ่งในการพัฒนาคุณภาพของวิทยาลัยนิพนธ์ฉบับนี้ ซึ่งคณะกรรมการสอบวิทยาลัยนิพนธ์นั้น ประกอบไปด้วย รศ.ดร. สมชาย ประสิทธิ์จตุระกุล และ รศ. ภาวดี สมภักดี

ขอขอบคุณทุน 90 ปีจุฬาลงกรณ์มหาวิทยาลัย กองทุนรัชดาภิเษกสมโภช ปี พ.ศ. 2551 ที่ให้เงินทุนสนับสนุนวิทยาลัยนิพนธ์เล่มนี้

ขอขอบคุณเพื่อน ๆ พี่ ๆ ที่ทั้งในห้องวิจัยคอมพิวเตอร์กราฟฟิก และเพื่อน ๆ ทุกคน ที่แบ่งปันความรู้ ช่วยเหลือกัน และเป็นกำลังใจให้กันเสมอมา

สุดท้ายนี้ขอขอบคุณมารดาและบิดาของข้าพเจ้าที่เลี้ยงดู สนับสนุน ทำให้ข้าพเจ้ามีโอกาสดำเนินการศึกษาที่นี้จนสามารถสร้างสรรค์วิทยาลัยนิพนธ์ฉบับนี้ขึ้นมาได้



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



# สารบัญ

หน้า

บทคัดย่อภาษาไทย .....	ง
บทคัดย่อภาษาอังกฤษ .....	จ
กิตติกรรมประกาศ .....	ฉ
สารบัญ .....	ช
สารบัญภาพ .....	ฌ
สารบัญตาราง .....	ฎ
บทที่ 1 บทนำ .....	1
1.1 ที่มาและความสำคัญของปัญหา .....	1
1.2 วัตถุประสงค์ของการวิจัย .....	3
1.3 ขอบเขตของการวิจัย .....	3
1.4 ประโยชน์ที่ได้รับ .....	4
1.5 วิธีดำเนินการวิจัย .....	4
1.6 ผลงานตีพิมพ์จากงานวิจัย .....	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง .....	5
2.1 ขั้นตอนวิธีที่เกี่ยวข้อง .....	5
2.1.1 การขึ้นรูปแบบจำลองสามมิติจากภาพ .....	5
2.1.2 การขึ้นรูปแบบจำลองสามมิติจากภาพภาพเดียว .....	7
2.1.3 การแบ่งย่อยรูปภาพ .....	12
2.1.1 การจับกลุ่มค่าเฉลี่ยเค .....	13
2.1.2 การทำงานร่วมกันระหว่างการเลื่อนค่าเฉลี่ยและการตรวจหาเส้นขอบ .....	14
2.1.3 การแบ่งย่อยแบบสันปันน้ำ .....	15
2.1.4 การแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพ หรืออีจีบีเอส .....	16
2.1.5 การประมวลผลรูปภาพ .....	17
2.1.6 การจำแนกประเภท .....	19
2.1.7 การฉายภาพแบบทัศนมิติ .....	19
2.1.8 การขึ้นรูปสามมิติ .....	21
2.2 งานวิจัยที่เกี่ยวข้อง .....	23
2.2.1 การฝักรูปถ่ายขึ้นอัตโนมัติ โดย ดีเร็กโฮเอ็มและคณะ .....	23
2.2.2 การเรียนรู้โครงสร้างจากสามมิติจากรูปภาพนิ่งภาพเดียวโดยแซ็กชีนา และคณะ .....	25

บทที่ 3 วิธีการดำเนินการวิจัย .....	27
3.1 การแบ่งย่อยรูปภาพ.....	28
3.1.1 การนำขั้นตอนวิธีการแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพมาใช้ .....	30
3.1.2 ซูเปอร์พิกเซล .....	30
3.1.3 เพื่อนบ้านของซูเปอร์พิกเซล.....	31
3.1.4 การปรับปรุงคุณภาพของซูเปอร์พิกเซล .....	31
3.2 การวิเคราะห์ประเภทของความสัมพันธ์ .....	33
3.2.1 นิยามประเภทของความสัมพันธ์ .....	34
3.2.2 ตัวจำแนกประเภท.....	36
3.2.3 การสกัดค่าจุดเด่น .....	37
3.3 การสร้างแบบจำลองสามมิติ .....	42
3.3.1 การหาความลึกของแต่ละจุดภาพ.....	42
3.3.2 การขึ้นรูปแบบจำลองสามมิติ .....	46
3.3.3 การส่งออกแบบจำลองสามมิติ .....	48
3.4 วิเคราะห์ประสิทธิภาพความซับซ้อนเชิงเวลา .....	48
บทที่ 4 การทดลองและวิเคราะห์ผล.....	50
4.1 การทดสอบประสิทธิผลของค่าจุดเด่น .....	51
4.1.1 วิธีการทดสอบการกระจายตัวของค่าจุดเด่น .....	51
4.1.2 ผลการทดสอบประสิทธิผลของค่าจุดเด่น.....	53
4.2 ค่าการความแม่นยำในการจำแนกประเภทความสัมพันธ์ .....	57
4.2.1 วิธีการทดสอบความแม่นยำในการจำแนกประเภทความสัมพันธ์ .....	57
4.2.2 ผลการทดสอบความแม่นยำในการจำแนกประเภทความสัมพันธ์ .....	58
4.3 การทดสอบความแม่นยำการประมาณพิกัดสามมิติ .....	58
4.3.1 วิธีการทดสอบความแม่นยำการประมาณพิกัดสามมิติ .....	59
4.3.2 ผลการทดสอบความแม่นยำในการประมาณพิกัดสามมิติ .....	60
4.4 วิเคราะห์รูปร่างของแบบจำลองสามมิติ .....	66
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ .....	68
5.1 ข้อเสนอแนะ .....	69
รายการอ้างอิง .....	70
ภาคผนวก.....	73
ประวัติผู้เขียนวิทยานิพนธ์ .....	77



## สารบัญภาพ

หน้า

รูปที่ 1.1 แสดงตัวอย่างการขึ้นรูปจากสามมิติโดยใช้ความสัมพันธ์ระหว่างวัตถุ.....	3
รูปที่ 2.1 ตัวอย่างการขึ้นรูปจากภาพหลาย ๆ ภาพ .....	6
รูปที่ 2.2 แสดงตัวอย่างรูปภาพที่ได้จากอุปกรณ์พิเศษ .....	7
รูปที่ 2.3 แสดงการสร้างแบบจำลองโดยการวิเคราะห์จากค่าแสง .....	9
รูปที่ 2.5 แสดงผลจากสร้างแบบจำลองโครงหน้าคนโดยวิธีการหลอมภาพ .....	10
รูปที่ 2.6 แสดงผลการตรวจวัดขนาดของวัตถุต่าง ๆ .....	11
รูปที่ 2.7 ภาพต้นแบบสำหรับการทดสอบขั้นตอนวิธีการแบ่งย่อยรูปภาพ .....	13
รูปที่ 2.8 แสดงผลการแบ่งย่อยโดยวิธีการจับกลุ่มค่าเฉลี่ยเคเมื่อใช้ตั้งค่าจำนวนกลุ่มเป็นห้า ..	13
รูปที่ 2.9 แสดงผลการแบ่งย่อยด้วยระบบของเอตีสัน .....	15
รูปที่ 2.10 แสดงผลการแบ่งย่อยแบบสันปันน้ำ แต่ละส่วนย่อยจะเน้นสีต่างกัน 16	
รูปที่ 2.11 แสดงผลผลลัพธ์ที่ได้จากการแบ่งย่อยรูปภาพด้วยวิธีอีจีบีเอส .....	17
รูปที่ 2.12 แสดงการฉายภาพแบบทัศนมิติ .....	20
รูปที่ 2.13 แสดงตัวอย่างการแบ่งสามเหลี่ยม .....	23
รูปที่ 2.14 แสดงขั้นตอนการทำงานของการทำงานของการสร้างแบบจำลองสามมิติ .....	24
รูปที่ 2.15 แสดงข้อผิดพลาดของการสร้างแบบจำลองสามมิติ 25	
รูปที่ 2.16 แสดงตัวอย่างผลลัพธ์การสร้างจากสามมิติโดยวิธีการเรียนรู้โครงสร้างจากสามมิติจากรูปภาพหนึ่งภาพเดียว.....	26
รูปที่ 3.1 แผนภูมิแสดงขั้นตอนการทำงานทั้งหมด .....	27
รูปที่ 3.2 แสดงผลการแบ่งย่อยโดยวิธีการแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพ .....	31
รูปที่ 3.3 แผนภาพแสดงค่าความเรียวและความขรุขระของซูเปอร์ฟิสิกเซลและขอบเขตขั้นต่ำในการตัดสินใจเป็นขอบส่วนเกิน .....	33
รูปที่ 3.4 ตัวอย่างแสดงความสัมพันธ์ของซูเปอร์ฟิสิกเซล.....	35
รูปที่ 3.6 แสดงการหาส่วนของเส้นตรงในภาพ 39	
รูปที่ 3.8 รหัสเทียมแสดงขั้นตอนวิธีการเลือกซูเปอร์ฟิสิกเซลอ้างอิง .....	45
รูปที่ 4.1 กราฟแสดงการกระจายตัวของค่าจุดเด่น “ค่าความต่างของสีเฉลี่ย” เส้นแต่ละสีแสดงถึงคลาสแต่ละคลาส .....	52
รูปที่ 4.2 แผนภูมิแสดงคะแนนค่าประสิทธิภาพการกระจายตัวของแต่ละค่าจุดเด่น.....	54
รูปที่ 4.3 แผนภูมิแสดงการกระจายตัวของค่าจุดเด่น.....	55
รูปที่ 4.4 แผนภูมิผังงานแสดงขั้นตอนการทดสอบโดยใช้ฉากที่จำลองขึ้น .....	59



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

## สารบัญตาราง

หน้า

ตารางที่ 3.1 แสดงการแบ่งคลาสของความสัมพันธ์ .....	34
ตารางที่ 3.2 สรุปค่าจุดเด่นสำหรับการจำแนกประเภทความสัมพันธ์ระหว่าง $S_i$ กับ $S_j$ .....	40
ตารางที่ 4.1 แสดงร้อยละความถูกต้องของการทดสอบการจำแนกประเภท .....	58
ตารางที่ 4.2 แสดงความผิดพลาดจากการจำแนกประเภทแบ่งตามคลาส .....	58
ตารางที่ 4.3 แสดงผลการประมาณความลึกที่ผลลัพธ์มีความคลาดเคลื่อนน้อยที่สุด 3 อันดับแรก .....	61
ตารางที่ 4.4 แสดงผลการประมาณความลึกที่ผลลัพธ์มีความคลาดเคลื่อนมากที่สุด 3 อันดับแรก .....	62
ตารางที่ 4.5 แสดงผลลัพธ์การประมาณความลึกของภาพบางส่วนที่น่าสนใจ .....	63
ตารางที่ 4.6 ตารางแสดงการเปรียบเทียบร้อยละความคลาดเคลื่อนในการประมาณค่าความลึกจากรูปภาพสังเคราะห์ 44 รูปภาพ .....	64
ตารางที่ 4.7 ตารางแสดงการเปรียบเทียบร้อยละความคลาดเคลื่อนในการประมาณค่าความลึกบางส่วน .....	65
ตารางที่ 4.8 สรุปลักษณะของแบบจำลองสามมิติเปรียบเทียบกับวิธีของโฮเอ็มและแซ็กซีนา .....	67

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

# บทที่ 1

## บทนำ

### 1.1 ที่มาและความสำคัญของปัญหา

ในปัจจุบันรูปภาพที่ถูกเก็บอยู่ในรูปแบบดิจิทัลทั่วโลกมีจำนวนเพิ่มขึ้นอย่างมหาศาล ซึ่งเป็นผลมาจากการพัฒนาของเทคโนโลยีในการถ่ายภาพดิจิทัลที่หลากหลายมากขึ้น ไม่ว่าจะเป็นการที่มีกล้องติดอยู่กับโทรศัพท์ กล้องที่ติดอยู่กับยานพาหนะ กล้องตามสถานที่ต่าง ๆ หรือแม้แต่การพัฒนาของกล้องให้มีขนาดพกพา นอกจากนี้เทคโนโลยีในการเผยแพร่รูปภาพและการเข้าถึงข้อมูลรูปภาพเหล่านี้ทำได้ง่ายมาก ดังจะเห็นได้จากการค้นหารูปในอินเทอร์เน็ต

ด้วยปริมาณรูปภาพที่มากและสามารถเข้าถึงได้ง่ายดังที่กล่าวไปข้างต้น การนำมาใช้ประโยชน์กลับยังทำได้ไม่เต็มที่นัก เนื่องจากข้อมูลที่คอมพิวเตอร์มองเห็นภาพนั้นไม่ต่างอะไรกับตารางของค่าสีเท่านั้น การตีความหมายของภาพที่มีอยู่มากมายเหล่านี้จึงเป็นงานวิจัยที่น่าสนใจยิ่ง หนึ่งในการตีความหมายเพื่อเพิ่มข้อมูลให้แก่ภาพที่สำคัญคือ การวิเคราะห์หามิติที่สามให้แก่ภาพสองมิติ

การสร้างแบบจำลองสามมิติจากภาพสองมิติสำหรับคอมพิวเตอร์นั้นเป็นงานที่ไม่ง่าย แม้คนเราจะสามารถเข้าใจตำแหน่งของวัตถุที่อยู่ในรูปได้โดยทันที แต่ก็เนื่องด้วยจากประสบการณ์ที่เราเรียนรู้มาตั้งแต่เกิด อาทิเช่น เรา รู้จักรถและจำขนาดของรถได้ ดังนั้นถ้าขนาดเล็กแสดงว่าอยู่ไกล หรือ เรา รู้ว่าตึกต้องติดกับพื้น เช่นนี้เป็นต้น เราจึงสามารถนำประสบการณ์เหล่านี้มาวิเคราะห์และจินตนาการถึงตำแหน่งของวัตถุภายในภาพได้ แต่คอมพิวเตอร์ไม่ได้เป็นเช่นนั้น คอมพิวเตอร์ไม่สามารถนำประสบการณ์มาใช้ได้ และการนำเอาประสบการณ์เหล่านั้นมาผสมและวิเคราะห์เป็นข้อมูลนั้นเป็นเรื่องที่คอมพิวเตอร์ไม่สามารถทำได้ง่าย ๆ

ในปัจจุบันมีเครื่องมือมากมายที่จะช่วยในการสร้างฉากสามมิติออกมาแต่เครื่องเหล่านี้ยังต้องอาศัยผู้ที่มีความรู้ความชำนาญ และยังใช้เวลามากมายในการสร้างฉากเหล่านี้ขึ้นมา งานวิจัยที่จะช่วยพัฒนาเครื่องมือเหล่านี้ให้มีใช้เวลาน้อยลง และใช้ง่ายงานขึ้นจึงเกิดขึ้นมา

วิธีหนึ่งที่ปัจจุบันนิยมใช้กันมากก็คือ การขึ้นรูปแบบจำลองสามมิติจากภาพ (Image-Based Modeling) ซึ่งเป็นการสร้างฉากสามมิติโดยมีรูปถ่ายเป็นต้นแบบ และสร้างแบบจำลองที่สอดคล้องกับรูปภาพนั้น ๆ ขึ้นมา

ขั้นตอนวิธีในการขึ้นรูปแบบจำลองสามมิติจากภาพนี้สามารถแบ่งแนวทางการสร้างได้เป็นสามแบบ ตามรูปภาพที่เป็นข้อมูลนำเข้า ทั้งจาก รูปภาพรูปเดียวซึ่งเป็นภาพถ่ายธรรมดา ๆ ที่หาได้ทั่วไป [1-5] หรือ รูปภาพหลายภาพ เช่น ภาพที่ถ่ายจากภาพต่างมุม

ภาพถ่ายคู่ขนาน (Stereo images) [6] หรือแม้แต่ รูปภาพที่ถ่ายมาจากวิธีพิเศษ เช่น ภาพที่มีการจัดโครงสร้างแสง (Structured light image) [7]

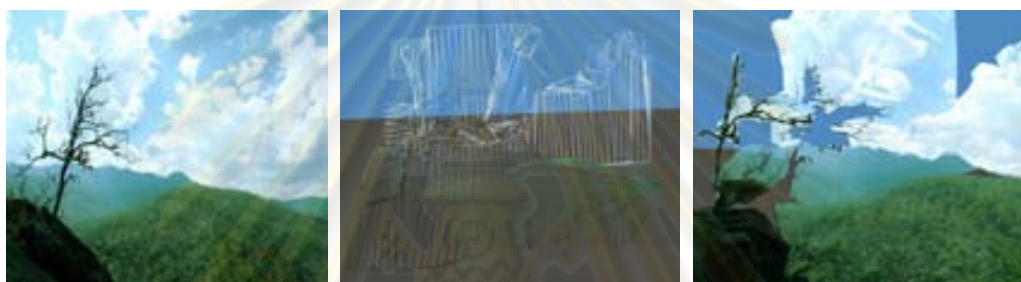
ภาพที่เก็บในอินเทอร์เน็ตส่วนใหญ่นั้นเกือบทั้งหมดเป็นรูปภาพเดี่ยว ๆ ที่ไม่มีข้อมูลตำแหน่งหรือการใช้เทคนิคพิเศษในการถ่าย ดังนั้นงานวิจัยนี้ไม่อาศัยข้อมูลอื่นใด ๆ ในการช่วยวิเคราะห์ตำแหน่งในพิกัดสามมิติ โดยสกัดข้อมูลจากรูปภาพธรรมดาเดี่ยวเท่านั้น ซึ่งในทางกลับกันในเมื่อสามารถวิเคราะห์ตำแหน่งของวัตถุโดยใช้ข้อมูลเพียงแค่อภาพเท่านั้น แต่มีข้อมูลอื่นหรือมีภาพจากมุมอื่น ๆ มาเพิ่มเติมก็ย่อมสามารถทำให้มีความแม่นยำเพิ่มขึ้นมากตามไปด้วย นอกจากนี้เพื่อที่จะให้สามารถลดเวลาในการสร้างรูปให้ได้มากที่สุด การวิเคราะห์จึงควรดำเนินไปอย่างอัตโนมัติ ซึ่งจะยิ่งทำให้ปัญหานี้กลายเป็นทั้งข้อมูลที่มีจำกัดคือมีแค่อรูปภาพเดี่ยว และยังคงให้เครื่องคอมพิวเตอร์ตัดสินใจโดยไม่มี การชี้แนะจากผู้ใช้งาน จึงทำให้ปัญหานี้ยังเป็นงานที่ท้าทายและมีการค้นคว้าอย่างต่อเนื่องมาจนถึงปัจจุบัน

ขั้นตอนวิธีในการวิเคราะห์ข้อมูลจากรูปภาพ ๆ เดียวถูกนำเสนอขึ้นมามากมาย ไม่ว่าจะเป็นการวิเคราะห์ระดับแสงสี (Shading) ลายบนพื้นผิว (Texture) เงา (shadows) การสะท้อนของแสง (Reflection) มุมและขอบของวัตถุ (Corner and Edges) และอื่น ๆ แต่ยังคงมีข้อจำกัดในสภาพแวดล้อมของรูปภาพที่ทำงานได้ แนวทางหนึ่งที่สามารถทำได้ในรูปภาพที่มีสภาพแวดล้อมหลากหลายและสามารถทำงานได้โดยอัตโนมัติคือการวิเคราะห์จากความสัมพันธ์ทางเรขาคณิตของวัตถุแต่ละชิ้นในรูปภาพ ซึ่งในปี ค.ศ. 2005 โฮเอิม (Hoiem) และคณะได้เสนอแนวทางในการแก้ปัญหาแนวทางนี้โดยอ้างอิงจากความสามารถของมนุษย์ โดยยึดหลักที่ว่าวัตถุไม่สามารถลอยอยู่เฉย ๆ ได้ และได้เสนอหลักการวิเคราะห์โดยใช้แปลงปัญหาให้อยู่ในรูปปัญหาทางสถิติ และสามารถทำให้คอมพิวเตอร์เรียนรู้และตัดสินใจตำแหน่งที่ถูกต้องของวัตถุต่างในภาพนั้น ๆ ได้ โดยจะใช้การแบ่งย่อย (Segmentation) รูปออกเป็นชิ้นส่วนย่อย ๆ โดยเรียกชิ้นส่วนย่อย ๆ นี้ว่า ซุปเปอร์พิกเซล (Super Pixel) เป็นตัวแทนของวัตถุและใช้การเรียนรู้โดยคอมพิวเตอร์ (Machine Learning) ในการตัดสินใจว่าวัตถุชิ้น ๆ คืออะไร จากนั้นจะกำหนดตำแหน่งประเภทของวัตถุ [5, 8, 9]

ในปี 2007 แซกซึนา (saxena) [3] และคณะ ก็ได้้นำแนวคิดนี้ไปพัฒนาโดยการใช้ มาร์คอฟแรนดอมฟิลด์หรือ เอ็มอาร์เอฟ (Markov Random Field, MRF) ช่วยในการเรียนรู้ และใช้ระบบตัวแปรเสริมของระนาบ (Plane Parameter) ในการสร้างแบบจำลองสามมิติของมุมมองขึ้นมาใหม่ ทำให้มีความแม่นยำเพิ่มขึ้นมาก แต่ยังมีข้อจำกัดอยู่ที่ ความสามารถในการพิจารณาจากที่ไม่ได้ถูกเรียนรู้โดยคอมพิวเตอร์ เพราะการพิจารณาส่วนใหญ่เน้นใช้สีในการระบุตำแหน่งของวัตถุ

จากข้อจำกัดดังกล่าว งานวิจัยนี้จึงได้นำเสนอข้อมูลอีกประเภทหนึ่งที่จะนำมาหาตำแหน่งสามมิติให้มีความแม่นยำเพิ่มมากขึ้น และสามารถใช้ได้กับรูปภาพที่หลากหลายไม่จำกัดอยู่กับเฉพาะรูป ขึ้น ซึ่งนั่นก็คือ การสร้างฉากสามมิติโดยการวิเคราะห์ความสัมพันธ์ของ

แต่ละวัตถุ โดยจะพิจารณาจากคุณลักษณะของกลุ่มของจุดสี และใช้ความสัมพันธ์นี้ในการสร้างฉากสามมิติขึ้นมา ซึ่งข้อดีในการสร้างฉากโดยอาศัยความสัมพันธ์นี้ก็คือ สามารถรู้ได้ว่าแบบจำลองสามมิติควรมีส่วนเชื่อมต่อกันในลักษณะใด นอกจากนี้การระบุความสัมพันธ์จะสามารถทำได้ง่ายกว่าการระบุว่าวัตถุนี้คืออะไร จึงจะทำให้ได้ความแม่นยำและสามารถนำไปประยุกต์ได้ในรูปภาพที่มีสภาพแวดล้อมต่าง ๆ กันออกไป นอกจากนี้การคำนวณตำแหน่งสามมิติเป็นการคำนวณแบบสัมพันธ์ระหว่างวัตถุจึงไม่จำเป็นต้องมีพื้นในการใช้อ้างอิงในการหาความลึกเป็นผลให้วิธีการนี้ลดข้อจำกัดลงไปได้มาก นอกจากนี้ยังสามารถใช้ความสัมพันธ์นี้ในการช่วยระบุรูปร่างลักษณะของแบบจำลองได้ ผลลัพธ์จากแบบจำลองสามมิติที่ได้จะมีลักษณะดังที่แสดงใน รูปที่ 1.1



รูปที่ 1.1 แสดงตัวอย่างการขึ้นรูปฉากสามมิติโดยใช้ความสัมพันธ์ระหว่างวัตถุ

## 1.2 วัตถุประสงค์ของการวิจัย

เพื่อนำเสนอแนวทางในการสร้างฉากสามมิติจากภาพเพียงภาพเดียวที่สามารถ

1. คำนวณตำแหน่งในภาพได้โดยอัตโนมัติ
2. ใช้งานกับรูปภาพที่มีสภาพแวดล้อมต่าง ๆ กัน
3. ใช้สร้างฉากจำลองสามมิติที่มีความถูกต้องหลายมุมมอง

## 1.3 ขอบเขตของการวิจัย

1. วัตถุต่าง ๆ ในรูปภาพไม่ได้เป็นวัตถุที่วางอยู่บนพื้นหรือติดกับพื้น วัตถุไม่ได้เป็นวัตถุที่กำลังลอยอยู่ในอากาศขณะบันทึกภาพ และพื้นในที่นี้คือระนาบที่วางตัวขนานไปกับแนวสายตาของกล้อง
2. เนื่องจากใช้ความสัมพันธ์ระหว่างวัตถุ รูปภาพที่เป็นต้นแบบจะประกอบด้วยวัตถุมากกว่าหนึ่งชิ้นขึ้นไป หรืออย่างน้อย วัตถุหนึ่งสามารถแบ่งเป็นชิ้นส่วนย่อย ๆ ได้



3. การวัดผลความถูกต้อง จำกัดเฉพาะส่วนที่สามารถเทียบได้จากรูปภาพที่เป็นข้อมูลนำเข้าเท่านั้น ในส่วนที่เพิ่มเติมออกมาทางด้านข้างและหลังจากวิธีการสร้างจะไม่นำมาคำนวณ
4. การวิจัยนี้จะใช้รูปภาพรูปเดียวที่เป็นภาพธรรมดา ๆ ไม่มีการเสริมแต่ง ในการสร้างแบบจำลอง

#### 1.4 ประโยชน์ที่ได้รับ

ได้แนวทางในการสร้างฉากจำลองสามมิติจากรูปภาพเพียงภาพเดียว ซึ่งสามารถทำงานได้โดยอัตโนมัติ และมีความแม่นยำสูง นอกจากนี้ยังสามารถนำไปประยุกต์ใช้ในด้านต่าง ๆ ได้เช่น การหาตำแหน่งสามมิติในรูปภาพ การเพิ่มมุมมองให้แก่รูปภาพ เป็นต้น

#### 1.5 วิธีดำเนินการวิจัย

1. ศึกษาเกี่ยวกับแนวทางการสร้างแบบจำลองจากภาพที่มีอยู่ในปัจจุบัน
2. วิเคราะห์และออกแบบแนวทางการใช้ความสัมพันธ์ระหว่างวัตถุในการสร้างแบบจำลอง
3. พัฒนาโปรแกรมสร้างแบบจำลองสำหรับทดสอบ
4. ทดสอบความแม่นยำและวิเคราะห์ผล
5. สรุป เรียบเรียง และจัดทำรายงานวิทยานิพนธ์

#### 1.6 ผลงานตีพิมพ์จากงานวิจัย

ส่วนหนึ่งของงานวิทยานิพนธ์นี้ ได้รับการตีพิมพ์เป็นบทความทางวิชาการสองเรื่อง ดังนี้

- “Enhanced Depth Estimation by Using Object Placement Relation” โดย ณัฏพล พุตระกูล และ พิษณุ คนองชัยยศ ในงานประชุมวิชาการ “2008 IEEE International Conference on Robotics and Biomimetics (ROBIO2008)” : ซึ่งจัดขึ้นที่ กรุงเทพฯ ประเทศไทย ในระหว่างวันที่ 22-26 กุมภาพันธ์ 2552
- “3-D Scene Approximation Using Object Placement Relation” โดย ณัฏพล พุตระกูล และ พิษณุ คนองชัยยศ ในงานประชุมวิชาการ “IV Iberoamerican Symposium in Computer Graphics (SIACG2009) จัดขึ้นที่เกาะมาการ์ตา ประเทศเวเนซุเอล่า ในเดือน มิถุนายน 2552

## บทที่ 2

### ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้มีจุดประสงค์หลักในการนำเสนอแนวทางการสร้างฉากสามมิติแบบใหม่ โดยเน้นสำคัญตรงที่ว่าขั้นตอนวิธีทั้งหมดสามารถทำงานได้โดยอัตโนมัติและสามารถสร้างฉากสามมิติโดยใช้รูปถ่ายที่เป็นต้นแบบเพียงแครูปเดียว การสร้างฉากสามมิติจากรูปภาพนั้นมีทฤษฎีที่เกี่ยวข้องและงานวิจัยที่เกี่ยวข้องดังนี้

#### 2.1 ขั้นตอนวิธีที่เกี่ยวข้อง

##### 2.1.1 การขึ้นรูปแบบจำลองสามมิติจากภาพ (Image-Based Modeling)

คือการนำรูปภาพสองมิติมาเป็นข้อมูลนำเข้า (Input) แล้วสร้างรูปแบบโครงสร้างแบบจำลองสามมิติขึ้นมาตามรูปภาพสองมิตินั้น ๆ แนวทางในการสร้างแบบจำลองสามารถแบ่งไปตามลักษณะของข้อมูลนำเข้าได้ดังนี้ [10, 11]

###### 2.1.1.1 รูปภาพภาพเดียว (Single Image)

ภาพเดียวคือภาพธรรมดาทั่ว ๆ ไป ที่อาจได้มาจากกล้องถ่ายรูปหรือจากฉากที่ผ่านการให้แสงและเงา (Rendered Image) ภาพประเภทนี้เป็นภาพที่หาได้ง่ายที่สุด ประกอบกับการเข้ามามีบทบาทของกล้องถ่ายรูปดิจิทัลในปัจจุบัน การสร้างแบบจำลองสามมิติจากรูปเหล่านี้จึงเป็นปัญหาที่น่าสนใจและท้าทายเนื่องจากการสร้างแบบจำลองสามมิติจากภาพเพียงภาพเดียวนั้นต้องใช้ข้อมูลอื่น ๆ มาทำนายตำแหน่งแบบสามมิติ ซึ่งจะกล่าวรายละเอียดในหัวข้อถัดไป

###### 2.1.1.2 รูปภาพหลาย ๆ ภาพ หลายมุมมอง (Multi-View Image)

หลักการคือการนำเอาภาพจากมุมต่าง ๆ มาประกอบกันเพื่อเติมข้อมูลในส่วนที่ไม่เห็น เช่น งานของคูทูลาคอส (Kutulakos) [12] ที่สร้างแบบจำลองสามมิติจากการถ่ายรูปวัตถุหนึ่ง ๆ ในหลาย ๆ มุมดังรูปที่ 2.1 อีกรูปแบบหนึ่งที่นิยมกันมากคือ รูปภาพที่เป็นมุมมองคู่ขนาน (Stereo Image) ซึ่งจะคล้ายกับกลไกการทำงานของ การมองเห็นของคน และถูกนำไปประยุกต์ใช้เป็นระบบการมองเห็นของหุ่นยนต์ แม้วิธีเหล่านี้จะได้ข้อมูลเพิ่มเติมมากกว่าแบบรูปภาพเดียว แต่รูปภาพแบบนี้จำเป็นต้องใช้อุปกรณ์พิเศษ หรือมีข้อจำกัดเมื่อองค์ประกอบของภาพมีการเปลี่ยนแปลงไป เช่น วัตถุอาจเคลื่อนที่ หรือ มีการเปลี่ยนของแสงและการสะท้อนระหว่างการถ่ายภาพแต่ละมุม



(ก)



(ข)

รูปที่ 2.1 ตัวอย่างการขึ้นรูปจากภาพหลาย ๆ ภาพ ก) รูปภาพที่ถ่ายจากหลาย ๆ มุม ข) ภาพแบบจำลองที่ได้ [12]

### 2.1.1.3 รูปภาพที่ใช้เทคนิคพิเศษ (Special Image)

คือการใช้อุปกรณ์พิเศษบางอย่างช่วยเมื่อทำการเก็บภาพ เช่นอาจเป็น

1. การฉายแสงโครงสร้าง (Structured light) การใช้แสงที่เป็นโครงสร้างแบบหนึ่งฉายลงไปยังฉากก่อน แล้วจึงเก็บภาพนั้น ๆ มา [13] ตัวอย่างแสดงการฉายแสงที่มีโครงสร้างแสดงในรูปที่ 2.2 ก) เช่นเดียวกับการใช้ภาพหลาย ๆ ภาพ แม้จะได้ข้อมูลเพิ่มขึ้น แต่ก็ต้องใช้อุปกรณ์พิเศษหรือจำกัดสภาพของแวดล้อมขณะถ่ายภาพ อย่างเช่นการใช้แสงฉายลงไปนั้นถ้าวัตถุสามารถสะท้อนแสงได้ การคำนวณอาจเกิดข้อผิดพลาดขึ้นได้มากมาย
2. การเก็บภาพด้วยกล้องที่เป็นเลนส์ออมนิ (Omni Camera) สามารถเก็บรูปได้ในหลายทิศทางพร้อมกันแต่ความละเอียดก็ลดลงตามไปด้วย
3. การขึ้นรูปจากภาพแบบร่าง (Sketch) [14] หรือจากภาพไวต์เฟรม (Wire frame) เป็นต้น ทั้งนี้งานประเภทนี้เป็นการนำเสนอส่วนติดต่อกับผู้ใช้ จึงเน้นการทำงานแบบโต้ตอบกับผู้ใช้โดยตรง ตัวอย่างเช่นการในรูปที่ 2.2 ข)



รูปที่ 2.2 แสดงตัวอย่างรูปภาพที่ได้จากอุปกรณ์พิเศษ ก) ตัวอย่างภาพที่มีการฉายแสงแบบมีโครงสร้าง [13] ข) ตัวอย่างการสร้างแบบจำลองสามมิติจากภาพแบบร่าง [14]

ถัดจากการแบ่งแยกตามประเภทของข้อมูลนำเข้าแล้ว แต่ละประเภทยังแบ่งได้ตามลักษณะการดำเนินงานต่อไปอีกว่า เป็นแบบอัตโนมัติ คือเครื่องคอมพิวเตอร์เมื่อรับข้อมูลรูปภาพเข้าไปก็สามารถประมวลผลและสร้างแบบจำลองสามมิติขึ้นมาได้เลย เช่นในงาน Automatic Photo Pop-up ของ โฮเอิม (Hoeim) ในปี 2005 [5] หรือแบบกึ่งอัตโนมัติ (Semi-Autonomous) คือจำเป็นต้องได้การป้อนข้อมูลบางอย่างจากผู้ใช้ระหว่างการทำงาน เช่นในงาน Tour into Picture ของโยอิชิ (Youichi) ในปี 1997 [1]

นอกจากนี้ยังแบ่งแยกตามลักษณะของผลลัพธ์ที่ออกมาอีกด้วยคือ แบบที่มีการเปลี่ยนแปลงต่อสภาพแวดล้อม (Active) เช่น อาจจะเปลี่ยน หรือ ลบการสะท้อนแสงออกไป อีกแบบหนึ่งคือ แบบที่ไม่มีการเปลี่ยนแปลงต่อสภาพแวดล้อม (Passive) วิธีนี้จะเก็บสภาพแวดล้อมทั้งหมดมาด้วย

ทั้งนี้งานวิจัยนี้ให้ความสำคัญที่การประมวลผลรูปภาพเพียงรูปภาพเดียวที่พบได้ทั่วไปจากการถ่ายภาพด้วยกล้องธรรมดา เนื่องจากต้องการนำข้อมูลรูปที่มีอยู่มากมายในอินเทอร์เน็ต มาใช้ให้เกิดประโยชน์สูงสุด ด้วยปริมาณที่มหาศาลวิธีที่ใช้จึงต้องการวิธีที่สามารถทำงานได้โดยอัตโนมัติ ดังนั้นหัวข้องานวิจัยที่สนใจคือ การขึ้นรูปแบบจำลองสามมิติจากภาพเดี่ยว

### 2.1.2 การขึ้นรูปแบบจำลองสามมิติจากภาพเดี่ยว (Image Based Modeling from Single Image)

ในงานวิจัยนี้ได้เลือกค้นคว้าในด้านที่ใช้รูปภาพเพียงภาพเดี่ยวเนื่องจากการที่มันมีอยู่ทั่วไป สามารถหาได้ง่าย และยังเป็นปัญหาที่ทำทายนมากกว่า โดยเฉพาะการสร้างแบบจำลองสามมิติแบบอัตโนมัติ ซึ่งสามารถแบ่งวิธีได้เป็นแนวทางต่าง ๆ ได้ดังนี้

### 2.1.2.1 การวิเคราะห์จากค่าแสงและเงา (Shading Based or Shape from Shading)

วิธีนี้จะดูจากความสว่างของแต่ละจุดภาพหรือพิกเซล (Pixel) วิเคราะห์ตำแหน่งในพิกัดสามมิติโดยใช้หลักการของการสะท้อนของแสง ดังนั้นสิ่งที่จะต้องรู้ในการคำนวณหาตำแหน่งจากค่าความสว่างนั้นจำเป็นต้องรู้ลักษณะการสะท้อนของพื้นผิวนั้น ๆ (Bidirectional Reflectance Distribution Function, BRDF) และตำแหน่งของมุมมองผู้สังเกตการณ์ หรือส่วนใหญ่คือตำแหน่งของกล้อง การที่จะต้องรู้ทั้งสองอย่างนี้เป็นปัญหาที่ยากมาก ในการแก้ปัญหาจะต้องมีการตั้งข้อกำหนด หรือทำงานได้เฉพาะสภาวะจำกัดอย่างใดอย่างหนึ่ง และจะต้องเป็นฉากที่ไม่ซับซ้อนมากนัก [4]

วิธีมากมายได้ถูกเสนอขึ้นเพื่อให้การวิเคราะห์ตำแหน่งมีประสิทธิภาพมากขึ้น งานวิจัยหนึ่งในปี 1998 [15] ได้วิเคราะห์วัตถุที่น่าจะสามารถสร้างเป็นแบบจำลองสามมิติได้ด้วยรูปเพียงรูปเดียว เพราะจากภาพสองมิติมุมมองหนึ่งสามารถทำนายรูปร่างของวัตถุนั้นใน ส่วนที่มองไม่เห็นได้ส่วนหนึ่ง วัตถุที่สามารถทำได้คือ วัตถุที่สมมาตรกันสองด้าน (Bilateral Symmetry Object) [6] คือวัตถุที่สามารถหาระนาบหนึ่งมาแบ่งครึ่งวัตถุนั้นเป็นสองวัตถุที่มีเหมือนกันทุกประการได้ เช่น มนุษย์ รถยนต์ วัตถุอีกชนิดคือ วัตถุทรงหลายหน้า (Polyhedral Object) ซึ่งพบได้ไม่บ่อยนัก ตัวอย่างผลงานที่ได้จากแนวความคิดนี้แสดงให้เห็นในรูปที่ 2.3

วิธีนี้มีข้อจำกัดอยู่มาก ตัวอย่างเช่น ต้องรู้ตำแหน่งของแหล่งกำเนิดแสง พื้นผิววัตถุต้องมีความสามารถในการสะท้อนแสงที่เท่า ๆ กันทั้งวัตถุ นอกจากนั้นยังมีปัญหาในการหา ส่วนที่สัมผัสกันของแต่ละวัตถุ [6] และที่สำคัญถ้าวัตถุในฉากมีความซับซ้อนประกอบกับการสะท้อนแสงหรือสภาพแสงที่ไม่ปรกติด้วยนั้น วิธีนี้แทบจะเป็นไปไม่ได้เลย

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



ก)



ข)

รูปที่ 2.3 แสดงการสร้างแบบจำลองโดยการวิเคราะห์จากค่าแสง [6] ก) รูปภาพต้นแบบ ข) ภาพผลลัพธ์ที่ได้จากวิเคราะห์ค่าแสง

#### 2.1.2.2 วิเคราะห์จากรูปร่าง (Model Based or Featured Based)

วิธีการนี้จะตรวจสอบลักษณะของวัตถุในรูปภาพแล้วนำไปเทียบกับวัตถุสามมิติในฐานข้อมูล และจะส่งคืนวัตถุที่ใกล้เคียงกับวัตถุในรูปภาพมากที่สุด ในการเปรียบเทียบรูปภาพกับวัตถุสามมิตินั้นมีอยู่หลายวิธีตั้งแต่แบบที่เรียบง่ายที่สุดคือ การใช้ข้อความในการค้นหาซึ่งจำเป็นต้องกำหนดคำอธิบายวัตถุแต่ละวัตถุในฐานข้อมูล ถัดมาคือการวิจัยในการใช้รูปร่างของวัตถุในแบบสองมิติเทียบกับภาพถ่ายของวัตถุสามมิติในฐานข้อมูล จะเห็นได้ว่าวิธีนี้ใช้เวลาในการค้นหานานมากและถ้าไม่ใช้วัตถุเดียวกันที่เหมือนกันทุกประการย่อมเกิดการผิดพลาดมากจากนั้นได้มีการนำเสนอวิธีการค้นหาโดยการการลักษณะเด่น (Feature) ในการค้นหาซึ่งทำให้การค้นหาทำได้ทั้งเร็วขึ้นและมีความแม่นยำเพิ่มมากขึ้น [16, 17] ตัวอย่างโปรแกรมนี้มีลักษณะการใช้งานดังรูปที่ 2.4

จุฬาลงกรณ์มหาวิทยาลัย





รูปที่ 2.4 ตัวอย่างโปรแกรมที่มีพื้นฐานจากการค้นหาแบบจำลองที่ใกล้เคียงจากในฐานข้อมูล [17]

นอกจากการวิเคราะห์ลักษณะของวัตถุในรูปภาพแล้วไปค้นหาในฐานข้อมูลแล้ว มีอีกวิธีหนึ่งคือการหาแบบจำลองมาตรฐานโดยการหาค่าเฉลี่ยจากแบบจำลองในฐานข้อมูล จากนั้นนำแบบจำลองมาตรฐานนี้ไปทำการหลอมภาพ (Morphing) ตามรูปภาพที่เป็นข้อมูลนำเข้า [18] โดยการสร้างแบบจำลองคร่าว ๆ แล้วให้แสงและเงา (Render) เก็บรูปภาพที่ได้มา เปรียบเทียบ แล้วหลอมใหม่ วนซ้ำไปเรื่อย ๆ วิธีนี้ได้รับความนิยมมากในการสร้างแบบจำลอง หน้าคน เพราะมีความสมจริงมาก ดังที่แสดงในรูปที่ 2.5



รูปที่ 2.5 แสดงผลจากการสร้างแบบจำลองโครงหน้าคนโดยวิธีการหลอมภาพ [18]

ข้อจำกัดของวิธีการเหล่านี้คือ วัตถุที่ได้จะต้องเป็นวัตถุที่เก็บอยู่ในฐานข้อมูล เท่านั้น หรืออย่างน้อยที่สุดต้องมีความคล้ายคลึงกันตามทอพอโลยี (Topology) จึงทำให้แบบจำลองที่สามารถหามาได้ส่วนใหญ่จะมีลักษณะเป็นวัตถุ ๆ เดียวซึ่งไม่สามารถที่จะสร้าง จากที่ความซับซ้อนประกอบไปด้วยวัตถุหลายวัตถุได้

### 2.1.2.3 วิเคราะห์จากความสัมพันธ์ทางเรขาคณิต (Geometric Content Based)

ซึ่งจะวิเคราะห์ความสัมพันธ์ของวัตถุหลาย ๆ ชิ้นในรูปภาพ เช่น การที่เส้นสอง เส้นขนานกันแล้วเล็กลง การเรียงตัวหรือบดบังกัน การที่วัตถุจะต้องวางอยู่บนพื้น ตำแหน่งของ เส้นขอบฟ้า เหล่านี้เป็นต้น แนวความคิดนี้เริ่มจากงานวิจัยในปี 1997 [1] โดยโยอิจิ (Youichi)

ในการสร้างกล่องสี่เหลี่ยมที่มีจุดหายไป (Vanishing Point) วางลงไปในรูปแบบที่เป็นรูปภาพวิวทิวทัศน์ จากนั้นก็จะสามารถทำนายตำแหน่งการวางของวัตถุต่าง ๆ ในเชิงของความลึกได้ ถัดมาในปี 1999 [19] ได้พัฒนาจนสามารถวัดระยะทางของวัตถุต่าง ๆ ในฉากได้ซึ่งแสดงให้เห็นในรูปแบบที่ 2.6 โดยจำเป็นต้องรู้ การตั้งฉากกันของเส้นขนาน 3 ชุด ตำแหน่งบนพื้น 3 จุด และความสูงที่ใช้อ้างอิง 1 แห่ง จากนั้นปี 2005 โฮเอิม (Hoeim) ได้เสนอวิธีที่สามารถสร้างแบบจำลองโดยอัตโนมัติทั้งหมดขึ้นมาได้ วิธีที่นำเสนอก็คือการสร้างแบบจำลองสามมิติโดยการอิงความสามารถของมนุษย์ในการประมวลผลความจำและประสบการณ์และการนำมาวิเคราะห์เป็นภาพสามมิติขึ้นมา เมื่อพิจารณาเกี่ยวกับเทคโนโลยีในปัจจุบันก็พบว่าวิธีการดังกล่าวสามารถประยุกต์ให้ปัญหานั้นอยู่ในรูปแบบของปัญญาประดิษฐ์ (AI , Machine Learning) ได้ โดยจะระบุชิ้นส่วนในภาพว่าคืออะไรแล้วนำมาระบุตำแหน่งโดยใช้การอ้างอิงตำแหน่งกับระนาบของพื้น เช่นเดียวกับงานของโยอิชิ (Youichi) [5] ต่อมาได้มีการพัฒนาขึ้นมาเรื่อย ๆ โดยใช้เทคนิคมาร์คอฟแรนดอมฟิลด์ (Markov Random Field) และ ตัวแปรเสริมระนาบ (Plane Parameter) โดยแซ็กซีนา (Saxena) [3] ซึ่งจะกล่าวรายละเอียดในหัวข้องานวิจัยที่เกี่ยวข้องต่อไป



รูปที่ 2.6 แสดงผลการตรวจวัดขนาดของวัตถุต่าง ๆ

แต่ทั้งนี้การตรวจหาข้อมูลเหล่านี้ยังมีความไม่แน่นอนสูง ไม่เหมาะกับการสร้างแบบจำลองของวัตถุชิ้นใดชิ้นหนึ่งโดยเฉพาะจะเหมาะกับภาพที่เป็นฉากทิวทัศน์มากกว่า ถึงกระนั้นฉากเหล่านี้ก็กลับมีความสำคัญในหลาย ๆ โปรแกรมประยุกต์ (Application) ซึ่งยากแก่การออกแบบโดยการใช้มือ ยิ่งในโปรแกรมประยุกต์ ที่ต้องใช้ฉากมาก ๆ และสมจริง จะช่วยลดภาระในการทำงานของช่างกราฟิกลงไปได้มาก เพราะวิธีนี้สามารถทำงานได้ในรูปแบบอัตโนมัติทั้งหมด

ดังจะเห็นได้ว่ามีแนวโน้มการสร้างสามมิติจากรูปที่สามารถทำให้สมจริงได้มากมาย แต่งานวิจัยนี้จะเน้นที่ความสามารถนำไปทำกับรูปจำนวนมาก ๆ ที่มีอยู่ตามอินเทอร์เน็ตได้ และสามารถช่วยลดภาระในการสร้างฉากสามมิติให้กับผู้ใช้ได้ ดังนั้นวิธีที่เหมาะสมจึงควรจะเป็นวิธีที่สามารถพัฒนาให้ทำงานได้โดยอัตโนมัติได้ งานวิจัยนี้จึงได้เลือกแนวทางการใช้วิเคราะห์จากความสัมพันธ์ทางเรขาคณิตมาเป็นแนวทางหลักในการทำวิจัย

สำหรับการวิเคราะห์จากความสัมพันธ์ทางเรขาคณิตนั้นจะประกอบไปด้วยส่วนหลัก ๆ 2 ส่วนคือ ส่วนในการประมวลผลรูปภาพ (Image Process) เพื่อแบ่งภาพออกเป็น ส่วนย่อย ๆ เรียกว่า การแบ่งย่อยรูปภาพ (Image Segmentation) และการใช้การเรียนรู้โดยเครื่องจักร (Machine Learning) เพื่อให้เครื่องคอมพิวเตอร์สามารถทำงานได้โดยอัตโนมัติ ซึ่งในการที่เรียนรู้จำเป็นต้องมีการวิเคราะห์จุดเด่นที่สำคัญในรูปภาพ ซึ่งงานวิจัยนี้ได้เสนอในการใช้ลักษณะของสี ร่วมกับลักษณะรูปร่างของวัตถุ

### 2.1.3 การแบ่งย่อยรูปภาพ (Images segmentation)

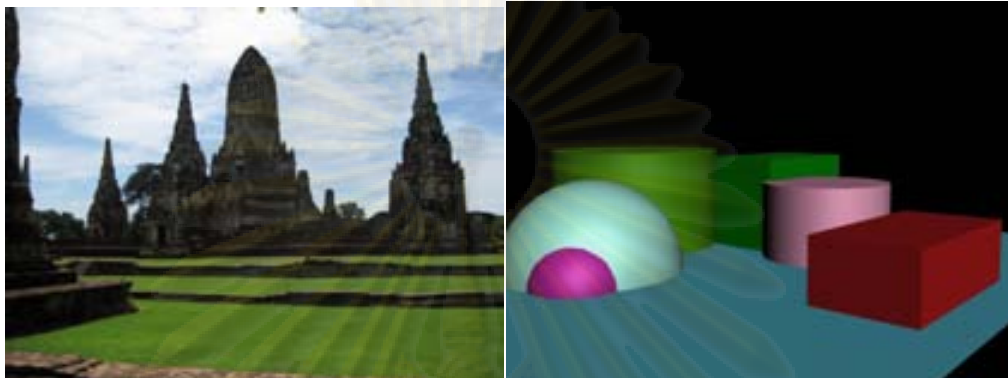
การแบ่งย่อยรูปภาพ คือ กระบวนการการแบ่งรูปภาพดิจิทัลออกเป็นเซตของจุดภาพ (Sets of pixels) โดยมีวัตถุประสงค์คือ เพื่อให้ข้อมูลดูง่ายขึ้น หรือ เปลี่ยนลักษณะการแสดงข้อมูลของภาพเป็นข้อมูลที่มีความหมายมากขึ้นหรือสามารถนำไปวิเคราะห์ได้ง่ายขึ้น [20] ขั้นตอนวิธีในการแบ่งย่อยรูปภาพเป็นงานด้านการประมวลผลรูปภาพ (Image process) ที่มีการวิจัยอย่างมาเป็นเวลานานและต่อเนื่อง เนื่องจากการแบ่งย่อยรูปภาพเป็นปัญหาที่ยาก และยังไม่ีผลเฉลยทั่วไปสำหรับปัญหาการแบ่งย่อยนี้ ขั้นตอนวิธีต่าง ๆ ถูกนำเสนอขึ้นมา มากมายและมีข้อดีและข้อเสียแตกต่างกันไป โดยในงานวิจัยนี้เน้นที่ว่าระบบควรจะสามารถทำงานได้โดยอัตโนมัติ และมีเป้าหมายในการแบ่งย่อยอยู่ที่ว่าต้องการให้แต่ละกลุ่มของจุดภาพ (Sets of pixel) หรือ ส่วนย่อย (Segment) นั้นสามารถตีความหมายได้ง่ายที่สุด

ในงานวิจัยนี้เป้าหมายคือการหาความสัมพันธ์ของแต่ละวัตถุตั้งนั้น ในแต่ละส่วนย่อยควรจะครอบคลุมวัตถุ ๆ หนึ่ง โดยสีในวัตถุนั้นอาจจะไม่เหมือนกันทั้งหมดก็ได้ และควรเก็บคุณลักษณะของขอบของวัตถุได้ด้วย หนึ่งการตีความหมายจะสามารถทำได้ง่ายขึ้นถ้าส่วนย่อยนั้นมีขนาดใหญ่แต่อย่างไรก็ตามไม่ควรจะใหญ่เกินวัตถุ ๆ หนึ่ง เพราะถ้าใหญ่เกินไปจะทำให้ในหนึ่งส่วนย่อยจะมีหลายวัตถุ ดังนั้นวิธีการแบ่งย่อยที่เหมาะสมที่สุดคือ วิธีที่สามารถแบ่งจุดภาพได้ขนาดใหญ่ที่สุดโดยที่ไม่ให้ใหญ่เกินกว่าขนาดของวัตถุ

งานวิจัยนี้จึงได้ทำการทดสอบขั้นตอนวิธีการแบ่งย่อยบางส่วนเพื่อทดสอบความเหมาะสมกับงานวิจัยนี้ ขั้นตอนวิธีการแบ่งย่อยที่ได้นำมาทดสอบนั้นมีทั้งหมดสี่วิธีด้วยกัน โดยได้แก่

1. การจัดกลุ่มค่าเฉลี่ยเค (K-mean Clustering) [20]
2. การทำงานร่วมกันระหว่างการเลื่อนค่าเฉลี่ยและการตรวจหาเส้นขอบ (Mean shift and edge detection synergy) [21]
3. การแบ่งย่อยแบบสันปันน้ำ (Watershed segmentation) [22]
4. การแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพ (Efficient graph-based image segmentation) [23]

ผลการทดสอบวิธีเหล่านี้จะจำแนกเป็นหัวข้อ ๆ โดยจะแสดงผลการแบ่งย่อยจากรูปภาพต้นแบบตัวอย่างสองประเภทดังรูปที่ 2.7 ซึ่งในรูปที่ 2.7 ก) เป็นรูปถ่ายจากกล้องทั่วไป และ รูปที่ 2.7 ข) เป็นรูปที่สร้างจากโปรแกรมออกแบบสามมิติ



รูปที่ 2.7 ภาพต้นแบบสำหรับการทดสอบขั้นตอนวิธีการแบ่งย่อยรูปภาพ ก) รูปภาพที่ถ่ายจากกล้องทั่วไป ข) รูปของแบบจำลองที่สร้างจากโปรแกรมออกแบบสามมิติ

### 2.1.1 การจับกลุ่มค่าเฉลี่ยเค (K-mean Clustering)

คือวิธีในการจับกลุ่มย่อยจำนวน  $k$  โดยมีเป้าหมายในการทำให้ความแปรปรวนภายในกลุ่มย่อย  $k$  ให้มีน้อยที่สุด โดยจะหยุดเมื่อค่าความต่างน้อยลงกว่าค่าหนึ่ง งานวิจัยนี้ได้ทดลองนำวิธีมาใช้ซึ่งต้องกำหนดตัวแปรสำคัญคือจำนวนจุดเริ่มที่ใช้ ผลที่จากวิธีนี้แสดงในรูปที่ 2.8



รูปที่ 2.8 แสดงผลการแบ่งย่อยโดยวิธีการจับกลุ่มค่าเฉลี่ยเคเมื่อใช้ตั้งค่าจำนวนกลุ่มเป็นห้า

จากผลลัพธ์จะเห็นได้ว่า วิธีการนี้สามารถแบ่งส่วนสำคัญในรูปภาพออกมาได้ดี และมีขอบที่คมชัด แต่เนื่องจากวิธีการนี้จำเป็นต้องกำหนดตำแหน่งเริ่มต้นโดยผู้ใช้ซึ่งขัดกับ



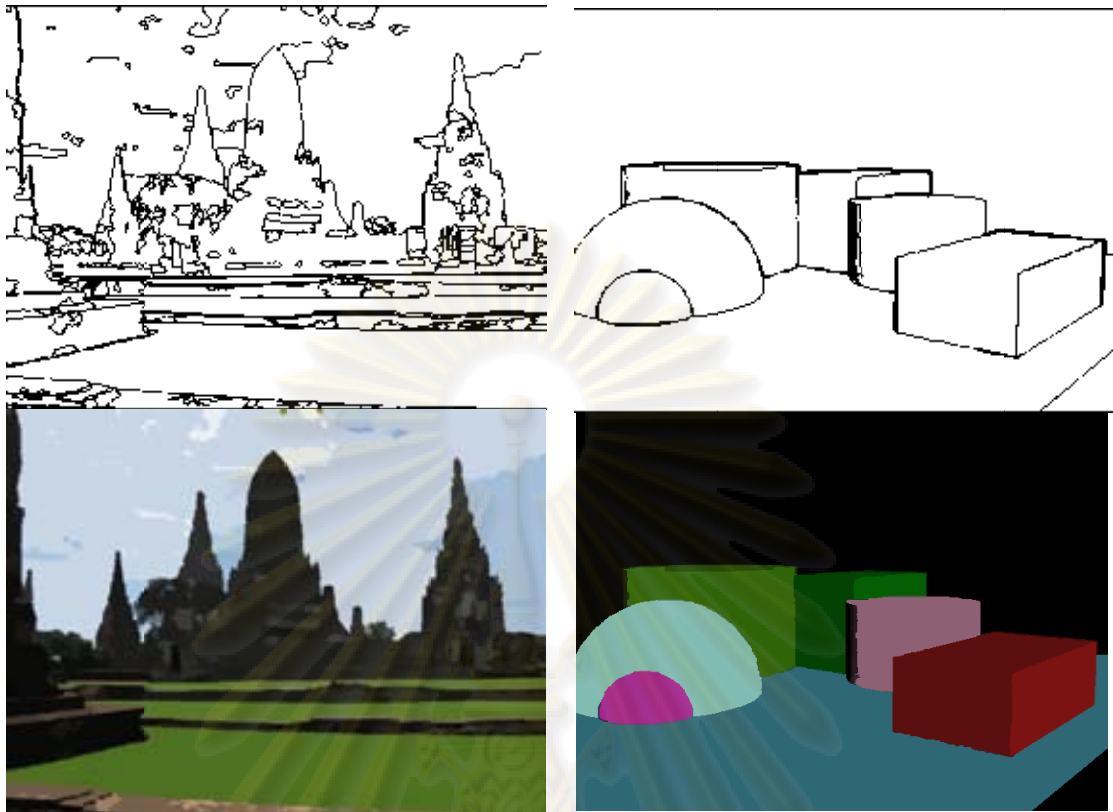
เป้าหมายของงานวิจัยนี้ซึ่งคือการออกแบบระบบให้สามารถทำงานได้ในอัตโนมัติ และการกำหนดตำแหน่งเริ่มต้นที่ผิดพลาดจะทำให้ผลการแบ่งย่อยแย่งมากดังที่แสดงในรูปที่รูปที่ 2.8 ที่กำหนดจุดเริ่มต้นห้าจุด และมีสอง จุดที่ได้ส่วนพื้นที่สี่เหลี่ยมที่เป็นหน้าเหมือนกัน ทำให้พื้นที่หน้าถูกแบ่งเป็นสองส่วน ทั้งนี้ส่วนย่อยที่ได้ยังมีลักษณะไม่ต่อเนื่องกันซึ่งจะทำให้การสร้างแบบจำลองสามมิตินั้นทำได้ยาก

### 2.1.2 การทำงานร่วมกันระหว่างการเลื่อนค่าเฉลี่ยและการตรวจหาเส้นขอบ (Mean shift and Edge detection synergy, EDISON system)

เป็นการนำขั้นตอนวิธีการเลื่อนค่าเฉลี่ยมาประยุกต์ใช้ร่วมกับการหาขอบ [21] โดยการจะเริ่มจากการหาขอบ ซึ่งวิธีที่ระบบนี้ได้นำเสนอเป็นส่วนที่พัฒนาต่อมาจากการตรวจหาเส้นขอบแค่นี้ (Canny edge detection) โดยจะเพิ่มการตัดสินใจเลือกเส้นขอบด้วยความมั่นใจ (Edge detection with embedded confidence) จากนั้นนำเส้นขอบที่ได้ขึ้นไปซ้อนทับกับรูปภาพต้นแบบ จากนั้นใช้วิธีการเลื่อนค่าเฉลี่ย โดยการปรับค่าเฉลี่ยนี้จะมีเป้าหมายในการจับกลุ่มวัตถุที่มีสีคล้ายกัน โดยจะปรับการกระจายจากที่คล้ายกันให้การเป็นสีเดียวกันไปเลย ซึ่งสีแต่ละสีก็แสดงถึงส่วนย่อยส่วนหนึ่งในรูปนั่นเอง ผลจากวิธีนี้แสดงอยู่ในรูปที่ 2.9

จากผลลัพธ์การแบ่งย่อยที่แสดงในรูปที่ 2.9 จะเห็นได้ว่าระบบของเอ็ดิสันนั้นมีข้อได้เปรียบที่เรื่องของขอบ เนื่องจากเป็นวิธีที่ผสมผสานระหว่างการใช้ขอบและตัวกรองร่วมกัน นอกจากนี้ระบบนี้ยังสามารถทำงานได้ในอัตโนมัติโดยไม่ต้องกำหนดค่าเริ่มต้น อย่างไรก็ตามแม้ขั้นตอนวิธีนี้จะสามารถแบ่งส่วนย่อยในรูปภาพที่สังเคราะห์ขึ้นได้ดีดังในรูปที่ 2.9 ข) แต่สำหรับรูปที่เป็นรูปถ่ายนั้นไม่เหมาะกับระบบที่ผู้วิจัยจะพัฒนาขึ้นนี้ เนื่องจากจำนวนของส่วนย่อย ๆ นั้นมีอยู่มาก (ประมาณ 200-300 ส่วนย่อยต่อภาพ) ดังที่แสดงในรูปที่ 2.9 ก) และส่วนย่อยที่มีขนาดเล็กก็มีอยู่เป็นจำนวนมาก สิ่งเหล่านี้จะทำให้การวิเคราะห์ความสัมพันธ์ของวัตถุทำได้ยากขึ้น

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



ก)

ข)

รูปที่ 2.9 แสดงผลการแบ่งย่อยด้วยระบบของเอ็ดิสัน ก) ผลลัพธ์เมื่อนำไปทดสอบกับภาพถ่าย  
ข) ผลลัพธ์เมื่อนำไปทดสอบกับภาพสังเคราะห์

### 2.1.3 การแบ่งย่อยแบบสันปันน้ำ (Watershed segmentation)

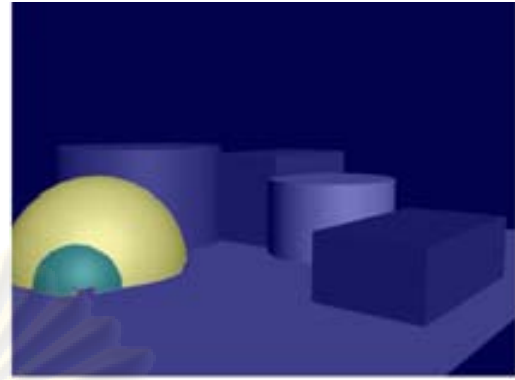
การแปลงแบบสันปันน้ำใช้ในการแบ่งย่อยรูปภาพออกเป็นส่วนย่อย ๆ ซึ่งมีแนวคิดที่สำคัญคือ การพิจารณารูปให้เหมือนภูมิประเทศที่มีระดับสูงต่ำ [22] โดยระดับสูงต่ำนั้นพิจารณาจากสีของรูปภาพแบบขาวดำ ถ้ามีระดับสีที่เท่า ๆ กันระดับความสูงก็จะมากขึ้นด้วยการแบ่งย่อยแบบสันปันน้ำ นั่นคือการทำให้มีระดับน้ำสูงขึ้นมาเรื่อย ๆ และพิจารณาขอบของภาพเป็นเหมือนเขื่อน ดังนั้นแอ่งน้ำที่เกิดขึ้นก็จะถือว่าเป็นส่วนย่อยเดียวกันนั่นเอง ทั้งนี้การจะได้เส้นขอบมาต้องผ่านการประมวลผลรูปภาพก่อน ผลจากขั้นตอนวิธีนี้ได้แสดงอยู่ในรูปภาพรูปที่ 2.10

จุฬาลงกรณ์มหาวิทยาลัย





ก)



ข)

รูปที่ 2.10 แสดงผลการแบ่งย่อยแบบสันปันน้ำ แต่ละส่วนย่อยจะเน้นสีต่างกัน ก) ผลลัพธ์ภาพส่วนย่อยเมื่อทดสอบกับภาพถ่าย ข) ผลลัพธ์ภาพส่วนย่อยเมื่อทดสอบกับภาพสังเคราะห์

วิธีการแบบสันปันน้ำนั้นให้ผลที่ดีสำหรับการแบ่งเส้นขอบ และยังสามารถรวมวัตถุที่มีลักษณะใกล้เคียงกันให้อยู่ในส่วนย่อยเดียวกันได้ทำให้จำนวนส่วนย่อยในรูปภาพมีไม่มาก นอกจากนี้ยังทำงานได้โดยอัตโนมัติ อย่างไรก็ตามวิธีการแบบสันปันน้ำ นั้นอาจเกิดการแบ่งย่อยที่เกินจริงได้ง่ายเมื่อรูปภาพมีสิ่งรบกวนอยู่มาก ซึ่งการแบ่งย่อยที่เกินจริงนี้จะทำให้ให้ส่วนย่อยส่วนหนึ่งประกอบไปด้วยหลายวัตถุดังเช่นในรูปที่ 2.10

#### 2.1.4 การแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพ หรืออีจีบีเอส (Efficient Graph-Based Image Segmentation, *EGBS*)

วิธีการนี้ใช้การแปลงจากรูปภาพให้แสดงอยู่ในรูปของกราฟแทน โดยแต่ละโนดแทนแต่ละจุดภาพในรูปภาพ [23] และเชื่อมต่อกันด้วยเส้นเชื่อมที่แสดงถึงความแตกต่างของแต่ละโนด โดยการแบ่งย่อยจะเป็นขั้นตอนที่มีพื้นฐานอยู่บนการเปรียบเทียบอาณาเขตเป็นคู่ ๆ ซึ่งผลจากการแบ่งย่อยด้วยวิธีจะให้ผลดังรูปที่ 2.11

จากผลลัพธ์ที่ได้แสดงในรูปที่ 2.11 จะเห็นได้ว่าวิธีนี้สามารถแบ่งภาพออกเป็นส่วนย่อย ๆ โดยที่แต่ละส่วนยังคงประกอบด้วยวัตถุเพียงแค่ชิ้นเดียว นอกจากนี้ยังตรงตามความต้องการที่ว่า แต่ละส่วนย่อยนั้นมีขนาดค่อนข้างใหญ่เมื่อเทียบกับขนาดของรูปและมีจำนวนไม่เยอะมากนักต่อรูป ๆ หนึ่ง ด้วยขนาดของส่วนย่อยที่มีขนาดใหญ่จะทำให้ง่ายต่อการวิเคราะห์ ซึ่งจะส่งผลต่อความแม่นยำในการวิเคราะห์หมวดหมู่ความสัมพันธ์ต่อไป

จุฬาลงกรณ์มหาวิทยาลัย



ก)



ข)

รูปที่ 2.11 แสดงผลลัพธ์ที่ได้จากการแบ่งย่อยรูปภาพด้วยวิธีอีจีบีเอส ก) แสดงผลลัพธ์โดยแต่ละส่วนย่อยแสดงด้วยสีที่แตกต่างกัน ข) แสดงส่วนย่อยที่มีตำหนิ

วิธีนี้ยังมีข้อจำกัดบางประการดังรูปที่ 2.11 ข) จะสังเกตเห็นว่าส่วนขอบของแต่ละส่วนย่อยจะมีลักษณะขรุขระ และเลเย่อออกมาจากเส้นขอบของวัตถุจริง ๆ ในภาพ ซึ่งตรงนี้จะมีผลอย่างมากในการตรวจสอบวิเคราะห์ลักษณะการวางตัวของวัตถุ

อย่างไรก็ตามวิธีที่งานวิจัยนี้จะเลือกใช้คือการแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพ เนื่องจากโดยรวมมีความเหมาะสมทั้งขนาด และส่วนประกอบภายในส่วนย่อย แม้ว่าจะมีข้อจำกัดในเรื่องของขอบของภาพที่ได้ แต่ก็สามารถแก้ไขได้ด้วยวิธีการวิเคราะห์คุณสมบัติของแต่ละส่วนย่อย

### 2.1.5 การประมวลผลรูปภาพ (Image process)

เนื่องด้วยในงานวิจัยจะต้องมีการวิเคราะห์ความหมายของส่วนต่าง ๆ ในภาพออกมาเป็นค่าตัวเลขที่สามารถชี้วัดได้ ซึ่งการวิเคราะห์นี้มีการศึกษาอย่างกว้างขวางในหัวข้อการประมวลผลรูปภาพ ในงานวิจัยนี้จะนำส่วนหนึ่งมาใช้ในการตีความหมายของภาพได้แก่ การตรวจหาขอบ (Edge detection) และการแปลงแบบฮัฟ (Hough transform)

#### 2.1.5.1 การตรวจหาขอบ (Edge detection)

คือการระบุตำแหน่งของเส้นขอบวัตถุในภาพ โดยขอบนี้อาจเกิดจากการเปลี่ยนแปลงของค่าสีระหว่างจุดสี ทั้งนี้การตรวจหาเส้นขอบที่งานวิจัยนี้ต้องการ จะแตกต่างกับการแบ่งย่อยเนื่องจากเส้นขอบที่ได้ไม่จำเป็นต้องครอบคลุมวัตถุใดวัตถุหนึ่ง ต้องการเพียงแค่ขอบที่เห็นเด่นชัดซึ่งแสดงถึงการวางตัวแค่นั้นก็เพียงพอ วิธีที่งานวิจัยนี้ได้นำมาทดสอบได้แก่วิธี การใช้ตัวกรองโซเบล (Sobel Filter) วิธีลาปลาซเซียนเกาเซียน (Laplacian of Gaussian) และ การ

ตรวจหาเส้นขอบแคเนนนี่ (Canny Edge detection) ในทั้งสามวิธีนี้วิธีการหาเส้นขอบแคเนนนี่นั้นจะให้จำนวนเส้นขอบมากที่สุดเพราะเป็นการรวมทั้งขอบเข้มและขอบอ่อน ซึ่งตรงกับความต้องการกับของงานวิจัย ซึ่งวิธีการหาขอบแคเนนนี่มีรายละเอียดดังนี้

#### การตรวจหาเส้นขอบแคเนนนี่ (Canny edge detection)

ขั้นตอนวิธีนี้มีรายละเอียดแบ่งเป็นสี่ขั้นตอนได้ดังนี้

1. การปรับภาพให้เรียบโดยตัวกรองเกาเซียน เพื่อเป็นการลดสัญญาณรบกวนและทำให้ขอบเด่นชัดขึ้น
2. คำนวณหาเส้นขอบโดยใช้ตัวกรอง 4 แบบสำหรับหาขอบแบบแนวนอน แนวตั้ง และแนวทแยงทั้งสองด้าน ตัวกรองเหล่านี้สามารถกำหนดได้ว่าเป็น โซเบล หรือพรีวิท (Sobel, Prewitt) โดยจะได้ผลเป็นเกรเดียน (Gradient) ซึ่งคืออนุพันธ์อันดับหนึ่งกับรูปภาพ ผลลัพธ์ที่ได้จะเป็นขอบเข้มคือขอบที่มีการแตกต่างของค่าสีเด่นชัด
3. การหยุดยั้งค่าที่ไม่สูงสุด (Non-maxima suppression) นำเกรเดียน ที่ได้มาใช้ในการตรวจหาขอบบาง ซึ่งเป็นเส้นขอบที่มีการเปลี่ยนแปลงไม่แนชดนักซึ่งจะมีอยู่กระจายอยู่ทั่วไปในภาพ
4. การตัดขอบเขตคู่ (Double thresholding) เพื่อเชื่อมต่อขอบเข้มกับขอบอ่อนเข้าด้วยกัน โดยหลักการคือให้ขอบเข้มเป็นหลัก และถ้าขอบเข้มยังไม่สามารถรวมเป็นพื้นผิวปิด (Closed surface) ได้จึงทำการเชื่อมต่อขอบเข้มเข้ากับขอบอ่อน

จะเห็นได้ว่าวิธีนี้สามารถตรวจหาขอบได้เป็นจำนวนมากและมีความใกล้เคียงกับขอบของวัตถุจริง ๆ ซึ่งเป็นผลดีเมื่อนำไปทำการตรวจหาเส้นตรงจากการแปลงแบบฮัฟ

#### 2.1.5.2 การแปลงแบบฮัฟ (Hough transform)

เป็นการแปลงพิกัดของตำแหน่ง  $x, y$  ให้อยู่ในรูปอื่นเพื่อให้ตีความหมายได้ง่ายขึ้น โดยเมื่อพิจารณาถึงสมการเส้นตรงดังสมการที่ (2.1)

$$y = mx + c \quad (2.1)$$

ในสมการนี้  $x$  และ  $y$  เป็นค่าที่เปลี่ยนแปลง ส่วน  $m$  และ  $c$  เป็นพารามิเตอร์ที่กำหนดค่า  $x, y$  ที่เป็นไปได้ ซึ่งการแปลงฮัฟนี้จะเปลี่ยนให้ค่า  $m$  และ  $c$  เป็นค่าที่เปลี่ยนแปลงและ ค่า  $x, y$  เป็นพารามิเตอร์ที่กำหนดค่า เขียนเป็นสมการได้ใหม่ดังนี้  $c = y - mx$  แต่ในทางปฏิบัติสมการที่เหมาะสมกว่าคือสมการที่เป็นสมการเชิงมุม เพราะจะไม่เกิดปัญหาที่ความชัน  $m$  เป็นอนันต์ ดังนั้นการแปลงฮัฟจะเป็นการแปลงพิกัด  $(x, y)$  ให้อยู่ในรูปสมการเชิงมุมที่ (2.2)

$$\rho = x \cos \theta + y \sin \theta \quad (2.2)$$

ซึ่งเมื่อแปลงให้อยู่ในระบบพิกัดนี้แล้ว ก็จะสามารถวิเคราะห์หาเส้นตรงในภาพได้ โดยการสังเกตว่าในกรณีของเส้นตรง จุด  $n$  จุดบนเส้นตรงแสดงถึงเส้นตรง  $n$  เส้นในมิติของพารามิเตอร์ เส้นตรงทั้งหมดนี้ลากผ่านจุด  $(\theta, \rho)$  ในมิติของพารามิเตอร์ มาก ๆ ก็แสดงว่าพิกัด  $(x, y)$  นั้นกำเนิดมาจากสมการเส้นตรง การแปลงแบบฮัฟฟ์นี้จึงสามารถช่วยหาเส้นตรงที่อยู่ในรูปภาพได้ซึ่งมีส่วนช่วยในการระบุประเภทความสัมพันธ์ของแต่ละวัตถุ

### 2.1.6 การจำแนกประเภท (Data classification)

การจำแนกประเภท เป็นการวิเคราะห์ข้อมูลดิบจำนวนหนึ่ง เพื่อตัดสินใจว่าข้อมูลที่ลักษณะใกล้เคียงกันนั้นอยู่ในกลุ่มเดียวกัน [24] ซึ่งวิธีการที่ได้นำมาทดสอบเพื่อใช้ในวิทยานิพนธ์นี้คือวิธีค่าเฉลี่ยของเพื่อนบ้านที่ใกล้ที่สุด (K-mean nearest neighbor) โดยขั้นตอนวิธีนี้มีรายละเอียดดังนี้

สมมติให้ค่าจุดเด่น (Feature) ของคลาส  $i$  คือ  $f_{ij}, j=1, \dots, N$  การคำนวณความคล้ายคลึงกันของค่าจุดเด่นจะทำโดยการคำนวณระยะห่าง ( $d_j$ ) และกำหนดคลาสที่มีระยะห่างน้อยที่สุดให้กับค่าจุดเด่นนั้น ๆ โดยระยะห่างอาจเป็นระยะห่างแบบใดก็ได้ซึ่งในตัวอย่างนี้จะยกระยะห่างยูคลิดีเนียน (Euclidean distance) มาใช้

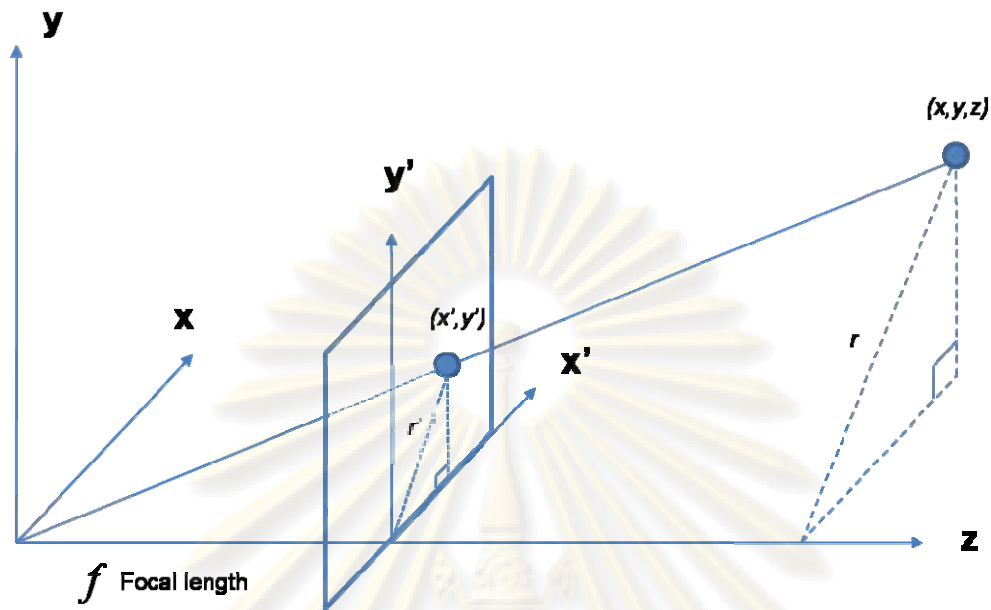
$$d_j = \left[ \sum_{i=1}^N (u_j - f_{ij})^2 \right]^{1/2} \quad (2.3)$$

โดย  $u_j$  คือ ค่าจุดเด่นที่นำมาพิจารณา ทั้งนี้  $f_{ij}$  ที่จะนำมาใช้อาจไม่ใช่ค่าทุกค่าก็ได้ แต่อาจใช้เป็นค่าตัวแทนเช่น เซนทรอยด์ (centroid) ค่าเฉลี่ย (mean) หรือ ฐานนิยม (mode) มาใช้แทน และค่าจุดเด่นจะอยู่ในคลาส  $R$  เมื่อ

$$d_R = \min_{j=1}^N [d_j] \quad (2.4)$$

### 2.1.7 การฉายภาพแบบทัศนมิติ (Perspective projection)

ภาพถ่ายเกิดจากการบันทึกค่าแสงที่สะท้อนจากวัตถุลงตัวเก็บแสงในกล้อง ซึ่งการเดินทางของแสงลงจากรับแสงในกล้องนั้นไม่ได้เดินทางในแนวขนาน แต่เป็นการเก็บค่าแสงแบบทัศนมิติ ดังที่แสดงในรูปที่ 2.12



รูปที่ 2.12 แสดงการฉายภาพแบบทัศนมิติ

ซึ่งจากรับแสงในกล้องนั้นเรียกว่าฉากภาพ (Image plane) จากรูปที่ 2.12 จะเห็นได้ว่าภาพที่ถูกฉายลงฉากภาพนั้นมีรูปร่างและขนาดที่เปลี่ยนไป ซึ่งขนาดที่เปลี่ยนไปนี้สามารถคำนวณได้ด้วยการใช้หลักการฉายภาพแบบทัศนมิติซึ่งมีหลักการดังต่อไปนี้

จุดปรากฏบนฉากภาพ (Image plane) ที่ตำแหน่ง  $(x', y')$  ซึ่งเกิดจากการบันทึกค่าแสงของจุดที่ตำแหน่ง  $(x, y, z)$  บนวัตถุ สามารถคำนวณค่าตำแหน่ง  $(x', y')$  ได้จากจุดตัดของเส้นลากผ่านจุดบนวัตถุที่ตำแหน่ง  $(x, y, z)$  และฉากภาพ

ระยะห่าง  $r$  ระหว่างจุด  $(x, y, z)$  บนวัตถุกับแกน  $z$  คำนวณได้ดังนี้  $r = \sqrt{x^2 + y^2}$  และระยะห่าง  $r'$  ระหว่างจุด  $(x', y')$  กับจุดกำเนิดบนฉากภาพคำนวณได้ดังนี้  $r' = \sqrt{x'^2 + y'^2}$  แกน  $z$  เส้นสายตาไปยังจุด  $(x, y, z)$  และส่วนของเส้นตรงความยาว  $r$  จากจุด  $(x, y, z)$  ไปยังแกน  $z$  (และตั้งฉากกันกับแกน  $z$ ) สร้างให้เกิดรูปสามเหลี่ยม แกน  $z$  เส้นสายตาไปยังจุด  $(x', y')$  บนฉากภาพและส่วนของเส้นตรงความยาว  $r'$  จากจุด  $(x', y')$  ไปยังแกน  $z$  (และตั้งฉากกันกับแกน  $z$ ) สร้างให้เกิดรูปสามเหลี่ยมอีกรูปหนึ่ง สามเหลี่ยมทั้งสองรูปนี้เป็นสามเหลี่ยมคล้าย ดังนั้นสัดส่วนของด้านของสามเหลี่ยมกับด้านคล้ายของสามเหลี่ยมอีกรูปหนึ่งมีค่าคงที่เหมือนกันดังสมการที่ (2.5) นี้

$$\frac{f}{z} = \frac{r'}{r} \quad (2.5)$$



สามเหลี่ยมที่เกิดจากเส้นตำแหน่งของแกน  $x$  เส้นตำแหน่งของแกน  $y$  และ ส่วนของเส้นตรงความยาว  $r$  กับสามเหลี่ยมบนฉากภาพซึ่งเกิดจากเส้นตำแหน่งของแกน  $x'$  เส้นตำแหน่งของแกน  $y'$  และส่วนของเส้นตรงความยาว  $r'$  เป็นสามเหลี่ยมคล้ายเช่นกัน ซึ่งสร้างให้เกิดสมการ(2.6)

$$\frac{x'}{x} = \frac{y'}{y} = \frac{r'}{r} \quad (2.6)$$

เมื่รวมสมการ (2.5) และ สมการ (2.6) เข้าด้วยกันทำให้เกิดสมการของการฉายภาพแบบทัศนมิติตั้งสมการที่ (2.7)

$$\frac{x'}{x} = \frac{f}{z} \text{ และ } \frac{y'}{y} = \frac{f}{z} \quad (2.7)$$

เพราะฉะนั้นตำแหน่งของจุดบนฉากภาพที่เกิดจากจุด  $(x, y, z)$  บนวัตถุสามารถคำนวณได้จากสมการ (2.8)

$$\begin{aligned} x' &= \frac{f}{z} x \\ y' &= \frac{f}{z} y \end{aligned} \quad (2.8)$$

### 2.1.8 การขึ้นรูปสามมิติ (3D Modeling)

หลักจากที่ได้ข้อมูลดิบมาจำนวนหนึ่ง จะนำข้อมูลเหล่านี้มาประยุกต์เพื่อนำมาสร้างเป็นแบบจำลองสามมิติ ทั้งนี้ในการสร้างสามมิตินั้นมีข้อจำกัดอยู่ปัญหาที่สำคัญและปรากฏเสมอในการสร้างแบบจำลองคือปัญหาลักษณะของรูปหลายเหลี่ยม (Polygon) ที่เป็นรูปเว้า

#### 2.1.8.1 ปัญหารูปนูน รูปเว้า (Convex Concave Problem)

รูปนูนและรูปเว้านั้นมาจากการนิยามรูปร่างทอพอโลยี (Topology) ของวัตถุ โดยรูปหลายเหลี่ยม ในระนาบหนึ่งจะเป็นรูปนูนได้ก็ต่อเมื่อมี ทุกจุดยอด (Vertex) ในรูปหลายเหลี่ยมสามารถลากเส้นต่อกันได้โดยไม่ตัดเส้นเชื่อม (Edge) ระหว่างจุดยอดใด ๆ โดยรูปหลายเหลี่ยมใดที่ไม่ได้เป็นรูปนูนก็ถือเป็นรูปเว้า [25] ซึ่งการตรวจสอบหาสามารถทำได้โดยการลากเส้นเชื่อมระหว่างจุดยอดทุกจุดและตรวจดูว่าผ่านพื้นที่ที่อยู่ด้านนอกของรูปหลายเหลี่ยมหรือไม่

เนื่องจากแบบจำลองรูปหลายเหลี่ยมที่นำมาใช้ในการแสดงผลนั้นข้อจำกัดว่าต้องเป็นรูปนูนเท่านั้น ซึ่งเป็นสาเหตุมาจากข้อจำกัดการทำให้เป็นจุดภาพ (Rasterization) ดังนั้นรูปร่างแบบจำลองจึงต้องมีการปรับให้อยู่ในรูปที่เป็นรูปนูนก่อน ที่จะส่งออกเป็นผลลัพธ์



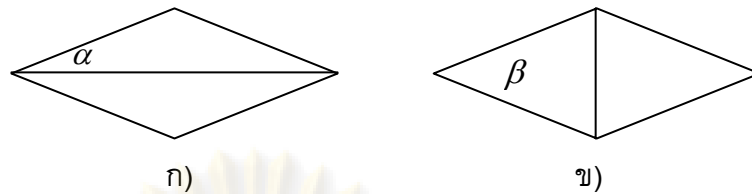
ทั้งนี้วิธีรูปหลายเหลี่ยมที่มีความเป็นรูปนูนไม่ว่าจะมีจุดยอดอยู่ในตำแหน่งใด ๆ นั่นคือรูปสามเหลี่ยม ทั้งนี้จุดยอดทั้งสามจุดจะต้องไม่อยู่ในแนวเส้นตรงเดียวกัน ซึ่งวิธีการนี้เรียกว่าวิธีการแบ่งสามเหลี่ยม (triangulation)

วิธีการแบ่งสามเหลี่ยม คือการสร้างสามเหลี่ยมจากเซตของจุด  $S \subset \mathbb{R}^2$  คือเซตของสามเหลี่ยมที่ซึ่งจุดยอด คือแสดงด้วยพิกัด  $(x,y)$  เป็นสมาชิกในเซตของ  $S$  และมีเส้นเชื่อมระหว่างคู่ของจุดใน  $S$  แต่ละจุดใน  $S$  ต้องมีส่วนร่วมอย่างน้อยในสามเหลี่ยมใดสามเหลี่ยมหนึ่ง เช่นเดียวกันถ้าเป็นการแบ่งสามเหลี่ยมในพิกัดสามมิติ  $S \subset \mathbb{R}^3$  จุดยอดจะแสดงด้วยพิกัด  $(x,y,z)$  จะเป็นการสร้างรูปพีระมิด (Tetrahedral) ขึ้นมาแทนแต่ในที่นี้ก็จะเรียกว่าเป็นการแบ่งสามเหลี่ยม เช่นเดียวกัน ในการสร้างสามเหลี่ยมนับมีกฎพื้นฐานอยู่ที่ว่าเส้นเชื่อมระหว่างคู่จุดจะต้องไม่ซ้อนทับกัน

วิธีการแบ่งสามเหลี่ยมนั้นได้มีการศึกษาค้นคว้ามาเป็นเวลานาน ผู้วิจัยได้ทดสอบขั้นตอนวิธีจำนวนหนึ่ง นั่นคือวิธีการ การแปลงเป็นรูปนูน (Convexize) การแบ่งสามเหลี่ยมแบบพัด (Fan Triangulation) และวิธีการแบ่งสามเหลี่ยมเดลาอูไน (Delaunay) ซึ่งวิธีการการแปลงเป็นรูปนูน นั้นสามเหลี่ยมรูปเว้าออกเป็นรูปนูนย่อย ๆ ได้แต่จะมีข้อผิดพลาดคือพื้นที่บางส่วนของรูปเว้าหายไป ทำให้ผลลัพธ์สุดท้ายกลายเป็นรูปทรงที่ขาดหายไป สำหรับวิธีแบบพัดนั้น จะกำหนดจุดตั้งต้นจุดหนึ่งในรูปเว้าและใช้จุดนั้นเป็นจุดกลางของพัด และสร้างกลีบของพัดที่เป็นสามเหลี่ยมไปจนทั่วรูปเว้า ซึ่งวิธีการนี้จะให้รูปหลายเหลี่ยมที่มีพื้นที่เกินออกมาจากความเป็จริง ซึ่งจากการทดสอบผู้วิจัยได้พบว่าวิธีที่เหมาะสมคือวิธีการแบ่งสามเหลี่ยมเดลาอูไน

#### 2.1.8.2 การแบ่งสามเหลี่ยมเดลาอูไน (Delaunay Triangulation)

การแบ่งสามเหลี่ยมที่ต้องการโดยทั่วไปคือการแบ่งสามเหลี่ยมที่ได้สามเหลี่ยมที่มีมุมภายในน้อยสุดให้มีค่ามากที่สุด (maximizes the minimum angle) ดังที่แสดงในรูปที่ 2.13 การแบ่งสามเหลี่ยมที่ดีควรจะได้ผลตามรูปที่ 2.13 ข. ซึ่งวิธีการสร้างสามเหลี่ยมเพื่อให้ได้มุมภายในน้อยสุดให้มีค่ามากที่สุด นี้เรียกว่า การแบ่งสามเหลี่ยมดีลาไน วิธีการนี้มีพื้นฐานอยู่บนการวิเคราะห์ไวโรนอย (Voronoi) และ การวนรอบวงกลม (Circumcircle) การวนรอบวงกลมคือวงกลมที่มีจุดยอดสามจุดอยู่บนเส้นของวงกลม การแบ่งสามเหลี่ยมเดลาอูไนจะมีคุณสมบัติว่าการวนรอบวงกลมของแต่ละสามเหลี่ยมจะไม่มีจุดยอดจุดอื่นอยู่ภายในเลย ซึ่งคุณสมบัตินี้จะถูกใช้เป็นพื้นฐานในการออกแบบขั้นตอนวิธีการก่อสร้างแบบเพิ่มขึ้นสำหรับการแบ่งสามเหลี่ยม (Incremental construction for the triangulation) [25]



รูปที่ 2.13 แสดงตัวอย่างการแบ่งสามเหลี่ยม ก) มุม  $\alpha$  คือมุมที่น้อยที่สุดในรูปหลายเหลี่ยมแรก ซึ่งน้อยกว่ามุม  $\beta$  ข) มุม  $\beta$  คือมุมที่น้อยที่สุดในรูปหลายเหลี่ยมที่สอง ซึ่งมากกว่ามุม  $\alpha$

## 2.2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้ซึ่งจะเสนอแนวทางในการใช้ความสัมพันธ์ระหว่างวัตถุเป็นข้อมูลหลักในการสร้างแบบจำลองสามมิติขึ้นมา จัดว่าอยู่ในหมวดหมู่การสร้างจากสามมิติจากภาพ โดยใช้ความสัมพันธ์ทางเรขาคณิต นอกจากนี้งานวิจัยนี้ยังเน้นสำคัญว่าต้องทำงานได้แบบอัตโนมัติ ทั้งนี้งานวิจัยที่ใช้แนวทางเดียวกัน จากการค้นคว้าของผู้วิจัย มีอยู่สองงานที่สำคัญได้แก่

### 2.2.1 การผุดรูปถ่ายขึ้นอัตโนมัติ (Automatic Photo Pop-up) โดย ดีเร็กโฮเอิม และคณะ (Derek Hoeim et al.) [5]

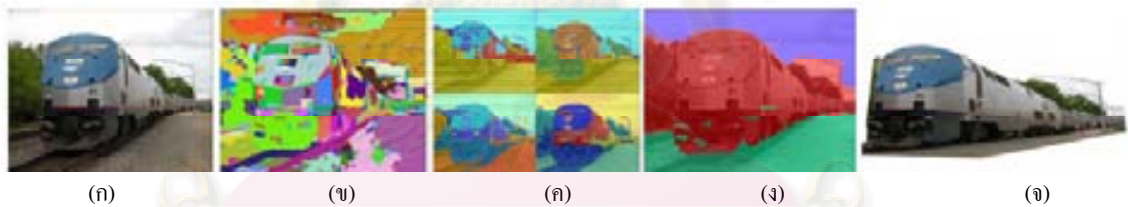
หัวใจหลักของงานนี้คือการระบุป้ายหรือเลเบล (Label) แต่ละส่วนของภาพว่าคืออะไร โฮเอิมได้แบ่งเลเบลที่ต้องการระบุเป็นเพียงแค่สามประเภทคือ พื้น วัตถุในแนวตั้ง และ ท้องฟ้า ในการระบุเลเบลจะเริ่มต้นจากนำรูปภาพที่เป็นข้อมูลนำเข้ามาใช้วิธีการแบ่งย่อยโดยเลือกใช้วิธีการแบ่งย่อยในเชิงกราฟอย่างมีประสิทธิภาพ (Efficient Graph-based Segmentation) [23] เพื่อให้ได้ส่วนย่อย ๆ ในภาพมา โดยกำหนดว่าส่วนย่อยหนึ่งประกอบด้วยวัตถุชิ้นเดียว จากนั้นจะพยายามรวบรวมกลุ่มของส่วนย่อยที่น่าจะมีเลเบลแบบเดียวกันขึ้นมา และจะทำการรวมกลุ่มอยู่หลายแบบเพื่อนำมาเป็นทางเลือกในภายหลัง จากนั้นในการระบุเลเบลของแต่ละกลุ่มจะใช้เทคนิคการตัดสินใจด้วยการเรียนรู้ของคอมพิวเตอร์ (Machine Learning)

ในการเรียนรู้จะใช้ลักษณะเด่นจำนวนหนึ่ง แบ่งเป็นกลุ่มหลัก ๆ ได้สี่กลุ่มดังนี้ กลุ่มแรก สี ท้องฟ้าควรจะมีสีฟ้า และพื้นอาจมีสีน้ำตาลหรือเขียวเป็นต้น กลุ่มที่สองคือพื้นผิว พื้นน้ำหรือพื้นดินจะลายที่แตกต่างกัน กลุ่มที่สี่คือตำแหน่ง พื้นอยู่ข้างล่างเป็นส่วนมาก และฟ้ามักจะอยู่ข้างบน และกลุ่มสุดท้ายคือลักษณะทางเรขาคณิต เช่นการหายไปของเส้นตรง การซ้อนทับกัน ใช้แบบจำลองความน่าจะเป็นในการเรียนรู้ (Probabilistic Model) และตัดสินใจระบุเลเบลว่ากลุ่มของส่วนย่อยนั้นเป็นอะไรระหว่าง พื้น วัตถุในแนวตั้ง และท้องฟ้า

การสร้างแบบจำลองสามมิติขึ้นมาใหม่นั้นจะใช้ระนาบพื้นดินที่ได้จากการตัดสินใจเป็นหลัก และยกวัตถุที่ระบุว่าเป็นวัตถุในแนวตั้งขึ้นมาในลักษณะของการผุดขึ้น (Pop-up) เหมือนการพับกระดาษขึ้นมา และท้องฟ้าจะถูกตัดออกไป การยกขึ้นมาแบบนี้จะทำให้เห็นระนาบพื้นดินนั้นนอนกับพื้นเมื่อยกวัตถุขึ้นมาตามตำแหน่งในภาพก็จะเห็นได้ว่ามีความลึกที่สอดคล้องกับรูปภาพ กลายเป็นภาพสามมิติขึ้นมา

ด้วยจุดเด่นที่กระบวนการสามารถทำได้อย่างอัตโนมัติทั้งหมดวิธีนี้จึงเป็นที่ยอมรับในวงกว้าง แต่กระนั้นวิธีนี้ยังมีข้อจำกัดมากมาย ยกตัวอย่างเช่นความแม่นยำในการประมาณตำแหน่งสามมิติ การที่วิธีนี้จำกัดว่าแต่ละฉากจะเบิ้ลได้ 3 แบบเท่านั้น ทำให้วัตถุในรูปภาพทั้งหมดก็จะถูกแบ่งเป็นสามชิ้นเท่ากับจำนวนเลเบล วัตถุที่อยู่ในฉากจึงมีข้อจำกัดในการวางตัว นอกจากนี้การใช้พื้นเป็นส่วนหลักในการตรวจหาตำแหน่งของพื้นภาพจะทำให้เกิดความผิดพลาดมากขึ้น เมื่อวัตถุในภาพไม่แสดงส่วนที่ติดกับพื้น และยังจำกัดอยู่ที่ว่าพื้นต้องอยู่ในระนาบเดียวเท่านั้นไม่สามารถนำไปประยุกต์ใช้กับภาพที่มีพื้นหลายระดับหรือมองไม่เห็นพื้นได้

สำหรับในส่วนของการสร้างแบบจำลองสามมิติที่ได้ จะเห็นได้ว่าแบบจำลองที่ได้เป็นลักษณะผุดขึ้นคือเป็นแผ่นตั้งขึ้นมาตรง ๆ ดังที่แสดงใน รูปที่ 2.15 โดยจะมีแผ่นพื้นที่เป็นแนวนอนแผ่นเดียวและส่วนที่เป็นท้องฟ้าจะถูกตัดออกไป



รูปที่ 2.14 แสดงขั้นตอนการทำงานของการสร้างแบบจำลองสามมิติ ก) ภาพต้นแบบ ข) ภาพที่ผ่านการทำการแบ่งย่อยแล้ว ค) ภาพแสดงกลุ่มของซูเปอร์พิกเซลหลาย ๆ แบบ ง) ภาพที่ถูกระบุประเภทแล้ว จ) ภาพตัวอย่างการเปลี่ยนมุมมอง

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.15 แสดงข้อผิดพลาดของการสร้างแบบจำลองสามมิติ ก) ภาพต้นแบบ ข) ภาพตัวอย่าง  
การเปลี่ยนมุมมอง

### 2.2.2 การเรียนรู้โครงสร้างฉากสามมิติจากรูปภาพนิ่งภาพเดียว (Learning 3-D Scene Structure from a Single Still Image) โดยแซ็กซีนาและคณะ (Saxena et al.) [3]

ใช้แนวคิดเดียวกันกับงานของโฮเอ้ม แต่วิธีในการตัดสินใจนั้นแตกต่างกัน ลินเซิงโดย งานวิจัยชิ้นนี้แทนที่จะทำการเลเบลแต่ละส่วนย่อยว่าเป็น พื้น วัตถุในแนวตั้ง หรือ ฟ้า แต่จะทำการระบุตำแหน่งของแต่ละระนาบที่แต่ละส่วนย่อยนั้นอยู่แทน ซึ่งการใช้ระนาบต่าง ๆ แทนแต่ละส่วนย่อยจะทำให้แบบจำลองมีความเป็นไปได้หลายรูปแบบมาก นอกจากนี้วิธียังสามารถทำงานได้โดยอัตโนมัติโดยการนำการเรียนรู้โดยเครื่องมาใช้เช่นกัน

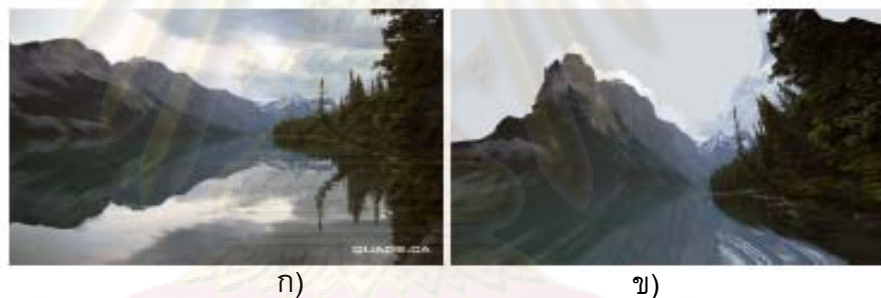
วิธีการหาตำแหน่งของระนาบนั้นแบ่งเป็นสองขั้นตอนใหญ่ได้แก่

1. การหาค่าความลึกของแต่ละจุดภาพ เริ่มจากการหาซูเปอร์พิกเซล คือด้วยวิธีการแบ่งย่อยในเชิงกราฟอย่างมีประสิทธิภาพ (Efficient Graph-based Segmentation) เช่นเดียวกันกับโฮเอ้ม และนำส่วนย่อยที่ได้นี้ไปหาค่าความลึก โดยใช้ค่าจุดเด่นที่ใช้ในการพิจารณาจะมาจากการวิเคราะห์ค่าสีโดยขั้นตอนวิธีลอว์มาร์คส์ (Law Mask) [26] โดยลอว์มาร์คส์ นี้จะถูกใช้กับรูปภาพที่ถูกแปลงความละเอียดทั้งสามระดับ ค่าสีที่ได้จะนำไปใช้ร่วมกับค่าจุดเด่นอื่น ๆ เช่น ตำแหน่ง ความแปรปรวน เป็นต้น โดยค่าความลึกที่ได้จะแตกต่างกันไปในแต่ละจุดภาพ
2. การกำหนดตำแหน่งและทิศทางของระนาบ หรือการกำหนดพารามิเตอร์ระนาบ ให้แต่ละซูเปอร์พิกเซล โดยจะทำการหาเงื่อนไขต่าง ๆ ให้แก่ระนาบ มาสร้างแบบจำลองมาคอฟแรนดอมฟิลด์ (Markov Random Field, MRF) เช่น การกำหนดจุดที่อยู่ระหว่างซูเปอร์พิกเซลขึ้นมาสองจุด ถ้าสองจุดใด ๆ บนรอยต่อนั้นมีระยะห่างเท่ากับศูนย์ตลอดก็แสดงว่าสองระนาบนี้เป็นระนาบที่ต่อกัน

(Connected Planar) เป็นต้น จากนั้นนำแบบจำลองนี้ มาเป็นตัวแปรในการเรียนรู้ และให้คอมพิวเตอร์เป็นตัวตัดสินใจ ได้ผลมาเป็นตำแหน่งและการวางตัวของระนาบสำหรับแต่ละซูเปอร์พิกเซลออกมา

จากวิธีดังกล่าวการสร้างฉากสามมิติที่มีพื้นหลาย ๆ ชั้นหรือไม่แน่นอนสามารถทำได้ เช่นในฉากของภูเขา ดังที่แสดงในรูปที่ 2.16 นอกจากนี้งานวิจัยนี้ยังได้มีการเปรียบเทียบความถูกต้องระหว่างวิธีของตนกับของโฮเอิมซึ่งระบุไว้ว่ามีความถูกต้องมากกว่าโดยเฉลี่ยร้อยละ 30 จากทดสอบกับข้อมูลทั้งหมด 134 รูป ซึ่งแสดงให้เห็นว่าวิธีนี้ความแม่นยำในการหาตำแหน่ง และสามารถสร้างแบบจำลองที่ไม่ว่าจะซับซ้อนเท่าไรก็ได้

อย่างไรก็ตามวิธีการนี้ยังคงมีข้อจำกัดอยู่ ในการหาค่าความลึกของแต่ละจุดภาพ เนื่องด้วยค่าจุดเด่นที่ใช้เป็นหลักนั้นขึ้นอยู่กับค่าสี เมื่อนำไปใช้กับรูปที่มีสภาพแวดล้อมต่างออกไปมาก ๆ จะทำให้ความแม่นยำลดลงมาก สำหรับการกำหนดความลึกที่ละจุดภาพจะทำให้ได้ค่าความลึกที่ไม่ต่อเนื่องเปลี่ยนไปมาแม้ว่าจะอยู่ในวัตถุเดียวกัน ทำให้ลักษณะเด่นของวัตถุหายไป นอกจากนี้ภาพที่ใช้จำเป็นต้องมีพื้นอยู่ในภาพ และในส่วนของท้องฟ้านั้นจะถูกตัดออกไปจากการประมาณแบบจำลองสามมิติ



รูปที่ 2.16 แสดงตัวอย่างผลลัพธ์การสร้างฉากสามมิติโดยวิธีการเรียนรู้โครงสร้างฉากสามมิติจากรูปภาพหนึ่งภาพเดียว ก) รูปภาพต้นแบบ ข) รูปภาพที่ได้จากการแสดงแบบจำลองสามมิติ

ดังจะเห็นได้ว่าการสร้างแบบจำลองโดยภาพเดี่ยวนั้นยังคงมีข้อจำกัดอยู่ โดยส่วนใหญ่แล้วจะเป็นวิธีที่อ้างอิงการใช้พื้นเป็นหลัก ทั้งในความจริงอาจจะไม่พื้นปรากฏอยู่ในรูปภาพเลย ซึ่งข้อจำกัดสามารถแก้ไขได้ด้วยวิธีการที่งานวิจัยนี้ได้นำเสนอ นั่นคือการใช้ความสัมพันธ์ระหว่างวัตถุ ในการบ่งบอกตำแหน่งในพิกัดสามมิตินั่นเอง

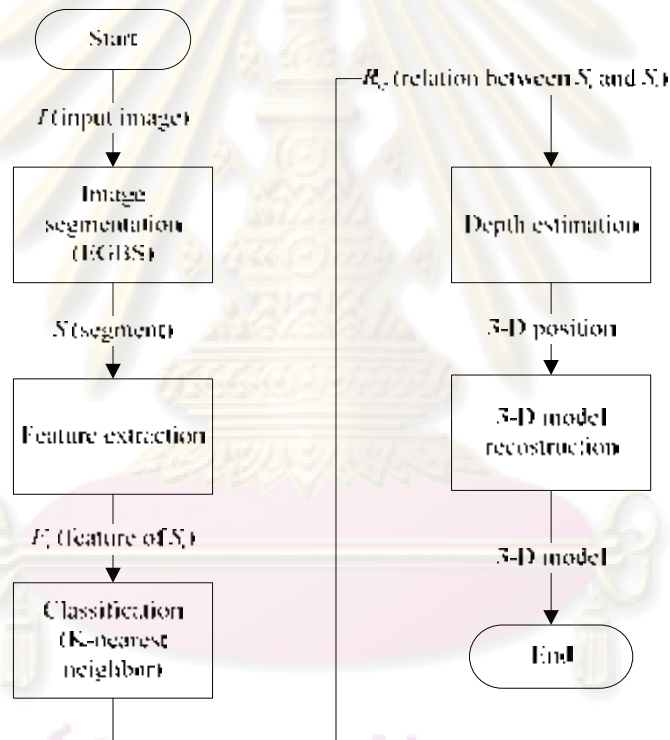
นอกจากนี้วิธีนี้ยังสามารถช่วยเพิ่มความแม่นยำให้แก่การประมาณความลึกของภาพได้เพราะการประมาณความลึกนั้นคำนวณโดยการวิเคราะห์เป็นวัตถุ แทนการวิเคราะห์เป็นที่ละจุดภาพ



### บทที่ 3

#### วิธีการดำเนินการวิจัย

งานวิจัยเป็นการนำเสนอแนวทางในการสร้างฉากสามมิติที่เน้นการทำงานที่สามารถทำได้โดยอัตโนมัติ มีความแม่นยำและสามารถใช้กับรูปที่มีสภาพแตกต่างกันไป โดยวิธีที่จะใช้นั้นคือ การแบ่งส่วนประกอบของภาพเป็นกลุ่มย่อย ๆ จากนั้นจะจำแนกความสัมพันธ์ระหว่างแต่ละกลุ่มย่อยเพื่อที่จะนำเอาความสัมพันธ์นั้นมาเป็นหลักเกณฑ์ในการสร้างแบบจำลองสามมิติ ขั้นตอนวิธีทั้งหมดสามารถแสดงได้ตามแผนภูมิรูปที่ 3.1 รายละเอียดของแต่ละขั้นตอนจะแสดงเป็นหัวข้อใหญ่สามข้อคือ การแบ่งย่อยรูปภาพ การวิเคราะห์จำแนกประเภทความสัมพันธ์ และสุดท้ายคือการสร้างแบบจำลองสามมิติ



รูปที่ 3.1 แผนภูมิแสดงขั้นตอนการทำงานทั้งหมด

ในแต่ละขั้นตอนของการทำงานจะมีการกำหนดตัวแปรต่าง ๆ ขึ้นมา เพื่อความเข้าใจผู้วิจัยจึงได้นำนิยามของแต่ละตัวแปรมาแสดงดังนี้

**นิยาม 1** รูปภพนำเข้า  $I$  คืออาร์เรย์สองมิติที่แต่ละคู่อันดับ  $(x,y)$  มีค่าสี 3 ค่าได้แก่ ค่าสีแดง  $R$  ค่าสีเขียว  $G$  และค่าสีน้ำเงิน  $B$

**นิยาม 2** ซูเปอร์พิกเซล  $S$  คือเซตของคู่อันดับ  $(x,y)$  ใน  $I$  ที่มีลักษณะคล้ายกันตามที่กำหนดในขั้นตอนการแบ่งย่อยรูปภาพ

**นิยาม 3** ค่าจุดเด่น  $F$  คือเซตของค่าจุดเด่นย่อย  $f$  สำหรับการวิเคราะห์ความสัมพันธ์  $R$

**นิยาม 4** ค่าจุดเด่นย่อย  $f$  คือค่าตัวเลขแสดงลักษณะเด่นของซูเปอร์พิกเซล

**นิยาม 5** ความสัมพันธ์  $R_{ij}$  คือคลาสแสดงลักษณะประเภทความสัมพันธ์ระหว่าง  $S_i$  กับ  $S_j$  ตามที่ได้กำหนดไว้ในหัวข้อ 3.2.1

**นิยาม 6** ฟังก์ชันความลึก  $d$  คือระยะห่างตามแนวมุมมองของกล้อง จากผู้สังเกตเข้าไปในรูปภาพ

### 3.1 การแบ่งย่อยรูปภาพ

เมื่อพิจารณารูปภาพสองมิติ  $I$  ที่เป็นรูปภาพนำเข้า เนื่องจากแนวทางที่งานวิจัยนี้จะนำเสนอคือความสัมพันธ์ของแต่ละวัตถุ ดังนั้น ขั้นแรกที่จะต้องทำคือ การระบุตำแหน่งของจุดภาพใน  $I$  ที่เป็นหน้าจะเป็นวัตถุเดียวกัน ซึ่งในขั้นแรกนั้นคือเพียงแค่วิเคราะห์ว่าวัตถุนั้นมีสีใกล้เคียงกันหรือไม่ ถ้าจุดภาพสองจุดที่มีตำแหน่งใกล้เคียงกันมีสีใกล้เคียงกันก็จะถือว่าเป็นวัตถุเดียวกัน และหลังจากนี้จะนำเอากลุ่มของจุดภาพนี้ไปหาความสัมพันธ์กับกลุ่มของจุดภาพอื่น ๆ ต่อไป

วิธีการแบ่งกลุ่มจุดภาพนี้เรียกว่าการแบ่งย่อยรูปภาพ (Image segmentation) ซึ่งเป็นส่วนหนึ่งของงาน การประมวลผลรูปภาพ (Image process) แต่ละกลุ่มย่อยที่เป็นผลจากการแบ่งนั้นจะเรียกว่า เซตของพิกเซล (Sets of pixels) หรือ ส่วนย่อย (segment) แทนด้วย ในงานวิจัยนี้จะเรียกว่า ซูเปอร์พิกเซล (Super-pixel) แทนด้วย  $S$  ตามนิยามดังนี้

ขั้นตอนวิธีในการแบ่งย่อยรูปภาพนั้นมีการค้นคว้าต่อเนื่องมาเป็นเวลานาน และมีอยู่เป็นจำนวนมาก งานวิจัยนี้จึงได้นำขั้นตอนวิธีบางส่วนมาทดสอบและเลือกวิธีที่เหมาะสมกับขั้นตอนวิธีของงานวิจัยนี้มากที่สุด

ในงานวิจัยนี้เป้าหมายคือการหาความสัมพันธ์ของแต่ละวัตถุ ดังนั้น ซูเปอร์พิกเซลแต่ละชั้นควรมีลักษณะที่เหมาะสมกับการนำไปวิเคราะห์เพื่อหาความสัมพันธ์ ซูเปอร์พิกเซลควรมีลักษณะดังนี้

1. ซูเปอร์พิกเซล  $S$  มีขนาดใหญ่ หรือ มีจำนวนน้อย ถ้าซูเปอร์พิกเซลของภาพมีขนาดใหญ่จะสามารถวิเคราะห์ตีความหมายได้ง่ายกว่าส่วนย่อยที่มีขนาดเล็ก รายละเอียดต่าง ๆ ในพื้นที่ของส่วนย่อยมีมากขึ้น การตัดสินใจระบุคุณลักษณะของซูเปอร์พิกเซลจึงมีโอกาสที่จะจำแนกได้แม่นยำมากขึ้น กลับกันถ้าส่วนย่อย

มีขนาดเล็กอาจจะเกิดจากรอยเล็ก ๆ บนวัตถุหนึ่ง ๆ จนทำให้ตีความหมายผิดไปได้ นอกจากนี้ถ้าส่วนย่อยมีขนาดใหญ่จะทำให้จำนวนรวมของซูเปอร์พิกเซลมีน้อยลง เป็นการลดภาระในการตัดสินใจโดยคอมพิวเตอร์ได้

2. ซูเปอร์พิกเซล  $S$  ประกอบด้วยวัตถุเพียงวัตถุเดียว เนื่องจากงานวิจัยที่นำเสนอเป็นการนำความสัมพันธ์ระหว่างวัตถุมาใช้ดังนั้นในการแบ่งย่อยในอุดมคติย่อมเป็นการแบ่งย่อยที่ได้ซูเปอร์พิกเซลเป็นวัตถุพอดี แต่ในทางปฏิบัตินั้นยังเป็นไปไม่ได้เนื่องได้ข้อจำกัดทางรูปภาพต่าง ๆ เพราะฉะนั้นวิธีที่ดีควรจะได้ซูเปอร์พิกเซลที่ใกล้เคียงกับวัตถุชิ้นหนึ่งที่สุด อย่างไรก็ตามการได้ซูเปอร์พิกเซลที่มีวัตถุมากกว่าหนึ่งประกอบอยู่ในส่วนย่อยนั้นถือว่าแย่กว่าการที่ได้ซูเปอร์พิกเซลที่มีเพียงส่วนหนึ่งของวัตถุ เพราะเมื่อวิเคราะห์ตำแหน่งของซูเปอร์พิกเซลจะทำให้ส่วนย่อยที่มีวัตถุมากกว่าหนึ่งจะได้ตำแหน่งที่ผิดทันทีเพราะวัตถุทั้งสองไม่ควรจะอยู่ในตำแหน่งเดียวกัน
3. ซูเปอร์พิกเซล  $S$  ประกอบด้วยขอบของวัตถุที่ชัดเจน งานวิจัยนี้จะใช้ขอบเป็นส่วนหนึ่งของการตัดสินใจความสัมพันธ์ของวัตถุการที่ได้ขอบของวัตถุที่ชัดเจนจะส่งผลให้การวิเคราะห์ความสัมพันธ์ระหว่างวัตถุมีความแม่นยำมากขึ้น

จากการทดสอบดังที่ได้แสดงในหัวข้อที่ 2.1.3 วิธีที่เหมาะสมกับงานวิจัยนี้คือการแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพหรืออีจีบีเอส (Efficient graph-based Image segmentation, ESRB) เนื่องจากผลที่ได้มีความใกล้เคียงกับคุณสมบัติที่ต้องการมากที่สุด

วิธีการแบ่งย่อยรูปภาพด้วยการใช้กราฟนั้นสามารถทำงานได้โดยอัตโนมัติ ผู้ใช้ไม่จำเป็นต้องกำหนดค่าต่าง ๆ สำหรับการแบ่งย่อยแต่ละครั้ง วิธีนี้สามารถแบ่งย่อยแล้วได้ซูเปอร์พิกเซลที่มีขนาดค่อนข้างใหญ่ คือส่วนใหญ่จะมีขนาดใหญ่จนใกล้เคียงกับวัตถุหนึ่งในภาพ และมีเพียงส่วนน้อยที่รวมสองวัตถุเอาไว้ในหนึ่งซูเปอร์พิกเซล ดังแสดงในรูปที่ 3.2 นอกจากนี้วิธีนี้สามารถรวมวัตถุได้แม้วัตถุนั้นจะมีลักษณะของการไล่เฉดสี (gradient color) หรือ วัตถุมีสัญญาณรบกวน (noise) ซึ่งสามารถช่วยในการวิเคราะห์จำแนกประเภทของความสัมพันธ์ได้มาก

อย่างไรก็ตามวิธีนี้ยังคงมีข้อจำกัดอยู่คือเมื่อใช้ขั้นตอนวิธีนี้กับภาพสีซึ่งเป็นภาพที่ประกอบไปด้วยสามช่องสี (3 channel) ซึ่งคืออาร์จีบี (RGB) เนื่องจากขั้นตอนวิธีนี้จะคำนวณการแบ่งย่อยภาพทีละ 1 ช่องสีแล้วจึงนำผลเฉลี่ยมารวมกันในภายหลัง ซึ่งจะทำให้เกิดข้อผิดพลาดขึ้นเมื่อวัตถุสองวัตถุมีสีต่างกันเพียง 1 ช่องสี ยกตัวอย่างเช่น สีแดงมีค่า อาร์จีบี คือ (255,0,0) และสีดำมีค่าอาร์จีบี เป็น (0,0,0) ผลที่ได้คือขั้นตอนวิธีนี้จะแบ่งส่วนย่อยออกจากกันที่ช่องสี R ส่วนใน G และ B นั้นจะรวมเป็นส่วนย่อยเดียวกัน เมื่อนำมาเฉลี่ยส่วนขอบจะเกินออกมาจากขอบของวัตถุจริง ๆ ซึ่งในงานวิจัยนี้เรียกขอบส่วนนี้ว่า ขอบส่วนเกิน ซึ่งงานวิจัยนี้ได้เสนอแนวทางในการลดข้อจำกัดนี้ในหัวข้อที่ 3.1.4

### 3.1.1 การนำขั้นตอนวิธีการแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพมาใช้

ในการนำขั้นตอนวิธีนี้มาใช้จำเป็นต้องกำหนดค่าตัวแปร 3 ตัว ได้แก่ ค่าสัมประสิทธิ์ของตัวกรองเกาส์เซียน ค่าขอบเขตการแบ่งแยก ค่าขนาดซูเปอร์พิกเซลเล็กสุด ตัวแปรแต่ละตัวมีรายละเอียดดังนี้

1. ค่าสัมประสิทธิ์ตัวกรองเกาส์เซียน ก่อนจะเริ่มกระบวนการแบ่งย่อยด้วย กราฟวิธีนี้ได้เสนอให้ใช้ตัวกรองเกาส์เซียนก่อน ซึ่งเมื่อผ่านตัวกรองเกาส์เซียนส่วนที่เป็นขอบของวัตถุจะเด่นชัดขึ้นมา ถ้าค่านี้ถูกกำหนดให้มีค่ามากจะทำให้มีเส้นขอบออกมาน้อย แต่ถ้าให้มีค่าน้อยเส้นขอบจะออกมามาก ผู้วิจัยได้ทำการทดลองปรับค่านี้ และพบว่าค่า 0.5 จะให้ผลลัพธ์เป็นส่วนย่อยที่อยู่ภายในมากที่สุด
2. ค่าขอบเขตการแบ่งแยก เป็นค่าที่ใช้สำหรับตัดสินใจว่าค่าพลังงานแต่ละโนดในกราฟนั้นควรจะถูกแยกออกจากกันหรือไม่ทั้งนี้ ในตามงานวิจัยที่เสนอโดยพีโตร [23] ได้แนะนำให้ตั้งค่านี้ไว้ที่ 500 ซึ่งผลการแบ่งที่ได้ก็ตรงกับความต้องการของงานวิจัยนี้
3. ค่าขนาดซูเปอร์พิกเซลเล็กสุด กำหนดค่าขนาดพิกเซลที่เล็กที่สุดที่จะเป็นส่วนย่อยได้ เนื่องจากขนาดนั้นต้องการในหน่วยเป็นพิกเซล เมื่อขนาดของรูปเปลี่ยนค่านี้จึงควรจะเปลี่ยนตามไปด้วย ผู้วิจัยได้กำหนดค่านี้เป็น 1/200 ของขนาดรูป เพื่อให้ได้ซูเปอร์พิกเซลที่มีขนาดใหญ่พอสมควร

### 3.1.2 ซูเปอร์พิกเซล (Super-pixel)

ซูเปอร์พิกเซลที่  $i$  แทนด้วย  $S_i$  โดยที่  $S_i$  เป็นเซตที่ประกอบด้วยพิกัด  $(x,y)$  ที่ได้มาจากการแบ่งย่อยโดยขั้นตอนวิธีการแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพหรืออีจีบีเอส โดยที่ กำหนดให้เซตของจุดภาพในรูปภาพข้อมูลนำเข้าคือ  $I$  ถูกแบ่งโดยวิธีการแบ่งย่อยรูปภาพอย่างมีประสิทธิภาพหรืออีจีบีเอส และ  $S_i$  คือส่วนย่อยหรือซูเปอร์พิกเซล ดังสมการที่ (3.1)

$$S_1, S_2, \dots, S_n = EGBS(I)$$

$$S_i = \{(x, y) \in I \mid (x, y) \text{ have a similar property defined by EGBS}\} \quad (3.1)$$

$$\text{โดยที่ } \forall i \forall j (S_i \cap S_j) = \emptyset, \bigcup_{i=1}^n S_i = I$$



รูปที่ 3.2 แสดงผลการแบ่งย่อยโดยวิธีการแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพ (Efficient graph-based Image segmentation) ก) รูปภาพต้นแบบ ข) ภาพที่แสดงซูเปอร์พิกเซล โดยแต่ละซูเปอร์พิกเซลแสดงด้วยสีที่ต่างกันไป

### 3.1.3 เพื่อนบ้านของซูเปอร์พิกเซล (super-pixel neighbor)

กำหนดให้ซูเปอร์พิกเซล  $S_i$  จะเป็นเพื่อนบ้านของ ซูเปอร์พิกเซล  $S_j$  ก็ต่อเมื่อ มีจุดภาพใดจุดภาพหนึ่งที่เชื่อมต่อกับ  $S_i$  ซึ่งสามารถแสดงได้ดังสมการที่ (3.2)

$$S_j \in Nb(S_i) \leftrightarrow \begin{cases} \exists(x_j \pm 1, y_j) \in S_i \\ \exists(x_j, y_j \pm 1) \in S_i \end{cases} \quad (3.2)$$

การตรวจสอบการเชื่อมต่อแบบนี้เรียกอีกอย่างหนึ่งว่าการเชื่อมต่อแบบสี่ (4-connected)

### 3.1.4 การปรับปรุงคุณภาพของซูเปอร์พิกเซล

ดังที่ได้กล่าวไปในหัวข้อที่ 2.3.3 วิธีนี้ยังมีส่วนย่อยของภาพที่ ผิดไปจากที่ได้ นิยามไว้ นั่นคือมีขอบที่เกินออกมาจากวัตถุดังที่ได้แสดงในรูปที่ รูปที่ 2.11 ข) ซึ่งขอบที่เกิน ออกมานี้จะทำให้เกิดความผิดพลาดระหว่างการทำการจำแนกประเภทเพราะไม่ตรงกับ สมมติฐานที่ตั้งไว้ว่า ส่วนย่อยส่วนหนึ่งควรจะใกล้เคียงกับวัตถุ ๆ หนึ่งมากที่สุด นอกจากนี้ยัง ทำให้คุณลักษณะขอบของส่วนย่อยไม่เป็นไปตามขอบวัตถุจริงซึ่งส่งผลต่อการจำแนกประเภท เป็นอย่างมาก ดังนั้นผู้วิจัยจึงได้เสนอแนวทางในการลดข้อจำกัดนี้

ขอบที่เกินออกมานี้เกิดจากแนวทางในการวิเคราะห์ภาพสีของขั้นตอนวิธีการ แบ่งย่อยด้วยกราฟอีจีบีเอส ซึ่งได้เสนอแนวทางที่พิจารณาแยกแต่ละช่องสี จึงทำให้เกิด ข้อผิดพลาดขึ้นเมื่อสีไม่ได้แตกต่างกันในทุก ๆ ช่องสี ค่าที่นำมาเฉลี่ยในภายหลังจึงรวมขอบ ของช่องสีที่ไม่แตกต่างกันที่ติดกัน ทำให้เกิดเป็นขอบส่วนเกินออกมา ทั้งนี้เราสามารถ ตรวจสอบหาขอบส่วนเกินเหล่านี้ได้ เพราะขอบส่วนเกินที่เกินขึ้นจากสาเหตุนี้จะมีลักษณะพิเศษ



อยู่สองประการนั้นคือ เส้นขอบมีความขรุขระมาก และมีรูปทรงที่เรียวยาวจนเกือบจะเป็นเส้นตรง ดังที่แสดงในรูปที่ 2.11 ข) ส่วนย่อยที่เป็นขอบส่วนเกินคือส่วนย่อยที่มีสีฟ้าอ่อน

งานวิจัยนี้จึงได้เสนอแนวทางในการวิเคราะห์หาขอบส่วนเกินโดยยึดหลักตามสมมติฐานที่ว่าขอบส่วนเกินจะมีลักษณะมีความขรุขระมาก และมีลักษณะเรียวยาว

การตรวจสอบความขรุขระ การตรวจสอบความขรุขระจะใช้วิธีการนับจำนวนจุดวิกฤตตามเส้นขอบของส่วนย่อยเป็นอัตราส่วนต่อความยาวเส้นขอบ ซึ่งค่าแสดงความขรุขระ  $cpp$  นี้สามารถหาได้จากสมการที่ (3.3)

$$cpp(S_i) = \frac{N(C)}{p(S_i)} \quad (3.3)$$

$$C = \Delta(\Delta(E_i)) \quad (3.4)$$

โดย  $C$  คือเซตของจุดวิกฤตที่หาได้จากอนุพันธ์อันดับสองของเส้นรอบรูป ตามสมการที่ (3.4) และ  $p$  คือฟังก์ชันสำหรับหาเส้นรอบวงของซูเปอร์ฟิกเชล  $S_i$

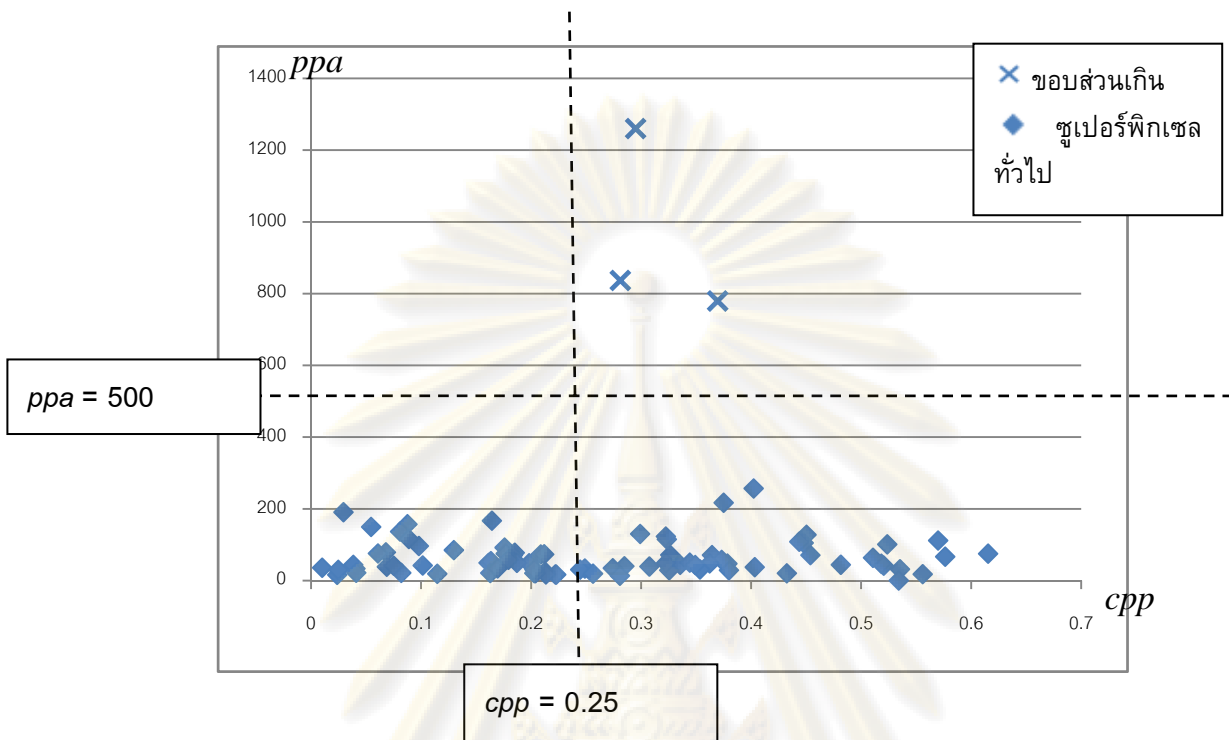
ค่าความขรุขระนี้จะช่วยแสดงถึงการเปลี่ยนแปลงทิศทางของขอบบ่อแคไหนในขอบทั้งหมด ถ้าค่าความขรุขระมากแสดงว่ามีการเปลี่ยนแปลงบ่อยเส้นก็จะขรุขระมาก และถ้าค่าความขรุขระน้อยเส้นก็มีแนวโน้มเป็นเส้นตรงมาก ทั้งนี้ค่านี้เพียงค่าเดียวจะไม่สามารถบ่งบอกว่าส่วนย่อยนี้เป็นส่วนเกินได้ เพราะถ้าวัตถุมีเส้นขอบเป็นวงกลมก็จะมีค่าความขรุขระนี้มากเช่นกัน เนื่องจากเส้นขอบของวงกลมก็มีการเปลี่ยนอยู่อย่างต่อเนื่อง

การตรวจสอบความเรียวยาว แนวทางการตรวจสอบนี้จะใช้วิธีคล้ายคลึงวิธีการตรวจความเป็นวงกลมภาพพื้นผิวปิด (Circularity of closed-surface) โดยการตรวจสอบจะใช้อัตราส่วนระหว่างความยาวเส้นรอบวงกับพื้นที่ของส่วนย่อย ค่าแสดงความเรียวยาว  $ppa$  สามารถหาได้จากสมการดังนี้ (3.5) โดย  $p$  คือฟังก์ชันหาความยาวเส้นรอบวง  $a$  คือฟังก์ชันหาพื้นที่

$$ppa(S_i) = \frac{p(S_i)^2}{a(S_i)} \quad (3.5)$$

ค่านี้สามารถแสดงความไม่เป็นวงกลมได้อย่างชัดเจน ถ้าค่านี้มีค่ามากส่วนย่อยจะมีความเรียวยาวมาก และถ้าค่านี้มีค่าน้อยก็จะมีลักษณะใกล้เคียงวงกลมมากงานวิจัยนี้ได้ทดสอบค่าทั้งสองนี้กับส่วนย่อยจำนวน 69 ส่วนย่อย และนำเสนอแสดงเป็นกราฟได้ตามรูปที่ 3.3 จากกราฟจะเห็นได้ชัดเจนว่า ส่วนขอบเกินจะมีลักษณะแตกต่างวัตถุทั่วไปอยู่พอสมควรงานวิจัยนี้จึงได้กำหนดขอบเขตต่ำสุดของทั้งค่าความขรุขระเป็น 2.5 และความเรียวยาวเป็น

500 ทั้งนี้เป็นการตั้งขอบเขตโดยคำนึงว่ายอมให้มีการผิดพลาดที่จะตรวจเจอส่วนย่อยที่ไม่ใช่ขอบมากกว่าการตรวจหาไม่เจอไปเลย



รูปที่ 3.3 แผนภาพแสดงค่าความเร็วยาวและความขรุขระของซูเปอร์ฟิสิกเซลและขอบเขตขั้นต่ำในการตัดสินใจเป็นขอบส่วนเกิน โดยแกนตั้งคือค่าแสดงความเร็วยาว  $ppa$  แกนนอนคือค่าแสดง ความขรุขระ  $cpp$  ซึ่งขอบส่วนเกิน โดยเครื่องหมายกากบาทคือซูเปอร์ฟิสิกเซลที่เป็นขอบ ส่วนเกิน

ขอบส่วนเกินที่ตรวจเจอจะนำไปรวมกับส่วนย่อยเพื่อนบ้าน โดยวิธีการเลือกคือใช้ส่วนย่อยที่มีใกล้เคียงกับขอบส่วนเกินมากที่สุด ซึ่งหมายความว่าขอบส่วนเกินนี้จะต้องถูกรวมไปกับส่วนย่อยเพื่อนบ้านขึ้นได้ชิ้นหนึ่งแน่นอน แม้จะไม่ใช่ว่าจากวัตถุเดียวกัน

### 3.2 การวิเคราะห์ประเภทของความสัมพันธ์

ในการวิเคราะห์แยกแยะความสัมพันธ์  $R_{ij}$  ระหว่างซูเปอร์ฟิสิกเซล  $S_i$  และ  $S_j$  นั้นมีเป้าหมายเพื่อที่จะนำความสัมพันธ์เหล่านี้ไปช่วยในการตัดสินใจของลักษณะแบบจำลองสามมิติ ทั้งนี้ความสัมพันธ์นี้จะต้องมีลักษณะที่เครื่องคอมพิวเตอร์จะสามารถเรียนรู้และตัดสินใจได้เองโดยอัตโนมัติ โดยรูปแบบการตัดสินใจในที่นี้จะจะเป็นแบบการเรียนรู้แบบมีการควบคุม (Supervised learning) กล่าวคือมีตัวอย่าง (Sample) ให้เรียนรู้ก่อนจากนั้นจะใช้กลุ่มตัวอย่างนี้สร้างเป็นฟังก์ชันขึ้นมา และนำฟังก์ชันนี้ไปใช้ในการตัดสินใจต่อไป

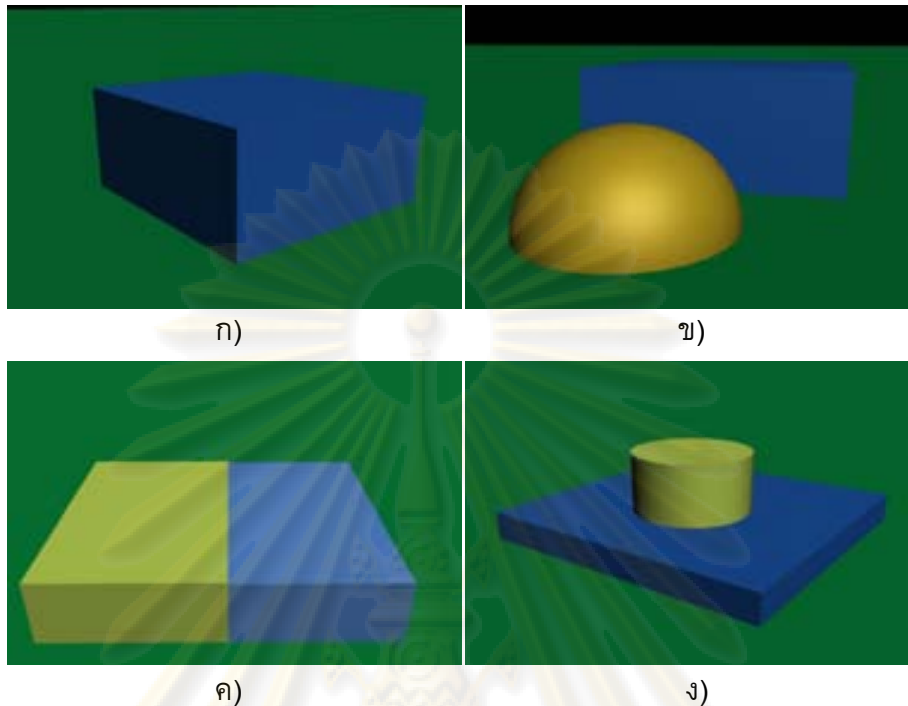
### 3.2.1 นิยามประเภทของความสัมพันธ์

ในการแยกแยะความสัมพันธ์  $R_{ij}$  ระหว่างซูเปอร์ฟิกเชล  $S_i$  และ  $S_j$  นั้น จำเป็นต้องมีการนิยามกลุ่มที่จะใช้เป็นเป้าหมายไว้ งานวิจัยนี้ได้ค้นคว้าและสังเกตจึงได้นิยามความสัมพันธ์ของแต่ละซูเปอร์ฟิกเชล โดยใช้ความลึกของซูเปอร์ฟิกเชล  $d(S)$  เป็นเกณฑ์หลัก ซึ่งสามารถแบ่งความสัมพันธ์ออกมาเป็นทั้งหมด 6 คลาส ซึ่งมีรายละเอียดดัง ตารางที่ 3.1 และ ตัวอย่างแสดงความสัมพันธ์แต่ละคลาสในรูปที่ 3.4

ตารางที่ 3.1 แสดงการแบ่งคลาสของความสัมพันธ์

สัญลักษณ์ คลาส	ชื่อคลาส	ค่าความลึก	รายละเอียด
1 (SO)	วัตถุ เดียวกัน	$d(S_i) = d(S_j)$	$S_i$ และ $S_j$ อยู่บนวัตถุชั้น เดียวกัน
2 (OT)	ปิดบัง	$d(S_i) < d(S_j)$	$S_i$ บังทับ $S_j$ อยู่ หรือ $S_i$ อยู่ ข้างหน้า $S_j$
3 (OB)	ถูกปิดบัง	$d(S_i) > d(S_j)$	$S_i$ ถูกบังทับโดย $S_j$ อยู่ หรือ $S_i$ อยู่ข้างหลัง $S_j$
4 (SP)	ระนาบ เดียวกัน	$d(S_i) = d(S_j)$	$S_i$ และ $S_j$ อาจอยู่บนวัตถุ เดียวกันหรือคนละวัตถุก็ได้ แต่จะอยู่ในแนวระนาบ เดียวกัน
5 (PT)	วางอยู่	$\min(d(S_j)) < d(S_i) < \max(d(S_j))$	$S_i$ วางอยู่บน $S_j$ หรือ $S_i$ มี ตำแหน่งอยู่ระหว่าง $S_j$
6 (PB)	ถูกวาง ทับ	$\min(d(S_i)) < d(S_j) < \max(d(S_i))$	$S_i$ เป็นฐานวางของ $S_j$ หรือ $S_i$ ครอบคลุมตำแหน่งของ $S_j$

ตามตารางที่ 3.1 จะเห็นได้ว่าครอบคลุมความลึกทั้งหมด ซึ่งหมายความว่าในการวิเคราะห์แต่ละครั้งจะต้องสามารถแยกแยะประเภทลงประเภทใดประเภทหนึ่งใน 6 คลาสนี้ได้ รายละเอียดและตัวอย่างสำหรับความสัมพันธ์ทั้ง 6 ประเภทมีดังนี้



รูปที่ 3.4 ตัวอย่างแสดงความสัมพันธ์ของซูเปอร์พิกเซล ก) แสดงความสัมพันธ์แบบวัตถุขึ้น  
 เดียวกัน ข) แสดงความสัมพันธ์แบบปิดบังและถูกปิดบัง ค) แสดงความสัมพันธ์แบบระนาบ  
 เดียวกัน ง) แสดงความสัมพันธ์แบบวางอยู่และถูกวางทับ

### 1. วัตถุเดียวกัน (Same Object, SO)

ซูเปอร์พิกเซล  $S_i$  และ  $S_j$  ทั้งสองอยู่บนวัตถุขึ้นเดียวกัน ทั้งนี้จะเป็นซูเปอร์พิกเซลที่อยู่ในระนาบเดียวกันหรือไม่ก็ได้ แต่คุณลักษณะโดยรวมแล้วคล้าย ๆ กัน ค่าความลึกของซูเปอร์พิกเซลทั้งสองจะเท่ากันหรือต่อเนื่องกัน ตัวอย่างความสัมพันธ์นี้แสดงในรูปที่ 3.4 ก)

ลักษณะที่ชัดเจน ได้แก่ ลักษณะของสี การกระจายตัวของสีระหว่างสองซูเปอร์พิกเซลจะมีค่าใกล้เคียงกัน หรือมักจะต่างกันเฉพาะค่าความเข้มแสง (Intensity) ซึ่งถ้าพิจารณาในปริภูมิเอชเอสวี (HSV space) ก็จะสามารถพิจารณาได้ง่าย โดยดูแค่ค่าสีสัน (Hue) กับ ความอิ่มตัว (Saturation) ถ้าใกล้เคียงกันก็จะมีโอกาสอยู่ในคลาสนี้ได้มาก

### 2. ปิดบังและถูกปิดบัง (Occlude, OT and be Occluded by ,OB)

ซูเปอร์พิกเซล  $S_i$  และ  $S_j$  ทั้งสองบดบังกันกล่าวคือมีซูเปอร์พิกเซลหนึ่งอยู่ข้างหน้าและซูเปอร์พิกเซลหนึ่งอยู่ข้างหลัง ในการนิยามคลาสได้แบ่งคลาสการปิดบังออกเป็นสองคลาสคือ คลาส  $S_i$  ปิดบัง  $S_j$  และคลาส  $S_j$  ถูกปิดบังโดย  $S_i$  เพราะคุณลักษณะเด่นนั้นแตกต่างกันออกไป ค่าความลึกของระหว่างซูเปอร์พิกเซลทั้งสองนี้จะมีค่ามากกว่าหรือน้อยกว่าแบบไม่ต่อเนื่องกัน ซึ่งค่าจะแตกต่างกันน้อยนั้นขึ้นอยู่กับตำแหน่งของซูเปอร์พิกเซลทั้งสอง ตัวอย่างความสัมพันธ์นี้แสดงในรูปที่ 3.4 ข) คลาสนี้สามารถช่วยระบุท้องฟ้าได้คือซูเปอร์

พิกเซลที่ถูกระบุว่าอยู่ข้างหลังสุดส่วนใหญ่มีสีฟ้าหรือขาว และอยู่ตำแหน่งบนสุดของรูปภาพ มีจำนวนเส้นตรงน้อยมาก

ลักษณะที่ชัดเจน ลักษณะของเส้นขอบของซูเปอร์พิกเซลทั้งสองชั้น กับลักษณะของเส้นขอบของส่วนที่ต่อเนื่องกัน ดังที่แสดงในรูปที่ 3.4 ข) นอกจากนี้วัตถุลักษณะนี้มักจะมีคุณลักษณะของสีที่แตกต่างกัน ลักษณะเส้นตรงที่เชื่อมต่อกันก็มักจะไม่ชี้ไปในทิศทางเดียวกัน

### 3. ระนาบเดียวกัน (Same Plane, SP)

ซูเปอร์พิกเซล  $S_i$  และ  $S_j$  อยู่ในระนาบเดียวกัน แต่อยู่บนวัตถุต่างชนิดกันหรือมีความแตกต่างของลวดลายมาก ค่าความลึกของซูเปอร์พิกเซลทั้งสองนี้จะมีค่าเดียวกัน หรือต่อเนื่องกัน ความสัมพันธ์แบบนี้พบบ่อยบนลายของถนน ทางเท้า ป้ายประกาศที่มีลายต่าง ๆ เป็นต้น ตัวอย่างความสัมพันธ์นี้แสดงในรูปที่ 3.4 ค)

ลักษณะที่ชัดเจน ได้แก่ เส้นตรงที่อยู่ในส่วนย่อย โดยถ้าในส่วนย่อยประกอบด้วยเส้นตรงที่มีทิศทางเดียวกันมาก ๆ ก็จะมีโอกาสอยู่ในคลาสนี้มากขึ้น โดยค่าลักษณะของสีอาจจะเหมือนหรือต่างกันได้

### 4. วางอยู่และถูกวางทับ (Place On, PO and be Placed By, PB)

ซูเปอร์พิกเซล  $S_i$  และ  $S_j$  วางทับกันอยู่ เช่นเดียวกับกับคลาสปิดบัง คลาสนี้จะแบ่งเป็นสองคลาสย่อยคือ  $S_i$  วางอยู่บน  $S_j$  รวมทั้ง  $S_i$  ถูกวางทับโดย  $S_j$  ค่าความลึกของทั้งสองซูเปอร์พิกเซลจะต่อเนื่องกันและมีค่าอยู่ระหว่างซูเปอร์พิกเซลที่เป็นฐาน คลาสนี้พบได้บ่อยเมื่อพิจารณาพื้น หรือ ชั้นบันได ตัวอย่างความสัมพันธ์นี้แสดงในรูปที่ 3.4 ง)

ลักษณะที่ชัดเจน คือ ความยาวการเชื่อมต่อระหว่างสองซูเปอร์พิกเซลโดยเทียบกับขนาดของซูเปอร์พิกเซลเอง โดยวัตถุที่อยู่ภายในมักจะมีอัตราส่วนที่เยาะกว่าประเภทอื่น ๆ นอกจากนี้ค่า  $\text{Max}(Y_i) < \text{Max}(Y_j)$  และ  $\text{Min}(Y_i) < \text{Min}(Y_j)$  เสมอ คลาสนี้สามารถตรวจพบได้ง่ายเมื่อสามารถรู้ได้ว่าซูเปอร์พิกเซลนั้นคือพื้น ลักษณะเด่นของพื้นคือตำแหน่งที่ปรากฏในภาพ ส่วนใหญ่จะปรากฏในส่วนล่าง ๆ ของภาพ ลักษณะของสีจำนวนหนึ่งจะเป็นพื้นและทิศทางของเส้นตรง เส้นตรงมักจะไม่ชี้ไปในแนวตั้งขึ้น

#### 3.2.2 ตัวจำแนกประเภท (Classifier)

ในหัวข้อนี้จะกล่าวถึงการพิจารณาวิธีการจำแนกประเภทความสัมพันธ์ จากนิยามของประเภทความสัมพันธ์ทั้ง 6 คลาสที่ได้นิยามไว้ในหัวข้อ 3.2.1 ขึ้นต่อไปคือการวิเคราะห์ตีความหมายจากรูปภาพนำเข้าและแยกแยะประเภทความสัมพันธ์เป็นคลาสใดคลาสหนึ่งใน 6 คลาสนั้น



วิธีการจำแนกที่งานวิจัยนี้ได้หยิบยกมาทดสอบนั้นคือวิธีเคเนียร์เอส เนียเบอร์ หรือเคเอ็นเอ็น (K-nearest neighbor Classifier, *Knn*) ซึ่งหลายละเอียดของขั้นตอนวิธีนี้ได้กล่าวไว้ในหัวข้อที่ 2.1.6 เนื่องจากประเด็นในการทำงานวิจัยนี้อยู่ที่การทดสอบจุดเด่นที่บ่งชี้ความสัมพันธ์ระหว่างซูเปอร์พิกเซล การนำวิธีนี้มาใช้จึงช่วยให้การวิเคราะห์ความแม่นยำนั้นง่ายขึ้น เพราะการจำแนกประเภทนั้นตรงไปตรงมา เพียงวัดค่าว่าต่างจากที่เป็นไปเท่าไรนั่นเอง

กำหนดให้  $F_i$  คือเซตของค่าจุดเด่น (Set of Features) ที่ได้จากการวิเคราะห์ซูเปอร์พิกเซล  $S_i$  และ ให้  $f_{ij}$  เป็นค่าจุดเด่นแต่ละค่าในเซต  $F_i$  โดย  $f_{ij} \in F_i$  ค่าระยะห่างที่จะใช้ในขั้นตอนวิธีเคเอ็นเอ็น จะสามารถเขียนให้อยู่ในรูปสมการที่ (3.6)

$$d_k = \left[ \sum_{j=1}^N (f_{i,j} - f_{k,j})^2 \right]^{1/2} \quad (3.6)$$

โดย  $k$  คือคลาสที่ 1 ถึง 6 ตามที่นิยามไว้  $d_k$  คือระยะห่างระหว่างค่าจุดเด่นของ  $S_i$  กับค่าจุดเด่นของ คลาสที่  $k$  และสามารถตัดสินใจ  $S_i$  ให้อยู่ในคลาส  $R$  โดยที่

$$d_R = \min_{k=1}^6 [d_k] \quad (3.7)$$

สำหรับการสกัดค่าจุดเด่นสำหรับนำมาใช้ในการจำแนกประเภทนี้จะกล่าวในหัวข้อถัดไป

### 3.2.3 การสกัดค่าจุดเด่น (Feature Extraction)

งานวิจัยนี้ได้เสนอการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุภายในรูปภาพ เนื่องจากความสัมพันธ์แต่ละประเภทมีจุดเด่นที่สามารถช่วยแยกแยะประเภทความสัมพันธ์ได้ แม้จะมีข้อมูลเป็นเพียงภาพสองมิติเพียงภาพเดียว งานวิจัยชิ้นนี้ได้นำเสนอจุดเด่นเหล่านั้นและวิธีที่สำหรับคำนวณค่าเหล่านี้จากภาพ จุดเด่นเหล่านี้สามารถแบ่งได้เป็น สามประเภทได้แก่

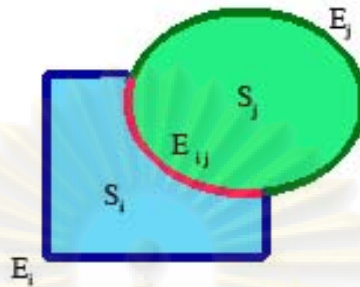
#### 3.2.3.1 จุดเด่นทางลายพื้นผิว (Texture feature)

คือจุดเด่นที่มาจากการวิเคราะห์คุณลักษณะค่าสีภายในซูเปอร์พิกเซล มีส่วนสำคัญในการช่วยในการระบุวัตถุในรูปภาพว่ามีโอกาสเป็นวัตถุอะไร หรือเป็นวัตถุเดียวกันหรือไม่ ค่าจุดเด่นทางลายพื้นผิวนี้อธิบายได้จากทั้งจากปริภูมิอาร์จีบี และปริภูมิเอชเอสวี

#### 3.2.3.2 จุดเด่นทางด้านเส้นขอบ (Edge feature)

คือจุดเด่นที่ได้มาจากการวิเคราะห์เส้นขอบของซูเปอร์พิกเซล วิเคราะห์ความยาว การเปลี่ยนแปลง ความยาวของเส้นเชื่อมต่อกัน มีส่วนสำคัญในการวิเคราะห์ว่าวัตถุปิดบังหรือถูกวางอยู่ นอกจากนี้ยังใช้ในการวิเคราะห์ว่าวัตถุอยู่ด้านหน้าหรือด้านหลัง ดังเช่นใน รูปที่ 3.5

เส้นขอบ  $E_j$  นั้นมีลักษณะคล้ายกับเส้นขอบ  $E_i$  ซึ่งเป็นเส้นขอบในส่วนรอยต่อระหว่าง  $S_i$  และ  $S_j$  แสดงให้เห็นว่า  $S_j$  นั้นอยู่ด้านหน้า  $S_i$  หรือ  $S_j \cap S_i$  นั่นเอง



รูปที่ 3.5 แสดงลักษณะของเส้นขอบในกลุ่มประเภทการปิดบัง

กำหนดให้  $E_i$  คือเซตของคู่อันดับ  $x, y$  ที่เป็นขอบของซูเปอร์พิกเซล  $S_i$  ซึ่งสามารถเขียนให้อยู่ในรูปสมการดังสมการที่ (3.8)

$$E_i = \{(x, y) | (x \pm 1, y \pm 1) \in S_i\} \quad (3.8)$$

จากเซตของคู่อันดับ  $E_i$  งานวิจัยนี้ได้นำมาแปลงเป็นค่าการเปลี่ยนแปลงหรือเรียกกันทั่วไปว่ารหัสลูกโซ่ (Chaincode) ซึ่งคือผลจากการหาดิฟเฟอเรนเชียลนั่นเอง (3.9)

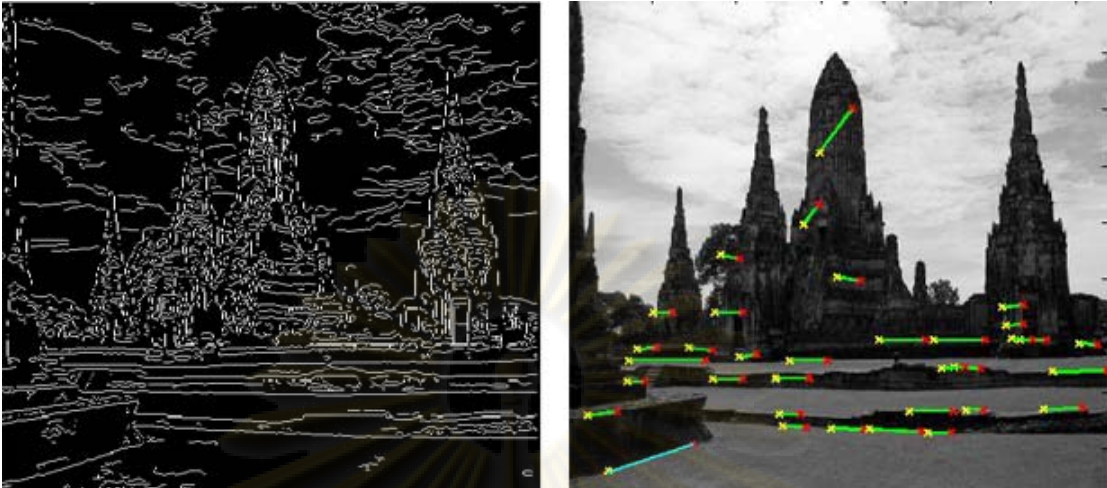
$$C_i = \Delta E_i \quad (3.9)$$

รหัสลูกโซ่จะบ่งบอกถึงทิศทางของเส้นขอบได้ว่าไปในทิศทางใด ผู้วิจัยใช้รหัสลูกโซ่นี้ในการบ่งบอกความต่อเนื่องของเส้นขอบ ซึ่งความต่อเนื่องคือการเก็บค่าการเปลี่ยนแปลงของทิศทางเส้นขอบ ว่าเกิดขึ้นบ่อยเพียงไร มีความยาวของเส้นตรงเฉลี่ยเท่าไร เป็นต้น

### 3.2.3.3 จุดเด่นสำหรับเส้นตรง (Straight line feature)

คือจุดเด่นที่ได้มาจากการวิเคราะห์ส่วนของเส้นตรง ซึ่งส่วนของเส้นตรงนี้ไม่ใช่ขอบที่เกิดมาจากซูเปอร์พิกเซล แต่เป็นเส้นตรงที่ได้มาจากการตรวจหาขอบ โดยใช้วิธีการตรวจหาเส้นขอบแคนนี่ (Canny edge detection)

ผู้วิจัยได้เลือกวิธีการตรวจหาเส้นขอบแคนนี่มาประยุกต์ใช้กับงานวิจัย เนื่องจากวิธีนี้มีจะได้ทั้งขอบเข้ม (Strong edge) และขอบอ่อน (Weak edge) ซึ่งเป็นผลมาจากที่ขั้นตอนวิธีนี้จะพยายามค้นหาทั้งขอบเข้มและขอบอ่อนที่สามารถนำมาประกอบกันเป็นพื้นผิวปิดได้ (Closed surface) ซึ่งเป็นผลทำให้ขั้นตอนวิธีนี้ได้ผลออกมาเป็นจำนวนมากดังที่แสดงในรูปที่ 3.6



รูปที่ 3.6 แสดงการหาค่าของเส้นตรงในภาพ ก) รูปผลลัพธ์จากการตรวจหาเส้นขอบโดยวิธีการตรวจหาเส้นขอบแบบนี้ ข) รูปผลลัพธ์การตรวจหาเส้นตรงโดยการใช้การแปลงฮัฟ

จากรูปที่ 3.6 ก) จะเห็นได้ว่าผลที่ได้จากการตรวจหาเส้นขอบแบบนี้มีจำนวนมากและมีลักษณะโค้งไปมา แต่เส้นที่ผู้วิจัยต้องการนำมาวิเคราะห์คือเส้นตรงเท่านั้น ดังนั้นในการตรวจหาผู้วิจัยจึงได้นำวิธีการตรวจหาเส้นตรงโดยการแปลงฮัฟ (Hough transform) มาตรวจหาเส้นตรงภายในรูป ซึ่งผลลัพธ์ที่ได้จากการตรวจหาได้แสดงในรูปที่ 3.6 ข)

#### 3.2.3.4 จุดเด่นทางตำแหน่ง (Position feature)

คือจุดเด่นที่มาจากวิเคราะห์ตำแหน่งของซูเปอร์พิกเซลนั้น ค่าตำแหน่งที่สำคัญคือค่า  $y$  หรือค่าตามแนวแกนตั้ง เนื่องจากการถ่ายภาพแบบทัศนมิติจะได้รูปภาพที่ตำแหน่งความลึกแปรผันตรงกับค่า  $y$  กล่าวคือถ้า  $y$  มีค่ามากแสดงว่าวัตถุนั้นมีโอกาสที่อยู่ลึกเข้าไปมาก ดังนั้นขณะวิเคราะห์ความสัมพันธ์ค่า  $y$  จึงมีส่วนช่วยในการคาดคะเนว่าซูเปอร์พิกเซลโดยอยู่ด้านหน้าหรือด้านหลังได้

ทั้งนี้ในการสกัดค่าจุดเด่น เป็นการสกัดค่าจุดเด่นจากทั้ง  $S_i$  และ  $S_j$  เพื่อนำมาหาความสัมพันธ์ระหว่างกัน งานวิจัยนี้จึงได้แบ่งย่อยค่าจุดเด่นทั้งสามประเภทออกเป็นส่วนย่อย ๆ อีก กล่าวคือ เซตของค่าจุดเด่นที่ได้จาก  $S_i$  แทนโดย  $F_i$  เซตของค่าจุดเด่นที่ได้จาก  $S_j$  แทนโดย  $F_j$  ซึ่ง  $F_i$  กับ  $F_j$  นั้นจะมีค่าจุดเด่นแบบเดียวกัน สำหรับเซตของค่าจุดเด่นที่ได้จากการพิจารณา  $S_i$  และ  $S_j$  พร้อมกัน แทนโดย  $F_{ij}$  จุดเด่นทั้งหมดสามารถสรุปได้ดังตารางที่ 3.2

ตารางที่ 3.2 สรุปค่าจุดเด่นสำหรับการจำแนกประเภทความสัมพันธ์ระหว่าง  $S_i$  กับ  $S_j$

	$F_i$ หรือ $F_j$	$F_{ij}$
Texture Feature	mean RGB std RGB prctile(10) prctile(30) prctile(70) prctile(90) Random H in sup	TextureDiff-colorStd TextureDiff-colorMean TextureDiff-hMean TextureDiff-hStd
Edge Feature	min of EdgeStat max of EdgeStat mean of EdgeStat std of EdgeStat length of EdgeStat Critical Point count	min of EdgeStat max of EdgeStat mean of EdgeStat std of EdgeStat length of EdgeStat Critical Point count Connection Angle B to AB Connection Angle A to AB length of edge AB EdgeDiff Mean-a2inter EdgeDiff Mean-b2inter EdgeDiff Mean-a2b2inter EdgeDiff Max-a2inter EdgeDiff Max-b2inter EdgeDiff Max-a2b2inter
Straight Line Feature	Mean Line angle Direction	Diff Mean Line angle Direction
Position Feature	Min Y Max Y Mean Y	yDiff min-min yDiff min-max yDiff max-min yDiff max-max yDiff mean-mean

รายละเอียดวิธีการสกัดค่าจุดเด่นของแต่ละค่าจุดเด่น

1. (*min of EdgeStat, max of EdgeStat, mean of EdgeStat, std of EdgeStat, length of EdgeStat, Critical Point count*) ค่าจุดเด่นที่แสดงคุณลักษณะของเส้นขอบ ค่าเหล่านี้คำนวณจากการเปลี่ยนแปลงของรหัสสีของ โดเมนที่คิดว่าขอบมีการเปลี่ยนทิศทางบ่อยแค่ไหน และรักษาทิศทางเดิมได้ นานเท่าไร ค่าที่นำมาใช้ได้แก่ ค่าต่ำสุด ค่ามากที่สุด ค่าความเบี่ยงเบนเฉลี่ย ความยาวเส้นขอบทั้งหมดและ จุดวิกฤต ถ้าขอบวัตถุมีลักษณะกลมเส้นขอบจะมีการเปลี่ยนแปลงเสมอ ตรงกันข้ามถ้าขอบเป็นเส้นตรงก็จะมีการเปลี่ยนแปลงน้อย
2. (*Mean RGB, std RGB, Random H*) ค่าจุดเด่นแสดงลักษณะของสี นำค่าของ RGB แต่ละช่องสีมาหาค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานโดยแยกทีละช่องสี รวมทั้งสุ่มค่าสีระบบ HSV เฉพาะช่องสี h มาหนึ่งตำแหน่ง
3. (*percentile 10<sup>th</sup> 30<sup>th</sup> 50<sup>th</sup> 70<sup>th</sup> 90<sup>th</sup> RGB*) ค่าจุดเด่นแสดงลักษณะของสี นำค่าสีมาสร้างฮิสโทแกรมจากนั้นคำนวณค่าของสีที่เปอร์เซ็นต์ไทล์ที่ 10 30 50 70 90 ของแต่ละช่องเปอร์เซ็นต์มาใช้
4. (*Min Y, Max Y, Mean Y*) ค่าจุดเด่นแสดงถึงตำแหน่ง ค่าตำแหน่ง Y ของช่องเปอร์เซ็นต์ เป็นพิกัด Y จากภาพ ค่าที่ใช้คือ ค่าสูงสุด ต่ำสุดและค่าเฉลี่ยของแต่ละช่องเปอร์เซ็นต์
5. (*Mean Line Angle Direction*) ค่าจุดเด่นแสดงถึงส่วนของเส้นตรง คือค่าเฉลี่ยของเส้นตรงทุกเส้นที่อยู่ในช่องเปอร์เซ็นต์นั้น ๆ
6. (*min of EdgeStat, max of EdgeStat, mean of EdgeStat, std of EdgeStat, length of EdgeStat, Critical Point count, length of edge AB*) ค่าจุดเด่นระหว่างช่องเปอร์เซ็นต์ที่แสดงลักษณะของเส้นขอบ ค่าเหล่านี้คำนวณจากการเปลี่ยนแปลงของรหัสสีของ แต่เป็นเส้นขอบที่เป็นรอยต่อระหว่างช่องเปอร์เซ็นต์
7. ค่าจุดเด่นระหว่างช่องเปอร์เซ็นต์ที่แสดงลักษณะการเชื่อมต่อ (*Connection Angle B to AB, Connection Angle A to AB*) วัดระยะของขอบที่เป็นรอยต่อเป็นมุมที่ทำกับจุดศูนย์กลางของแต่ละช่องเปอร์เซ็นต์
8. (*TextureDiff-colorStd, TextureDiff-colorMean, TextureDiff-hMean, TextureDiff-hStd*) ค่าจุดเด่นระหว่างช่องเปอร์เซ็นต์ที่แสดงความแตกต่างของลายผิว คำนวณจากค่าเฉลี่ยของสีของแต่ละช่องเปอร์เซ็นต์แล้วนำมาหาค่าที่ใกล้เคียงกัน ค่าที่ใช้คือ ค่าเฉลี่ย RGB ค่าเบี่ยงเบนมาตรฐาน RGB ค่าเฉลี่ย hue จาก HSV และ ค่าเบี่ยงเบนมาตรฐาน hue ของ HSV
9. (*EdgeDiff Mean-a2inter, EdgeDiff mean-b2inter, EdgeDiff Mean-a2b2inter, EdgeDiff Max-a2inter, EdgeDiff Max-b2inter, EdgeDiff Max-a2b2inter*) ค่าจุดเด่นระหว่างช่องเปอร์เซ็นต์ที่แสดงความแตกต่างของเส้นขอบ คำนวณจาก



ผลต่างของค่าเฉลี่ย และผลต่างของค่าสูงสุดของค่าความต่อเนื่องของเส้นขอบ มาเปรียบเทียบกับกันระหว่าง  $E_i$ ,  $E_j$  และ  $E_{ij}$

10. (Diff Mean Line angle Direction) ค่าจุดเด่นระหว่างซูปเปอร์ฟิกเชลที่แสดงถึงความแตกต่างของเส้นตรง คำนวณจากผลต่างของค่าเฉลี่ยของเส้นตรงที่อยู่ในซูปเปอร์ฟิกเชล

ค่าจุดเด่นที่ใช้ในงานวิจัยนี้มีทั้งหมด 83 ค่า แบ่งเป็นของ  $F_i$ ; 29 ค่า  $F_j$ ; 29 ค่า และ  $F_{ij}$ ; 25 ค่า ในการแบ่งแยกประเภทจะนำทั้ง  $F_i$ ;  $F_j$  และ  $F_{ij}$  มาใช้ทั้งหมดและทดสอบประสิทธิภาพของแต่ละค่าจุดเด่นต่อไป

### 3.3 การสร้างแบบจำลองสามมิติ

ในหัวข้อนี้จะกล่าวถึงการนำความสัมพันธ์ระหว่างวัตถุมาสร้างเป็นฉากสามมิติขึ้นมา เริ่มจากการนำความสัมพันธ์ของแต่ละวัตถุมาประมาณค่าความลึกในฉาก จากนั้นนำค่าความลึกและความสัมพันธ์นี้ไปสร้างเป็นแบบจำลองโดยมีค่าพิกัดเป็นสามมิติ สุดท้ายคือการส่งออกฉากสามมิตินี้ให้อยู่ในรูปแบบวีอาร์เอ็มแอล 2.0 (VRML 2.0)

#### 3.3.1 การหาความลึกของแต่ละจุดภาพ

งานวิจัยนี้ได้นำเสนอขั้นตอนวิธีในการใช้ความสัมพันธ์ของแต่ละวัตถุเพื่อหาตำแหน่งในพิกัดสามมิติ ทั้งนี้ในการหาค่าความลึกของจุดภาพจะแบ่งเป็นสองขั้นตอนนั่นคือการหาค่าความลึกสัมพัทธ์ (Relative depth) และการหาค่าความลึกสัมบูรณ์ (Absolute depth)

##### 3.3.1.1 การหาค่าความลึกสัมพัทธ์

คือความลึกที่ใช้สำหรับบอกตำแหน่งของซูปเปอร์ฟิกเชลอย่างคร่าว ๆ สามารถบอกได้เพียงว่าซูปเปอร์ฟิกเชลนี้อยู่ด้านหน้า ด้านหลัง หรืออยู่ภายในเท่านั้น งานวิจัยนี้ได้สร้างกราฟแสดงความสัมพันธ์ระหว่างซูปเปอร์ฟิกเชลขึ้นมา โดยแต่ละโหนดแสดงถึงซูปเปอร์ฟิกเชล และแต่ละเส้นเชื่อมคือความสัมพันธ์ระหว่างซูปเปอร์ฟิกเชล

ในแต่ละซูปเปอร์ฟิกเชล  $S$  จะถูกกำหนดค่าดัชนีความลึก  $di(S)$  การกำหนดค่าจะทำจากซูปเปอร์ฟิกเชลที่อยู่ด้านบนลงล่าง (Top down) โดยจะกำหนดค่าดัชนีความลึก  $di(S_0)$  ให้กับซูปเปอร์ฟิกเชลด้านบนซ้ายสุดและจะใช้เป็นตัวอ้างอิงสำหรับซูปเปอร์ฟิกเชลอื่นต่อไป โดยการกำหนดดัชนีความลึกนี้จะกำหนดตามความสัมพันธ์  $R_{ij}$  ของแต่ละซูปเปอร์ฟิกเชล ซึ่งเป็นไปตามที่นิยามไว้ในคลาสความสัมพันธ์ ดังนี้

1.  $S_i$  เป็นวัตถุเดียวกันกับ  $S_j$  ( $R_{ij} = SO$ ) กำหนดให้  $di(S_i) = di(S_j)$
2.  $S_i$  ปิดบัง  $S_j$  ( $R_{ij} = OT$ ) กำหนดให้  $di(S_i) < di(S_j)$
3.  $S_i$  ถูกปิดบังโดย  $S_j$  ( $R_{ij} = OT$ ) กำหนดให้  $di(S_i) > di(S_j)$

4.  $S_i$  อยู่ระนาบเดียวกับกับ  $S_j$  ( $R_{ij} = SP$ ) กำหนดให้  $di(S_i) = di(S_j)$
5.  $S_i$  วางอยู่บน  $S_j$  ( $R_{ij} = PT$ ) กำหนดให้  $\min(di(S_j)) < di(S_i) < \max(di(S_j))$
6.  $S_i$  ถูกทับโดย  $S_j$  ( $R_{ij} = PB$ ) กำหนดให้  $\min(di(S_i)) < di(S_j) < \max(di(S_i))$

หมายเหตุ จะสังเกตได้ว่าถ้ามีความสัมพันธ์แบบ  $PT$  หรือ  $PB$  ซูเปอร์พิกเซลนั้นจะ ได้ดัชนีความลึกแบบเป็นช่วงซึ่งเรียกว่า ดัชนีความลึกแบบช่วง (range depth index) ซึ่งเป็นส่วนสำคัญในการพิจารณาค่าความลึกสัมบูรณ์ต่อไป

### 3.3.1.2 การหาค่าความลึกสัมบูรณ์

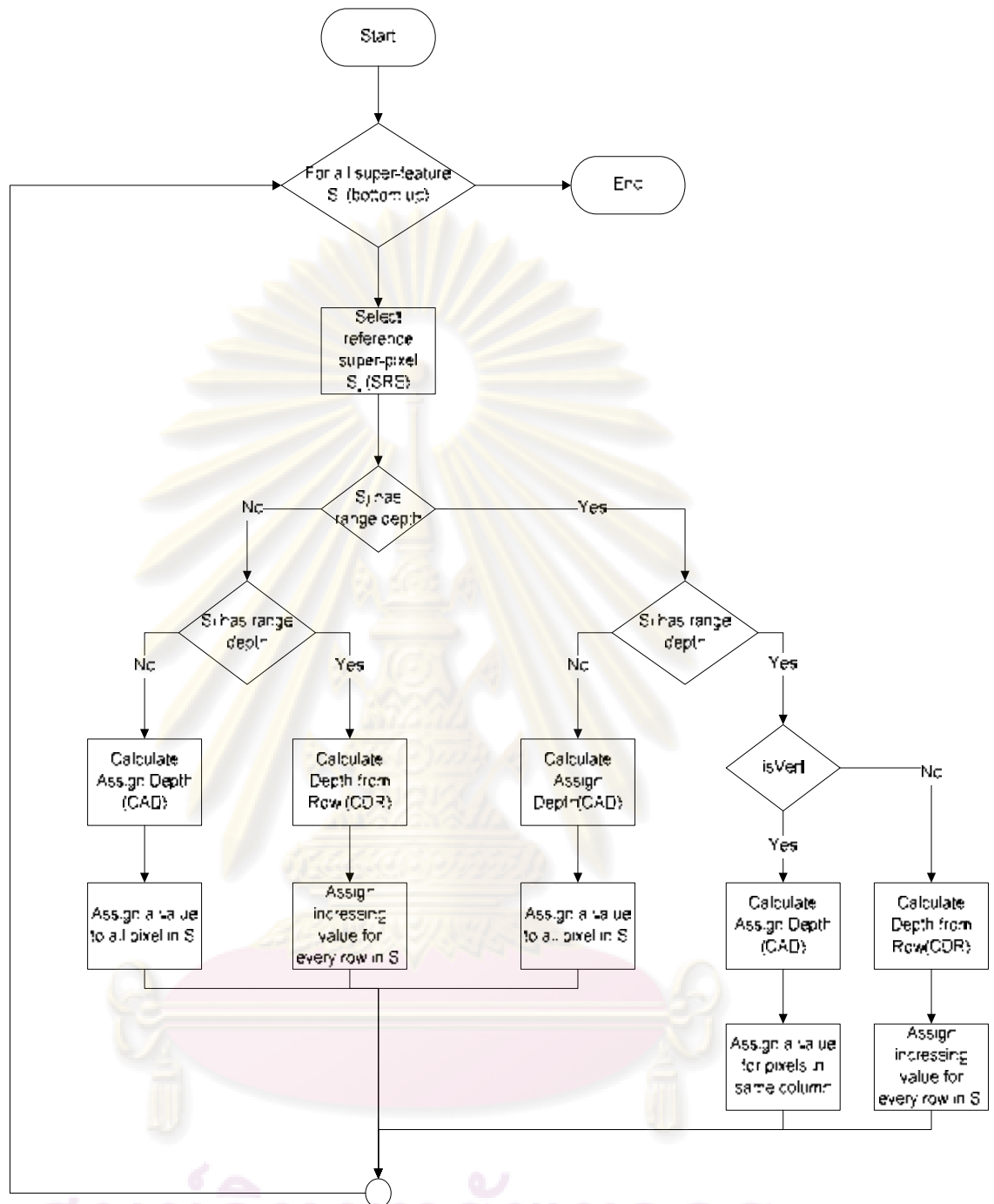
การหาค่าความลึกสัมบูรณ์  $d$  สำหรับทุกจุดภาพ ใช้หลักการการประมาณ ตำแหน่งโดยอ้างอิงถึงตำแหน่งที่วัตถุเทียบกับตำแหน่งของพื้น ซึ่งงานวิจัยนี้ได้ตั้งสมมติฐานว่า วัตถุในรูปภาพนำเข้ามาทุกชั้นไม่ได้ลอยอยู่ ค่าความลึกสัมบูรณ์บนพื้นระนาบ  $d(x,y)$  ที่ตำแหน่ง  $x,y$  สามารถหาได้จากสมการที่ (3.10)

$$d(x, y) = \frac{h_c \times f}{|y_h - y|} \quad (3.10)$$

โดยที่  $x,y$  คือตำแหน่งใด ๆ ในภาพ  $h_c$  คือความสูงของกล้องถ่ายภาพวัดจากพื้น  $f$  คือ ระยะจุดรวมสายตาของกล้อง (camera's focal length) และ  $y_h$  คือความสูงของจุดรวมสายตา (horizon point of view) [27] โดยตำแหน่งจุดรวมสายตานี้ตำแหน่งที่เราเห็นพื้นดินสิ้นสุดและ กลายเป็นท้องฟ้าไป การที่รูปถ่ายแสดงให้เห็นว่าท้องฟ้าและพื้นดินที่ควรจะขนานกันกลับมามีบรรจบกันนั้นเป็นผลมาจากการฉายภาพแบบทัศนมิตินั่นเอง ทั้งนี้จุดรวมสายตาก็มีผลอย่างมากในความถูกต้องของการหาความลึก

งานวิจัยนี้ได้เสนอแนวทางการประมาณจุดรวมสายตา ทั้งนี้เนื่องจากรูปภาพ ปกติมักจะมีวัตถุบังจุดที่เป็นจุดรวมสายตาอยู่ ดังนั้นจึงไม่สามารถหาตำแหน่งที่แน่นอนได้ วิธีที่ นำเสนอคือการใช้ค่าเฉลี่ยตำแหน่งของซูเปอร์พิกเซลที่น่าจะเป็นท้องฟ้า โดยซูเปอร์พิกเซลที่เป็นท้องฟ้า นั้นสามารถดูได้จาก การที่วัตถุอื่นบังหมด หรือมีความสัมพันธ์แบบถูกปิดบัง ( $R_{ij} = OB$ ) กับวัตถุอื่น และมีสีใกล้เคียงกับสีฟ้า

หลักการของขั้นตอนวิธีในการหาค่าความลึกสัมบูรณ์คือการกำหนดค่าความลึกสัมบูรณ์ให้แก่ซูเปอร์พิกเซลจะกระทำที่ละซูเปอร์พิกเซล โดยอ้างอิงจากดัชนีที่ได้จากความลึกสัมพัทธ์ ขั้นตอนวิธีในการหาความลึกสัมบูรณ์นี้สามารถแสดงเป็นแผนภูมิสายงานได้ในรูปที่ 3.7 จากแผนภูมิสายงานจะเห็นได้ว่าค่าที่มีผลในการเลือกกรณีคือค่าดัชนีความลึกเป็นช่วง (depth range) เมื่อพิจารณาแล้วซูเปอร์พิกเซลที่จะมีค่าดัชนีความลึกเป็นช่วงนั้นจะมีค่าพื้นและวัตถุที่ตั้งอยู่บนพื้นเท่านั้น



รูปที่ 3.7 แผนภูมิสายงานแสดงกระบวนการประมาณความลึก

จากในแผนภูมิจะมีขั้นตอนย่อย ๆ ที่สำคัญ 3 ขั้นนั้นคือ *SRS*, *CAD*, *CDR* ซึ่งแต่ละขั้นตอนมีรายละเอียดดังนี้

1. การเลือกซูเปอร์พิกเซลอ้างอิงหรือเอสอาร์เอส (Select reference super-pixel, *SRS*) เป็นกระบวนการคำนวณหาซูเปอร์พิกเซลที่จะใช้อ้างอิง ซูเปอร์พิกเซลจะเป็นซูเปอร์พิกเซลอ้างอิงได้ก็ต่อเมื่อซูเปอร์พิกเซลมีค่าความลึกสมบูรณ์แล้ว และจุดศูนย์กลางทำมุมระหว่าง  $-\pi/4$  ถึง  $-3\pi/4$  และจะหาไปเรื่อย ๆ

จนกระทั่งถึงตำแหน่งล่างสุดของภาพ ทั้งนี้ในการคำนวณจะเพิ่มซูเปอร์ฟิกเซล สมมติขึ้นมาที่ด้านล่างสุดขึ้นมาหนึ่งแถว และซูเปอร์ฟิกเซลนี้จะมีค่าความลึก เริ่มต้นอยู่แล้ว ดังที่สามารถแสดงเป็นรหัสเทียมดังรูปที่ 3.8

---

**Algorithm 1 : SRS( $S$ )**


---

```

1:  Foreach  $Nb \in \text{neighbor of } S$ 
2:    If  $Nb$  is below
3:      Mem  $Nb$ .connection length
4:    endIf
5:  EndFor
6:  Reference Sup  $\leftarrow \text{Max}(Nb$ .connection length)
7:  If ( $Nb$  has absolute Depth)
8:    Return Reference Sup
9:  Else
10:   Call SRS(Reference Sup)
15: EndIF

```

---

รูปที่ 3.8 รหัสเทียมแสดงขั้นตอนวิธีการเลือกซูเปอร์ฟิกเซลอ้างอิง

2. การคำนวณค่าความลึกแต่งตั้งหรือซีเอดี (Calculate Assigning Depth, CAD) เป็นกระบวนการหาค่าความลึกที่เป็นตัวแทน สำหรับนำไปใช้เป็นค่าความลึกให้กับ ซูเปอร์ฟิกเซลส่วนหนึ่ง ในการหานั้นจะขึ้นอยู่กับความสัมพันธ์ระหว่างซูเปอร์ฟิกเซลนี้กับซูเปอร์ฟิกเซลที่ใช้อ้างอิงแบ่งเป็น 3 กรณีตามความสัมพันธ์  $R_{ij}$

- $R_{ij} = SO, SP$  ค่าความลึกค่าเดียวกันแต่ถ้า
- $R_{ij} = PT, PB$  ค่าความลึกจะใช้ค่าของฟิกเซลที่อยู่พื้นฐาน
- $R_{ij} = OT, OB$  จะกำหนดค่าความลึก  $d$  ของซูเปอร์ฟิกเซล  $S_i$  ตามสมการที่ (3.11)

$$d_i = d_j \pm (w_1 \times ocd + w_2 \times \frac{h_c \times f}{|y_h - j|}) \quad (3.11)$$

โดยที่  $ocd$  คือค่าคงที่สำหรับการปิดบัง โดยผู้วิจัยกำหนดค่านี้ไว้ที่ 10 หน่วย ซึ่งหมายความว่าถ้ามีความสัมพันธ์แบบปิดบังจะมีความแตกต่างของความลึกอย่างน้อย 10 หน่วย สำหรับ  $w_1$   $w_2$  คือค่าถ่วงน้ำหนักผู้วิจัยกำหนดค่านี้ไว้ที่ 0.5 และ 0.5 นอกจากนี้ในการคำนวณค่าความลึกแต่งตั้ง งานวิจัยนี้ได้ ออกแบบระบบสำหรับตรวจสอบความขัดแย้งไปด้วย เช่น  $S_j$  ปิดบัง  $S_i$  อยู่แต่  $S_j$  มีค่าน้อยกว่า ระบบสามารถตรวจหาความขัดแย้งเหล่านี้ได้ แต่การแก้ไขให้ถูกต้องนั้นเป็นเรื่องที่ยาก การเปลี่ยนความสัมพันธ์ตำแหน่งหนึ่งจะกระทบกับความสัมพันธ์ที่ซูเปอร์ฟิกเซลอื่น ๆ การแก้ไขให้ถูกต้องจึงต้องแก้แล้ว สอดคล้องกันทั้งภาพ งานวิจัยนี้จึงได้เสนอวิธีแก้ไขในเฉพาะกรณีนี้

ความสัมพันธ์เป็นการปิดบัง และถูกปิดบังเท่านั้น โดยถ้าเกิดความขัดแย้งขึ้น ให้แก้โดยให้  $S_i$  นั้นอยู่ในระนาบเดียวกับ  $S_j$  แทน

3. การคำนวณค่าความลึกจากแถวหรือซีดีอาร์ (Calculate Depth form Row, CDR) เป็นวิธีคำนวณความลึกโดยประมาณว่าซูเปอร์พิกเซลนั้นวางตัวอยู่ในแนวระนาบเดียวกันกับพื้นตั้งนั้นค่าความลึกของจุดภาพที่  $x, y$  สามารถหาได้จากสมการการหาความลึกบนระนาบ สมการที่ (3.10) ค่าความลึกที่ได้จากการคำนวณแบบนี้จะมีลักษณะเพิ่มขึ้นแบบการเพิ่มแบบชี้กำลัง (exponential growth) เมื่อค่า  $y$  มีค่าเท่ากับ  $y_h$  ค่าความลึกที่จุดนั้นจะเป็นอนันต์ ซึ่งจากการทดลองของผู้วิจัยพบว่ากรณีนี้ที่จุดความลึกของภาพเป็นอนันต์นั้นพบไม่บ่อยนักและการเกิดกรณีนี้ขึ้นมีโอกาสทำให้มีความคลาดเคลื่อนสูง ผู้วิจัยจึงได้แก้ไขสมการที่ (3.10) เป็นสมการใหม่ดังสมการที่ (3.12)

$$d(x, y) = \frac{h_c \times f}{|c + y_h - y|} \quad (3.12)$$

โดย  $c$  คือค่าคงที่ผู้วิจัยกำหนด โดยทั่วไปค่านี้จะถูกกำหนดไว้ที่ 20

### 3.3.2 การขึ้นรูปแบบจำลองสามมิติ

หลังจากที่ได้ค่าความลึกสัมบูรณ์ของแต่ละจุดภาพมาแล้ว เมื่อนำไปรวมกับตำแหน่งของจุดภาพสองมิติ ก็ถือได้ว่าได้จุดตำแหน่งในพิกัดสามมิติมาแล้ว ทั้งต้องอย่าลืมว่ารูปภาพที่เป็นข้อมูลนำเข้านั้นเกิดจากการฉายภาพแบบทัศนมิติ ดังนั้นจึงต้องมีการฉายภาพย้อนกลับไปเป็นพิกัดสามมิติตามความเป็นจริง

หลักการการฉายภาพทัศนมิติย้อนกลับนั้นใช้หลักการของสามเหลี่ยมคล้ายมาคำนวณทั้งในแกน  $x$  และ แกน  $y$  กำหนดให้  $y'$  คือค่าตามแกน  $y$  ในรูปภาพนำเข้า และ  $x'$  คือค่าตามแกน  $x$  ในรูปภาพนำเข้า พิกัดสามมิติจริง  $(x_r, y_r)$  สามารถคำนวณได้จากสมการที่ (3.13)

$$x_r = \frac{x'}{f} z \quad (3.13)$$

$$y_r = \frac{y'}{f} z$$

โดยที่  $f$  คือระยะจุดรวมสายตาของกล้อง และ  $z = d(x', y')$

สิ่งที่ต้องคำนึงถึงต่อไปคือการสร้างแบบจำลองสามมิติคือการสร้างรูปหลายเหลี่ยม (polygon) ขึ้นมาจากกลุ่มของพิกัดสามมิติที่ได้มา ปัญหาของการสร้างรูปหลายเหลี่ยม



อยู่สองประการคือ การประมาณส่วนที่มองไม่เห็นจากภาพ และ วิธีการสร้างสามเหลี่ยมจากกลุ่มของพิกัดสามมิติ

### 3.3.2.1 การประมาณส่วนที่มองไม่เห็นจากภาพ

เนื่องจากงานวิจัยนี้ได้กำหนดให้ข้อมูลที่ใช้ในการสร้างแบบจำลองสามมิติคือรูปภาพเพียงรูปเดียวและไม่มีข้อมูลอื่น ๆ มาช่วย ดังนั้นในส่วนที่ถูกวัดถูบังอยู่จึงสามารถคาดคะเนถึงรูปร่างในส่วนนั้นได้ อย่างไรก็ดีตามด้วยการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุที่งานวิจัยนี้ได้นำเสนอ มีส่วนช่วยในการระบุการประมาณการเชื่อมต่อระหว่างส่วนย่อยของภาพได้ โดยวิธีการจะแบ่งตามคลาสดังนี้

- $SO, SP$  จะให้รูปหลายเหลี่ยมของทั้งสองเชื่อมต่อกัน แม้ว่าค่าความลึกจากทั้งสองซูเปอร์พิกเซลจะไม่เท่ากันก็จะใช้ค่าเฉลี่ยระหว่างซูเปอร์พิกเซล
- $OT, OB$  เส้นเชื่อมทั้งสองจะถูกแยกออกจากกันและสร้างให้ซูเปอร์พิกเซลที่อยู่ด้านหน้ามีความหนามากขึ้น
- $PT, PB$  โดยทั่วไปค่าความลึกที่ได้จากขั้นตอนวิธีที่เสนอนั้นส่วนฐานเชื่อมต่อกันและส่วนด้านบนจะแยกออกจากกันดังนั้น ก็จะสร้างความหนาให้กับส่วนที่แยกออกจากกันด้วย

### 3.3.2.2 วิธีการสร้างรูปหลายเหลี่ยมจากกลุ่มของพิกัดสามมิติ

การสร้างสามเหลี่ยมจากกลุ่มของพิกัดสามมิตินั้นเป็นปัญหาที่สำคัญ เนื่องจากในขั้นตอนการทำให้เป็นจุดภาพจากแบบจำลองสามมิติมาเป็นภาพสองมิตินั้นจะสามารถทำได้ก็ต่อเมื่อข้อมูลรูปร่างอยู่ในรูปแบบสามเหลี่ยมที่ต่อกันหรืออย่างน้อยมีลักษณะเป็นรูปหลายเหลี่ยมที่เป็นรูปนูน ผู้วิจัยได้ทดสอบวิธีการหาสามเหลี่ยมจำนวนหนึ่งและได้ผลสรุปที่ว่านำวิธีเดลลาอูใน มาประยุกต์ใช้จะให้ผลดีที่สุด เนื่องจากวิธีการอื่น ๆ จะทำให้รูปหลายเหลี่ยมที่ได้นั้นไม่ต่อเนื่องกัน

ขั้นตอนแรกของการสร้างรูปหลายเหลี่ยมคือการกำหนดจุดยอด (vertex) โดยผู้วิจัยจะกำหนดจุดยอดตามเส้นขอบของซูเปอร์พิกเซล โดยจุดยอดนี้คือจุดที่เป็นจุดวิกฤตของเส้นขอบทุกจุด ดังสมการที่ (3.14)

$$V_i = \{(x, y) \mid \Delta\Delta E_i(x, y) \neq 0\} \quad (3.14)$$

จุดยอดเหล่านี้ส่วนใหญ่จะมีลักษณะเป็นรูปเว้าไม่ใช่อรูปนูน หลักการคือต้องหาส่วนย่อย ๆ ที่เป็นรูปนูนมาประกอบกันให้ได้ตามรูปทรงที่ต้องการ ผู้วิจัยได้นำวิธีเดลลาอูในมาประยุกต์ใช้กับงานวิจัยนี้

เตลาอุโนจะได้ผลลัพธ์เป็นรูปพีระมิด (tetrahedral) สาเหตุที่เลือกใช้วิธีนี้ เนื่องจาก วิธีนี้จะได้สามเหลี่ยมที่มีลักษณะไม่เรียวยาว เนื่องจากวิธีการนี้ตั้งข้อกำหนดไว้ว่าให้ได้มุมน้อยสุดให้มีค่ามากที่สุด ซึ่งช่วยให้ได้ผลเป็นรูปหลายเหลี่ยมสมบูรณ์

แบบจำลองที่เป็นผลลัพธ์สุดท้ายนั้นจะอยู่ในรูปของรูปทรงหลายเหลี่ยมรูปเดียว และส่งออกไปเพื่อแสดงผลในรูปแบบของวีอาร์เอ็มแอล (VRML 2.0)

### 3.3.3 การส่งออกแบบจำลองสามมิติ

ผู้วิจัยได้เลือกรูปแบบวีอาร์เอ็มแอลมาเป็นรูปแบบในการส่งออกแบบจำลองสามมิติ ทั้งนี้เนื่องด้วยเป็นรูปแบบที่สามารถแสดงผลได้บนเว็บเพจ ซึ่งจะช่วยให้การเผยแพร่ผลงานนั้นทำได้ง่าย นอกจากนี้การรูปแบบวีอาร์เอ็มแอล นั้นง่ายต่อการตรวจสอบ เพราะจะเก็บอยู่ในรูปแบบข้อความ สามารถอ่านและวิเคราะห์ได้แม้ไม่มีเครื่องมือช่วย สำหรับวากยสัมพันธ์ของ วีอาร์เอ็มแอล 2.0 นั้นสามารถอ่านเพิ่มเติมได้ที่ [28]

### 3.4 วิเคราะห์ประสิทธิภาพความซับซ้อนเชิงเวลา

หัวข้อนี้จะวิเคราะห์ประสิทธิภาพโดยการพิจารณาความซับซ้อนเชิงเวลาในแต่ละขั้นตอน โดยจะแสดงอยู่ในรูปของสัญกรณ์โอใหญ่

- การแบ่งย่อยรูปภาพ การแบ่งย่อยรูปภาพด้วยวิธีอีจีบีเอสนั้นมีความซับซ้อนเชิงเวลาเป็น  $O(n \log n)$  [23] โดย  $n$  คือจำนวนจุดภาพทั้งหมดในรูปภาพ  $I$
- การจำแนกประเภทความสัมพันธ์ โดยขั้นตอนวิธีเคเอ็นเอ็น มีความซับซ้อนเชิงเวลาเป็น  $O(n_f)$  โดย  $n_f$  คือจำนวนค่าจุดเด่นซึ่งในงานวิจัยนี้ได้นำเสนอทั้งหมด 83 ค่าจุดเด่น ดังนั้นสามารถประมาณได้เป็น  $O(1)$
- การสกัดค่าจุดเด่น ในการตรวจหาค่าจุดเด่นที่เป็นลายฉนวนนั้นต้องคำนวณจากทุกจุดภาพดังนั้นจึงมีความซับซ้อนเชิงเวลาเป็น  $O(n)$  โดย  $n$  คือจำนวนจุดภาพทั้งหมดในรูปภาพ  $I$
- การหาความลึกแต่ละจุดภาพ ในกรณีที่แย่มากที่สุดคือ กรณีที่ใช้ขั้นตอนวิธีซีดีอาร์ สำหรับทุกซูเปอร์พิกเซลจะได้  $O(n)$  โดย  $n$  คือจำนวนจุดภาพและกรณีที่ดีที่สุดคือกรณีที่ทุกซูเปอร์พิกเซลใช้ขั้นตอนวิธี ซีเอดี นั่นคือ  $O(n_s)$  โดย  $n_s$  คือจำนวนซูเปอร์พิกเซล ซึ่งจำนวนซูเปอร์พิกเซลนั้นจะไม่เกิน 200 เนื่องจากได้ถูกกำหนดไว้ขณะทำการแบ่งย่อยรูปภาพด้วยวิธีอีจีบีเอส ดังนั้นในกรณีที่ดีที่สุดสามารถประมาณได้เป็น  $O(1)$

- การสร้างแบบจำลองสามมิติมีความซับซ้อนเชิงเวลาเป็น  $O(n_v^2)$  โดย  $n_v$  คือจำนวนจุดยอด ซึ่งกรณีที่แย่ที่สุดคือเมื่อ  $n_v$  มีจำนวนเท่ากับจำนวนจุดภาพ  $n$  ซึ่งจะได้เป็นความซับซ้อนเชิงเวลา  $O(n^2)$

สรุปโดยรวมแล้วขั้นตอนวิธีที่นำเสนอนี้มีประสิทธิภาพความซับซ้อนเชิงเวลาในกรณีที่ดีที่สุดอยู่ที่  $O(n \log n)$  และประสิทธิภาพความซับซ้อนเชิงเวลาที่แย่ที่สุดอยู่ที่  $O(n^2)$



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

## บทที่ 4

### การทดลองและวิเคราะห์ผล

งานวิจัยนี้มีจุดประสงค์เพื่อนำเสนอวิธีการประมาณฉากสามมิติโดยอัตโนมัติ ดังนั้นการทดสอบหลัก ๆ จึงเป็นการทดสอบเพื่อประเมินความแม่นยำของขั้นตอนวิธีที่ได้ นำเสนอไปและความสามารถในการนำไปใช้ในการประมาณฉากสามมิติได้จริง ซึ่งงานวิจัยนี้ได้ แบ่งการทดสอบเป็นสามขั้นตอนได้แก่

1. การทดสอบประสิทธิภาพของค่าจุดเด่นที่นำมาใช้ในการจำแนกประเภท
2. การทดสอบความแม่นยำของการจำแนกประเภทความสัมพันธ์
3. การทดสอบความแม่นยำการประมาณพิกัดสามมิติ

สำหรับความสวยงามของแบบจำลองสามมิตินั้นเป็นการตัดสินโดยตัวบุคคล จึงไม่มีการทดสอบแต่จะวิเคราะห์สิ่งที่แบบจำลองสามารถแสดงได้ เมื่อเทียบกับงานวิจัยอื่น

งานวิจัยนี้ได้สร้างโปรแกรมขึ้นมาเพื่อทดสอบตามแนวคิดที่ได้เสนอไป โดยงานวิจัยนี้ใช้โปรแกรม MATLAB 7.5.2 2007b เป็นเครื่องมือหลักในการสร้าง การทดสอบงานวิจัยนี้ทั้งหมดทำขึ้นบนเครื่องคอมพิวเตอร์ที่มีรายละเอียดดังนี้ Intel® Core™2 Duo Processor E7300 ความเร็ว 2.66 GHz เป็นซีพียูในการประมวลผล แรมขนาด 2 GB DDR2 ความเร็ว 667 MHz

รูปที่นำมาใช้ในการทดลองมีทั้งหมด 90 รูปภาพดังที่ได้แสดงไว้ในภาคผนวก สามารถแบ่งได้เป็นสองประเภทคือรูปที่เป็นรูปถ่ายกับรูปที่เป็นภาพสังเคราะห์ขึ้นจากโปรแกรมสามมิติ

- รูปที่เป็นรูปถ่ายมีจำนวนทั้งสิ้น 46 รูปภาพ เป็นรูปถ่ายทิวทัศน์ทั่วไป ทั้งที่ได้มาจากอินเทอร์เน็ต และรูปที่ผู้วิจัยถ่ายเอง
- รูปที่เป็นภาพสังเคราะห์จากโปรแกรมสามมิติมีจำนวนทั้งสิ้น 44 รูป รูปนี้ได้จากการสร้างแบบจำลองสามมิติแล้วจึงนำไปผ่านกระบวนการให้แสงและเงา (Render) โดยแบบจำลองนี้ถูกสร้างโดยโปรแกรม Autodesk 3ds max 2009 ซึ่งในแต่ละรูปจะมีลักษณะของวัสดุที่เป็นส่วนประกอบในภาพต่าง ๆ กันไป ประกอบด้วยทั้งแบบที่มีสีของวัตถุไม่ซ้ำซ้อน มีรูปทรงง่าย ๆ ไปจนถึงแบบจำลองมีสภาพแวดล้อมคล้ายของจริง

ขนาดของรูปภาพถูกกำหนดไว้ที่ 640x480 ซึ่งถ้ารูปภาพมีขนาดใหญ่หรือเล็กกว่าที่ได้กำหนดไว้จะถูกปรับขนาดจนมีขนาดเท่ากับที่ได้ตั้งไว้โดยอัตโนมัติ

อึ่งในการเรียนรู้ค่าจุดเด่นโดยวิธีการเคเอ็นเอ็นนั้นใช้รูปที่มีลักษณะความสัมพันธ์ชัดเจนกล่าวคือรูปที่มีความซับซ้อนไม่มากและตรงตามสมมติฐานของงานวิจัยนี้ รูปที่ใช้ในการเรียนรู้มีทั้งหมด 13 รูปประกอบด้วย รูปถ่าย 5 รูปและรูปที่สังเคราะห์ขึ้น 8 รูป มีค่าความสัมพันธ์ทั้งหมด 276 ค่า

#### 4.1 การทดสอบประสิทธิภาพของค่าจุดเด่น

การทดสอบประสิทธิภาพของค่าจุดเด่น คือการนำเอาค่าจุดเด่นที่สกัดได้จากรูปภาพ มาทดสอบความสามารถในการจำแนกประเภทความสัมพันธ์ของวัตถุว่ามีความแม่นยำมากน้อยเพียงใด ซึ่งความถูกต้องในการแยกแยะความสัมพันธ์นี้จะมีผลอย่างมากต่อความแม่นยำในการประมาณพิกัดสามมิติ จุดประสงค์ของการทดสอบนี้เพื่อแสดงให้เห็นว่าจากรูปภาพรูปหนึ่ง ๆ สามารถนำมาวิเคราะห์ความสัมพันธ์ของแต่ละวัตถุได้

##### 4.1.1 วิธีการทดสอบการกระจายตัวของค่าจุดเด่น

งานวิจัยนี้ได้ทดสอบการกระจายตัวของค่าจุดเด่นที่นำมาใช้ เพื่อสังเกตความเหมาะสมของแต่ละค่าจุดเด่นที่จะนำมาใช้ในการจำแนกประเภท สำหรับวิธีการทดสอบนั้นคือการหาค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน กำหนดให้  $F_n$  คือเซตของค่าจุดเด่น  $f_n$  และ  $f_k$  เป็นเซตของค่าจุดเด่นที่เป็นคลาส  $k$  ดังสมการที่ (4.1)

$$f_k = \{f_k \subset F_n \mid \text{class}(f_k) = k\} \quad (4.1)$$

ผู้วิจัยจะคำนวณค่าเฉลี่ยแยกตามคลาส ดังนั้นค่าเฉลี่ยจะมีค่าดังนี้

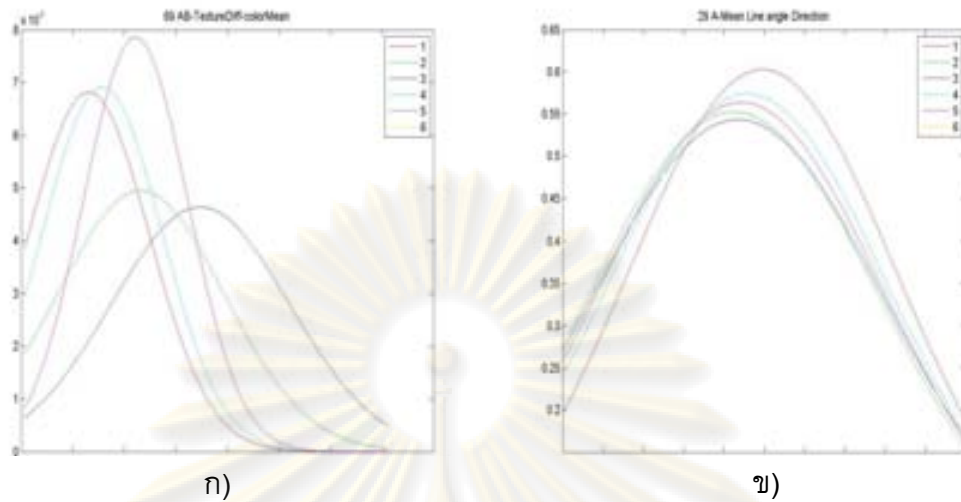
$$\mu_k = \frac{\sum_{i=1}^n f_{ki}}{n}, f_{ki} \in f_k \quad (4.2)$$

และค่าเบี่ยงเบนมาตรฐานคือ

$$\sigma_k = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_{ki} - \mu)^2}, f_{ki} \in f_k \quad (4.3)$$

สมมติให้ค่าจุดเด่นทุกประเภทมีการกระจายแบบปกติ (Normal distribution) จะสามารถนำค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานมาแสดงเป็นกราฟได้ดังรูปที่ 4.1





รูปที่ 4.1 กราฟแสดงการกระจายตัวของค่าจุดเด่น “ค่าความต่างของสีเฉลี่ย” เส้นแต่ละสีแสดงถึงคลาสแต่ละคลาส ก) กราฟแสดงถึงค่าจุดเด่นที่มีกระจายตัวที่ดี ข) กราฟแสดงถึงค่าจุดเด่นที่มีการกระจายตัวไม่ดี

สำหรับวิธีการประเมินความเป็นเอกลักษณ์หรือประสิทธิภาพของแต่ละค่าจุดเด่นนั้นจะใช้การเปรียบเทียบโดยอ้างอิงจากรูปนี้เป็นหลัก กำหนดนิยามค่าจุดเด่นที่ดีนั้นจะมีส่วนซ้อนทับกันน้อย รูปแบบที่ดีที่สุดใญ่สุดคือพื้นที่ใต้กราฟไม่มีการซ้อนทับกันเลย และรูปแบบที่แย่มากที่สุดคือกราฟทุกคลาสซ้อนทับกันพอดี การกระจายตัวที่ดีของค่าจุดเด่นนี้ไม่จำเป็นต้องทุกตัวกระจายกันหมด เพียงแค่มีคลาสใดคลาสหนึ่งแยกตัวออกมาอย่างชัดเจนก็ถือว่าเป็นการกระจายตัวที่ดี เนื่องจากในการวิเคราะห์จำแนกประเภทไม่ได้ค่าจุดเด่นเพียงค่าเดียว แต่การที่มีค่าจุดเด่นหนึ่งที่บ่งบอกเด่นชัดว่าเป็นคลาสใดจะสามารถช่วยในการจำแนกประเภทได้เป็นอย่างดี ตัวอย่างแผนภูมิการกระจายตัวที่ดีแสดงในรูปที่รูปที่ 4.1 ก) และตัวอย่างแผนภูมิการกระจายตัวที่ไม่ดีแสดงในรูปที่ รูปที่ 4.1 ข) ค่าคะแนนค่าประสิทธิภาพของเซตของค่าจุดเด่น  $F_n$  นั้นสามารถคำนวณได้จากสมการที่ (4.4)

$$s(F_n) = \sum_{i=1}^6 \sum_{j=1}^6 ttest(f_i, f_j)$$

$$ttest(f_i, f_j) = \frac{\mu_i - \mu_j}{\sqrt{\frac{\sigma_i^2}{n_i} + \frac{\sigma_j^2}{n_j}}} \quad (4.4)$$

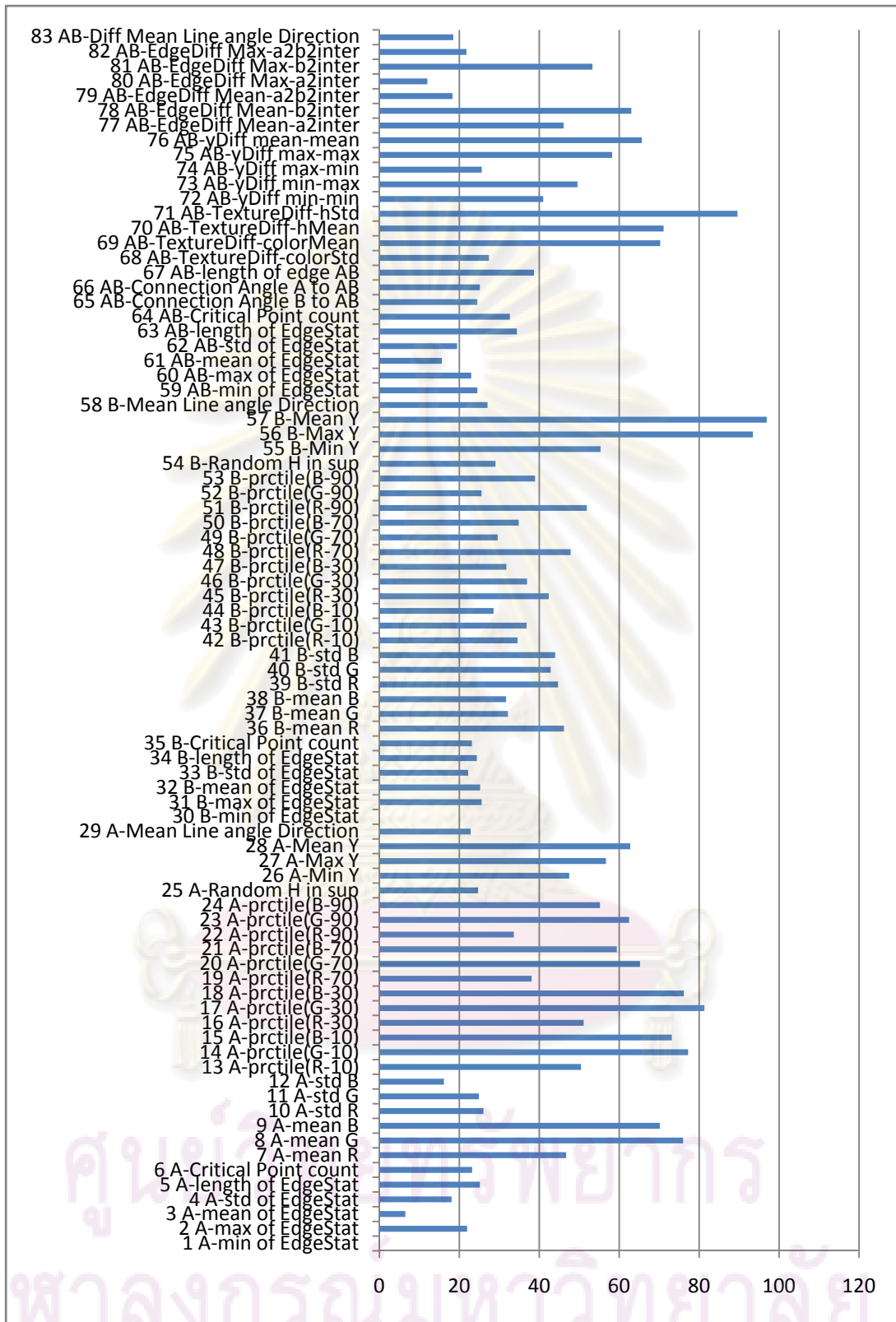
โดย  $\mu_i$  คือค่าเฉลี่ยของ  $f_i$  ,  $\sigma_i$  คือค่าเบี่ยงเบนมาตรฐานของ  $f_i$  และ  $n_i$  คือจำนวนตัวอย่างใน  $f_i$

#### 4.1.2 ผลการทดสอบประสิทธิภาพของค่าจุดเด่น

ผลการทดสอบการกระจายตัวของค่าจุดเด่นทั้ง 83 ค่า ซึ่งได้ระบุไว้ในหัวข้อที่ 3.2.3 โดยค่าจุดเด่นเหล่านี้ได้จากตัวอย่างที่ได้จากการจำแนกประเภทโดยผู้วิจัยจำนวน 276 ตัวอย่าง คะแนนการกระจายตัวแสดงในแผนภูมิรูปที่ 4.2 ทั้งนี้สำหรับคลาสที่ 6 PB นั้นจะไม่ถูกนำมาแสดงเนื่องจากคลาส PB ในรูปทั่วไปนั้นมีน้อย ทำให้การวัดการกระจายตัวมีความคลาดเคลื่อนมาก สาเหตุที่คลาส PB นั้นหาพบได้ยากเนื่องจากการกำหนดความสัมพันธ์จะมาจากส่วนย่อยด้านบนของรูปภาพลงมา แต่คลาส PB นั้นเกิดจากพื้นไปยังวัตถุ จึงเป็นสถานการณ์ที่พบได้ยาก

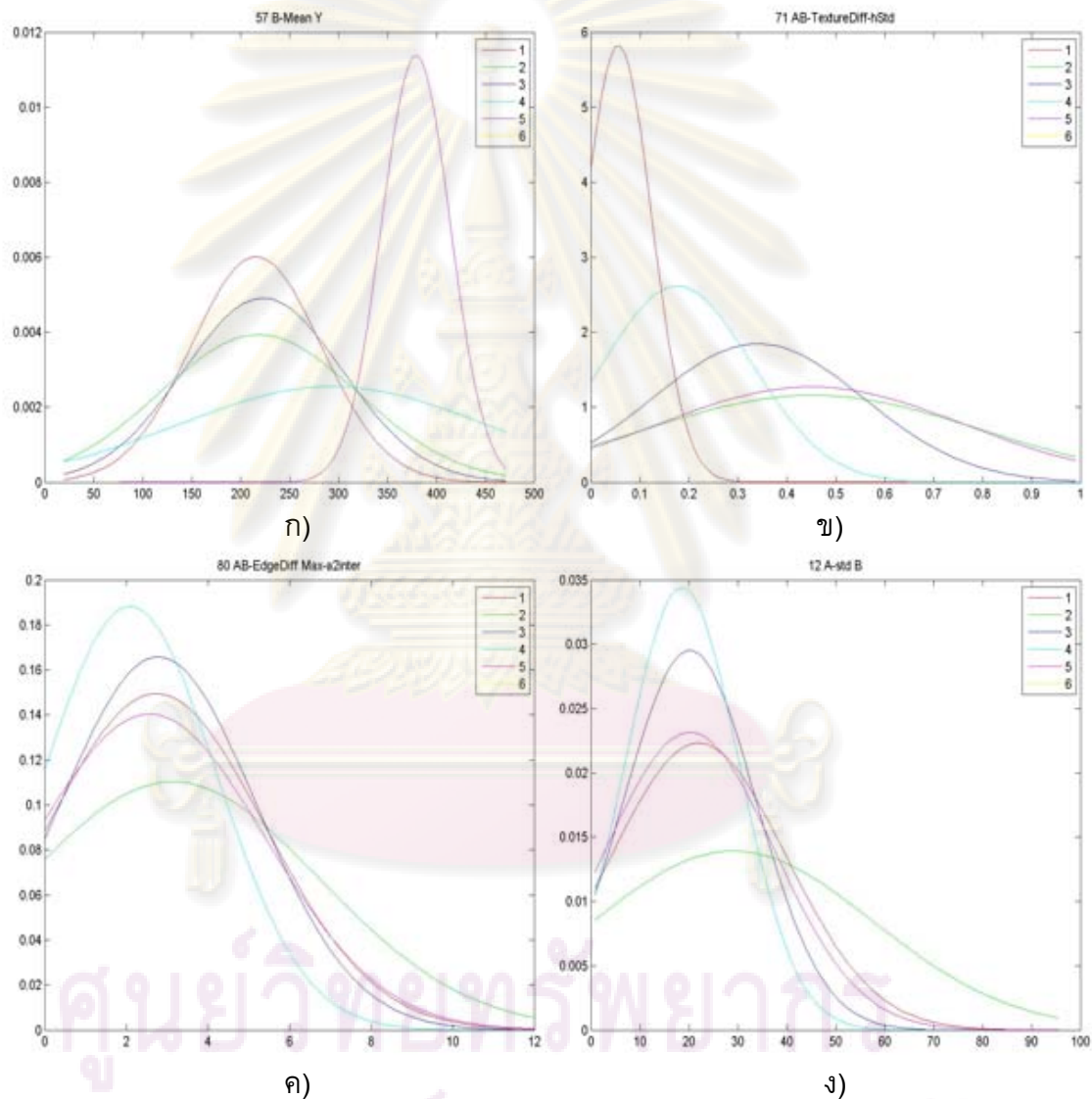


ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.2 แผนภูมิแสดงคะแนนค่าประสิทธิภาพการกระจายตัวของแต่ละค่าจุดเด่น

จากแผนภูมิที่แสดงในรูปที่ 4.2 ค่าจุดเด่นที่มีคะแนนมากที่สุด 5 อันดับแรก คือ 57 B-Mean Y, 56 B-max Y, 71 AB-TextureDiff-hstd , 17 A-prctile(G-30), 18 A-prctile(B-30) กราฟการกระจายตัวที่น่าสนใจของค่าจุดเด่น B-Mean Y และ AB-TextureDiff-hstd ผู้วิจัยได้นำมาแสดงในรูปที่ 4.3 ก) และข) ตามลำดับ และค่าที่มีคะแนนน้อยที่สุด 5 อันดับแรกได้แก่ 1 A-min of EdgeStat, 30 B-min of EdgeStat, 3 A-mean of EdgeStat, 80 AB-EdgeDiff Max-a2inter, 12 A-std B



รูปที่ 4.3 แผนภูมิแสดงการกระจายตัวของค่าจุดเด่น ก) แผนภูมิแสดงการกระจายตัวของค่าจุดเด่น B-mean Y ซึ่งมีคะแนน 96.90 ข) แผนภูมิแสดงการกระจายตัวของค่าจุดเด่น AB-TextureDiff-hstd ซึ่งมีคะแนน 89.58 ค) แผนภูมิแสดงการกระจายตัวของค่าจุดเด่น AB-EdgeDiff Max-a2inter ซึ่งมีคะแนน 11.98 ง) แผนภูมิแสดงการกระจายตัวของค่าจุดเด่น A-std B ซึ่งมีคะแนน 16.14 B

จากรูปที่ 4.3 จะเห็นว่าในแต่ละค่าจุดเด่นจะสามารถแบ่งคลาส ๆ ออกมาได้ชัดเจน ดังเช่น ยกตัวอย่างเช่น ค่าจุดเด่นที่ 57 *B-mean Y* ซึ่งแสดงการกระจายตัวในรูปที่ 4.3 ก) จะเห็นได้ว่าคลาส *PT* (วางอยู่) จะเด่นชัดออกมาและมีส่วนที่ทับกันน้อย โดยจะมีค่าเฉลี่ยอยู่ที่ประมาณ 400 และเบี่ยงเบนไม่มาก ดังนั้นถ้าวัตถุ *B* มีค่าใกล้เคียงกับ 400 มากก็มีโอกาสมากที่จะอยู่ในคลาสวางอยู่ ค่าจุดเด่นนี้สามารถแบ่งคลาสที่ *PT* ออกมาได้ชัดเจนเนื่องจากคลาสวางอยู่นั้นจะเกิดกับการวางอยู่บนพื้นน้อยที่สุด และพื้นมักจะอยู่ในส่วนด้านล่างของภาพค่า *Y* ของวัตถุ *B* จึงมีโอกาสมากที่จะมีค่ามากอยู่เสมอ

ค่าจุดเด่นที่ 71 *AB-TextureDiff-hstd* จากรูปที่ 4.3 ข. จะเห็นได้ว่าสามารถแบ่งคลาส *SO* ออกมาได้ชัดเจนเนื่องจาก วัตถุเดียวกันมักจะมีค่าเบี่ยงเบนเฉลี่ยของสีที่เหมือนกันโดยเฉพาะสำหรับค่า *Hue* ในช่องสี *HSV* นั้นจะสามารถบ่งบอกความเหมือนกันได้อย่างชัดเจน

ค่าจุดเด่นที่ 17 และ 18 ซึ่งเป็นค่าเปอร์เซ็นต์ที่ 30 ของสีเขียวและสีฟ้า ที่มีค่าชัดเจนจากการแบ่งคลาสที่ 3 เพราะคลาสที่สามส่วนใหญ่เกิดจากที่วัตถุปิดบังท้องฟ้าอยู่ ซึ่งท้องฟ้าก็จะกระจายตัวของค่าสีเขียวและสีฟ้าอยู่มากและที่เด่นชัดที่เปอร์เซ็นต์ที่ 30 เนื่องจากสีเขียวและสีฟ้านั้นมีค่าที่อ่อน

สำหรับค่าจุดเด่นที่มีการกระจายตัวที่ไม่ดี ยกตัวอย่างเช่นค่าจุดเด่นที่ 1 และ 30 ซึ่งเป็นค่าต่ำสุดของความต่อเนื่องของเส้นขอบ ที่ค่านี้ไม่ดีนั้นเพราะเกือบทุกเส้นขอบของส่วนย่อยจะมีค่านี้อยู่ที่ 1 ดังนั้นค่านี้จึงไม่สามารถบอกลักษณะเด่นได้ แต่ค่านี้ไม่มีผลมากนักเมื่อเทียบกับค่าจุดเด่นที่ 79 และ 80 (*AB-EdgeDiff Max a2inter, AB-EdgeDiff Mean a2b2inter*) ซึ่งเป็นค่าจุดเด่นผู้วิจัยตั้งสมมติฐานว่าจะช่วยในการระบุคลาสปิดบังได้ ซึ่งมีผลกระทบต่อการจัดประเภทของคลาสมาก การกระจายตัวของค่าจุดเด่นที่ไม่ดีนักถูกแสดงในรูปที่ 4.3 ค) ทั้งนี้อาจอธิบายสาเหตุได้ดังนี้ ประการแรกเป็นผลจากการแบ่งย่อยรูปภาพ การแบ่งย่อยบ่อยครั้งมักจะได้ไม่ได้เส้นขอบแท้ของวัตถุแต่จะได้เป็นเส้นขอบของส่วนย่อยเป็นส่วนใหญ่ ประการที่สองวัตถุที่อยู่ข้างหน้ากับข้างหลังอาจมีคุณลักษณะของเส้นขอบที่เหมือนกันได้ เช่นเดียวกับความสัมพันธ์แบบคลาสอื่น ๆ จึงทำให้ค่าจุดเด่นนี้ไม่สามารถจำแนกประเภทความสัมพันธ์ของวัตถุได้ อย่างไรก็ตามคลาสปิดบังนี้ยังมีค่าจุดเด่นที่ 77,78,81,82 ที่มีคะแนนความเด่นชัดพอประมาณช่วยแบ่งแยกได้บ้าง

นอกจากนี้ค่าที่มีการกระจายตัวไม่ดีนักคือค่าที่เป็นค่าเบี่ยงเบนเฉลี่ยของค่าสีต่าง ๆ เช่นค่าเบี่ยงเบนเฉลี่ยของสีฟ้า ซึ่งแสดงในรูปที่ 4.3 ง) เนื่องจากการกระจายของช่องสีใดช่องสีหนึ่งนั้นไม่ส่งผลต่อการจัดประเภทของความสำคัญมากนัก

โดยสรุปแล้วจากการวิเคราะห์การกระจายตัวของค่าจุดเด่นนี้ จะเห็นได้ว่า ค่าจุดเด่นที่นำเสนอสามารถบ่งบอกความสัมพันธ์ของวัตถุได้ โดยรูปภาพที่วิธีการนี้สามารถระบุ



ประเภทของคลาสที่ดีที่สุดคือ ค่าจุดเด่นที่เกี่ยวข้องกับตำแหน่ง  $Y$  และค่าจุดเด่นที่เกี่ยวข้องกับค่าสี กล่าวคือจะสามารถแบ่งคลาสที่ 1 SO และคลาสที่ 5 PT ออกมาได้ชัดเจนที่สุด

## 4.2 ค่าการความแม่นยำในการจำแนกประเภทความสัมพันธ์

เป็นการทดสอบความแม่นยำโดยรวมในการจำแนกประเภทความสัมพันธ์ ซึ่งความแม่นยำนี้จะขึ้นอยู่กับตัวแปรสองประการคือ ค่าจุดเด่นที่ใช้ และตัวจำแนกประเภท ซึ่งบทบาทของค่าจุดเด่นนั้นได้ถูกวิเคราะห์ไปในหัวข้อที่ผ่านมา และดังที่ได้กล่าวมาในขั้นตอนสาเหตุที่ผู้วิจัยได้เลือกตัวจำแนกประเภทเคเอ็นเอ็น มาใช้เพื่อจำกัดตัวแปรให้ตัวจำแนกประเภทเป็นแบบที่เรียบง่ายที่สุด ทั้งนี้เพื่อทดสอบจุดประสงค์หลักคือการทดสอบดูว่าเมื่อพิจารณาค่าจุดเด่นหลาย ๆ ค่าแล้วสามารถตัดสินความสัมพันธ์ของแต่ละส่วนย่อยได้แม่นยำมากน้อยเพียงใด หรือค่าจุดเด่นเหล่านี้มีความสอดคล้องกันมากน้อยเพียงใด

### 4.2.1 วิธีการทดสอบความแม่นยำในการจำแนกประเภทความสัมพันธ์

ในการทดสอบความแม่นยำนี้จะทำโดยการเปรียบเทียบผลลัพธ์ความสัมพันธ์  $R_{ij}$  ระหว่างความสัมพันธ์  $R_{ij}$  ที่ได้จากโปรแกรม กับ ความสัมพันธ์  $R_{ij}$  ที่ผู้วิจัยเป็นผู้ระบุด้วยมือตามความจริง และเปรียบเทียบเป็นร้อยละความถูกต้องของการตัดสินระบุความสัมพันธ์ในภาพทั้งหมด

นอกจากนี้ผู้วิจัยยังได้วิเคราะห์ว่าคลาสไหนเกิดการจำแนกผิดพลาดขึ้นมาก และคลาสใดเกิดความผิดพลาดขึ้นน้อย ค่าความผิดพลาดนั้นนับจากถ้า การจำแนกผิดพลาดจากคลาสใด และคลาสที่ถูกต้องควรเป็นคลาสใด ยกตัวอย่างเช่น การจำแนกประเภทโดยวิธีที่เสนอได้ผลเป็น  $k_m$  แต่คลาสที่ถูกต้องคือ  $k_t$  ก็จะนับเป็นความผิดพลาดจากทั้งสองคลาส ดังสมการที่ (4.5)

$$k_m \neq k_t \rightarrow err(k_m) + 1, err(k_t) + 1 \quad (4.5)$$

#### 4.2.2 ผลการทดสอบความแม่นยำในการจำแนกประเภทความสัมพันธ์

ผู้วิจัยได้ทดสอบกับรูปภาพ จำนวน 18 ภาพรวมทั้งสิ้น 360 ความสัมพันธ์ รูปภาพที่ใช้แบ่งเป็นรูปภาพจริงเก่าภาพและรูปภาพที่ได้จากโปรแกรมสามมิติเก่าภาพ ผลความแม่นยำในการจำแนกประเภทแสดงในตารางที่ 4.1

ตารางที่ 4.1 แสดงร้อยละความถูกต้องของการทดสอบการจำแนกประเภท

	ต่ำสุด	สูงสุด	โดยเฉลี่ย
ภาพจากโปรแกรมสามมิติ	37.5	100	67.55
ภาพจากรูปถ่ายจริง	40.74	83.33	63.01
ทั้งหมด	37.5	100	65.74

จากตารางที่ 4.1 แสดงให้เห็นว่าความแม่นยำที่ได้จากรูปภาพจริงและจากโปรแกรมสามมิตินั้นมีความแตกต่างกันไม่มากนัก โดยรูปภาพจากโปรแกรมสามมิตินั้นจะมีร้อยละความถูกต้องมากกว่าทั้งนี้เนื่องมาจากความชัดเจนของเส้นขอบและสีในภาพจากโปรแกรมสามมิติที่มีความซับซ้อนน้อยกว่า

ตารางที่ 4.2 แสดงความผิดพลาดจากการจำแนกประเภทแบ่งตามคลาส

Class	SO	OT	OB	SP	PT
$err(k)$	25	10	40	25	10
$N$	53	22	142	11	47
$Err(k) / N$	0.4717	0.4545	0.2817	2.2727	0.2127

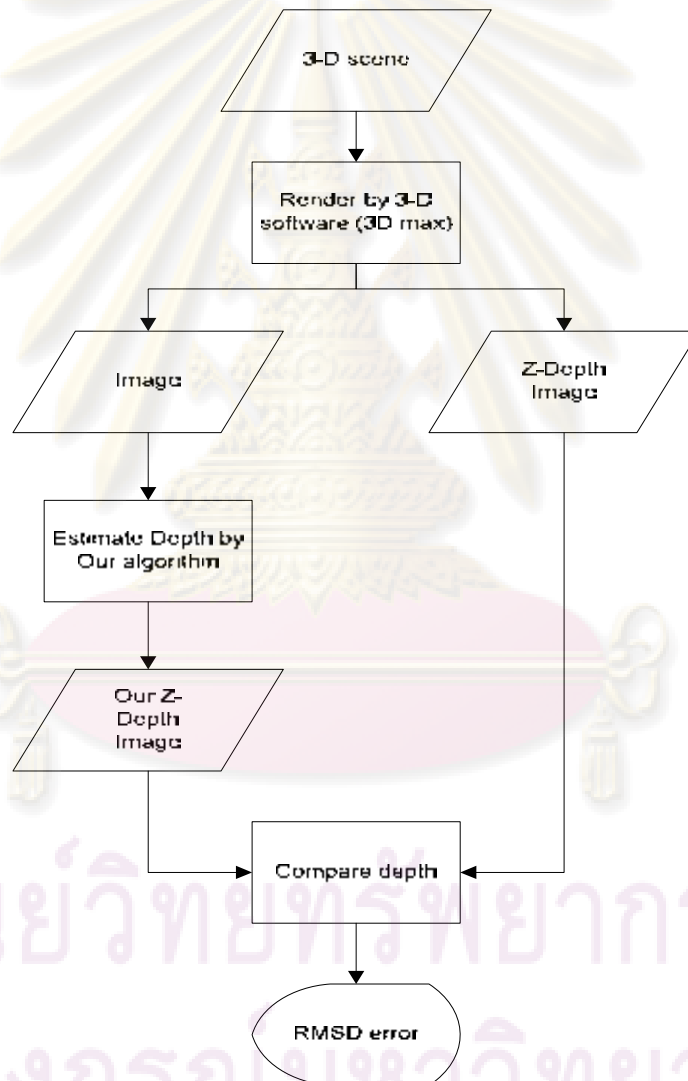
จากตารางที่ 4.2 จะเห็นได้ว่าคลาส  $PT$  นั้นมีความผิดพลาดน้อยซึ่งสอดคล้องกับค่าจุดเด่นในกลุ่มตำแหน่งที่ได้คะแนนมาก และคลาส  $OB$  ที่ส่วนใหญ่เป็นเกิดจากค่าจุดเด่นที่เป็นค่าสีที่ช่วยบ่งบอกว่าเป็นท้องฟ้าอยู่ ซึ่งพบได้เป็นจำนวนมากในภาพ สำหรับ  $SO$  และ  $OT$  นั้นก็มีค่าผิดพลาดไม่เยอะนัก แต่ที่น่าสังเกตคือ  $SP$  ที่ความผิดพลาดเกิดขึ้นมากทั้งนี้เนื่องจากค่าจุดเด่นที่ใช้เป็นตัวตัดสินการอยู่ระนาบเดียวกันคือส่วนของเส้นตรง ซึ่งจากคะแนนที่ได้ถือว่าเป็นค่าจุดเด่นที่ไม่ดีนักการตัดสินใจจำแนกประเภทของ  $SP$  นี้จึงมีโอกาสผิดพลาดมากกว่าคลาสอื่น ๆ

#### 4.3 การทดสอบความแม่นยำการประมาณพิกัดสามมิติ

ในหัวข้อนี้จะเป็นการทดสอบความแม่นยำของพิกัดสามมิติที่ได้จากขั้นตอนวิธีที่ได้นำเสนอ โดยจะแบ่งเป็นการทดสอบความแม่นยำที่ใช้ประเภทความสัมพันธ์ที่ได้จากการจำแนกโดยอัตโนมัติ และประเภทความสัมพันธ์ที่ตามจริง จุดประสงค์ของการทดสอบนี้เพื่อแสดงให้เห็นว่าสามารถนำเอาประเภทของความสัมพันธ์มาประมาณตำแหน่งพิกัดสามมิติได้

### 4.3.1 วิธีการทดสอบความแม่นยำการประมาณพิกัดสามมิติ

วิธีการทดสอบจะใช้ทำโดยการวาดแบบจำลองสามมิติ จากนั้นนำแบบจำลองนี้ไปผ่านการให้แสงและเงา หรือเรียกว่าการเรนเดอร์ (Render) โดยกำหนดให้มีข้อมูลส่งออกสองอย่างคือ รูปภาพ และค่าความลึก ซึ่งค่าความลึกนี้จะเป็ค่าความลึกจริง ๆ ตามแบบจำลองที่ได้สร้างขึ้นมา ส่วนรูปภาพที่ได้จะถูกนำไปเป็นให้แกระบบที่งานวิจัยนี้ได้นำเสนอออกไปและจะได้เป็นฉากสามมิติพร้อมพิกัดตำแหน่งสามมิติออกมา และการวัดความถูกต้องก็จะวัดโดยนำพิกัดสามมิติที่ได้จากขั้นตอนวิธีที่เสนอมาเปรียบเทียบกับค่าความลึกที่ได้จากการเรนเดอร์ฉากสามมิติจริง ๆ ขั้นตอนต่าง ๆ เหล่านี้สามารถแสดงเป็นแผนภูมิผังงานได้ดังรูปที่ 4.4



รูปที่ 4.4 แผนภูมิผังงานแสดงขั้นตอนการทดสอบโดยใช้ฉากที่จำลองขึ้น

ในการตรวจสอบความแม่นยำนั้นผู้วิจัยได้นำ ค่าเบี่ยงเบนรากกำลังสองเฉลี่ย (Root mean square deviation) มาใช้ในการทดสอบซึ่งค่านี้สามารถหาได้จากสมการที่ (4.6)

$$rmsd = \sqrt{\frac{\sum_{y=1}^n \sum_{x=1}^n (d(x,y) - ds(x,y))^2}{n}} \quad (4.6)$$

โดย  $ds(x,y)$  คือค่าความลึกที่ตำแหน่ง  $x,y$  ที่ได้จากการเรนเดอร์แบบจำลอง และ  $d(x,y)$  คือค่าความลึกที่ตำแหน่ง  $x,y$  ที่ได้จากขั้นตอนวิธีที่นำเสนอ นอกจากนี้เพื่อให้ตีความหมายได้ง่าย จึงแปลงค่าเบี่ยงเบนนี้ให้เป็นค่าบรรทัดฐานตามสมการที่ (4.7)

$$Normalized\ rmsd = \frac{rmsd}{\max(\nabla ds) - \min(\nabla ds)} \quad (4.7)$$

ค่าเบี่ยงเบนรากล้างสองเฉลี่ยที่เป็นค่าบรรทัดฐานนี้สามารถแสดงให้อยู่ในรูปร้อยละความคลาดเคลื่อนได้



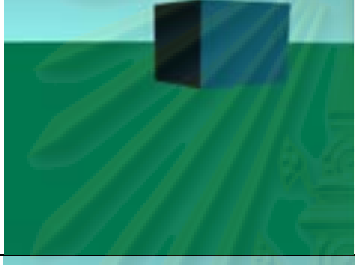

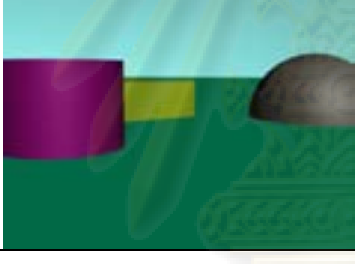

#### 4.3.2 ผลการทดสอบความแม่นยำในการประมาณพิกัดสามมิติ

ผู้วิจัยได้ทำการทดสอบกับรูปภาพสังเคราะห์ที่ได้จากแบบจำลองสามมิติจำนวน 44 รูป ตัวอย่างของรูปที่ใช้แสดงอยู่ในภาคผนวก ผลการทดสอบความคลาดเคลื่อนตามสมการที่ (4.7) ได้ร้อยละความคลาดเคลื่อนเฉลี่ย  $rmsd$  ร้อยละ 38.99 โดยมีร้อยละความคลาดเคลื่อนน้อยที่สุดอยู่ที่ร้อยละ 14.65 และมากที่สุดอยู่ที่ร้อยละ 195.31

รูปภาพที่นำมาทดสอบที่ขั้นตอนวิธีนี้สามารถประมาณค่าได้ใกล้เคียงที่สุดสามอันดับถูกแสดงในตารางที่ 4.3 และรูปภาพที่ให้ผลคลาดเคลื่อนจากความเป็นจริงไปมากที่สุดสามอันดับแรกแสดงใน ตารางที่ 4.4

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.3 แสดงผลการประมาณความลึกที่ผลลัพธ์มีความคลาดเคลื่อนน้อยที่สุด 3 อันดับแรก


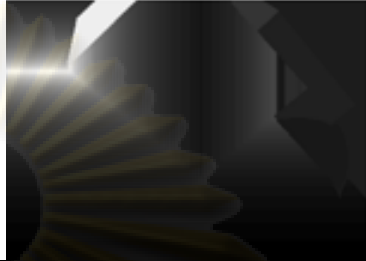

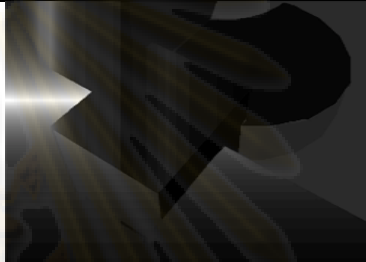


รหัสรูปภาพ	รูปภาพต้นแบบ	รูปภาพแสดงความลึกที่ได้จากขั้นตอนวิธีที่นำเสนอ	ร้อยละความคลาดเคลื่อน (rmsd)
s3-1			14.65
s2-2			15.59
s4-2			16.05

ในตารางที่ 4.3 จะแสดงรูปที่มีความคลาดเคลื่อนน้อยที่สุด โดยรูปที่ 1 ด้านบนสุดจะมีค่าความคลาดเคลื่อนร้อยละ 14.65 รูปตรงกลางจะมีความคลาดเคลื่อนร้อยละ 15.59 และรูปล่างสุดจะมีความคลาดเคลื่อนร้อยละ 16.05 ซึ่งจากรูปทั้งสามจะเห็นได้ว่าความแม่นยำนั้นไม่ได้ขึ้นอยู่กับจำนวนวัตถุในภาพแต่เป็นความชัดเจนของวัตถุแต่ละชิ้นในภาพ ทั้งนี้ความแม่นยำของค่าความลึกนั้นแปรผันโดยตรงกับความแม่นยำในการจำแนกประเภทของความสัมพันธ์

ศูนย์วิจัยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



ตารางที่ 4.4 แสดงผลการประมาณความลึกที่ผลลัพธ์มีความคลาดเคลื่อนมากที่สุด 3 อันดับแรก

รหัสรูปภาพ	รูปภาพต้นแบบ	รูปภาพแสดงความลึก	ร้อยละความคลาดเคลื่อน (rmsd)
s3-2			195.31
s3-3			89.00
s5-7			77.46

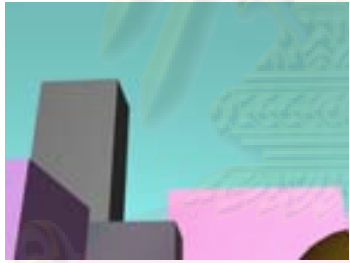


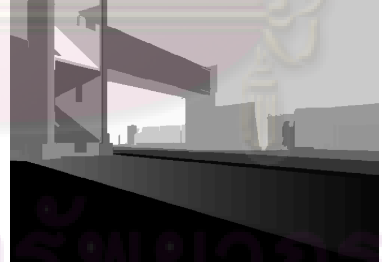
ในตารางที่ 4.4 ซึ่งแสดงรูปที่มีความคลาดเคลื่อนมากโดยในรูปด้านบนสุด (s3-2) นั้นจะมีความคลาดเคลื่อนมากที่สุดถึงร้อยละ 195.31 หรือเกือบสองเท่าของค่าที่ควรจะเป็น ซึ่งความผิดพลาดนี้เกิดจากภาพที่ใช้ในนั้นไม่เป็นไปตามสมมติฐาน ตัวอย่างเช่นในภาพด้านบนเป็นการถ่ายภาพโดยตั้งกล้องให้มีลักษณะเอียง ซึ่งขัดแย้งกับทำให้ค่าความลึกที่งานวิจัยนี้อ้างอิงจากพื้นที่มีลักษณะขนานกับแนวมุมมองของกล้อง หรือค่าความลึกตามสมการที่ (3.10)

นอกจากนี้สำหรับภาพตรงกลาง (s3-3) ที่มีซึ่งมีความคลาดเคลื่อนถึงร้อยละ 89 เป็นข้อผิดพลาดเกี่ยวกับการหาจุดรวมสายตาซึ่งเป็นปัญหาเดียวกันกับภาพด้านบนด้วย จากรูปแสดงค่าความลึกจะเห็นได้ว่ามีเส้นสีขาวที่แสดงความลึกของจุดรวมสายตาอยู่บนพื้น ซึ่งไม่ถูกต้อง และการที่ได้ค่าเช่นนี้ทำให้ค่าผิดไปจากความจริงมากเพราะค่าความลึกที่จุดรวมสายตาจะถือว่ามึระยะไกลมาก แต่ปัญหานี้สามารถแก้ไขได้โดยกำหนดจุดรวมสายตาที่ถูกต้อง ค่าความคลาดเคลื่อนก็จะลดลงเป็นอย่างมาก โดยเมื่อกำหนดความจุดรวมสายตาที่ -200 จะได้ค่าความคลาดเคลื่อนเพียงร้อยละ 43.27

สำหรับรูปด้านล่างสุด (s5-7) นั้น มีความคลาดเคลื่อนอยู่ที่ร้อยละ 77.46 ซึ่งเป็นสาเหตุมาจากการจำแนกประเภทความสัมพันธ์ที่ผิดพลาด โดยจำแนกว่าวัตถุที่ตั้งอยู่เป็นพื้นถึงสองวัตถุทำให้มีค่าความคลาดเคลื่อนไปมาก

นอกจากนี้ยังมีรูปภาพที่ให้ผลสำคัญ ๆ ดังแสดงในตารางที่ 4.5 รูปภาพ s3-4 นั้นมีค่าความผิดพลาดเพียงร้อยละ 23.09 แม้รูปนี้ไม่มีพื้นอ้างอิง ที่สามารถประมาณค่าได้ใกล้เคียง เนื่องจากการประมาณตำแหน่งโดยอ้างอิงว่าทุกวัตถุวางอยู่บนพื้นและวัตถุที่ปิดบังกันห่างกันโดยใช้การคำนวณจากตำแหน่งนั้น ๆ ซึ่งวิธีการอื่น ๆ นั้นไม่สามารถประมาณได้เนื่องจากไม่สามารถรู้ได้ว่าวัตถุเหล่านั้นห่างกัน นอกจากนี้ในรูปรหัส s7 นั้นแสดงให้เห็นว่าแม้ในรูปที่มีความซับซ้อนและสภาพสีที่แตกต่างออกไปวิธีการนี้ก็ยังสามารถประมาณความลึกได้แม่นยำในระดับหนึ่ง แสดงให้เห็นว่าขั้นตอนวิธีการที่งานวิจัยนี้ได้นำเสนอมีความสามารถในการประมาณตำแหน่งสามมิติได้กับรูปภาพที่ประกอบด้วยสภาพแวดล้อมหลายแบบ

ตารางที่ 4.5 แสดงผลลัพธ์การประมาณความลึกของภาพบางส่วนที่น่าสนใจ

รหัสรูปภาพ	รูปภาพต้นแบบ	รูปภาพแสดงความลึก	ร้อยละความคลาดเคลื่อน (rmsd)
s3-4			23.09
s7			33.98

งานวิจัยนี้ได้เปรียบเทียบความแม่นยำในการประมาณตำแหน่งสามมิติระหว่างวิธีที่งานวิจัยนี้แนะนำเสนอกับวิธีที่นำเสนอโดยเช็กซึน่า ทั้งนี้เพื่อเปรียบเทียบความแม่นยำของการประมาณค่าความลึกในจุดภาพ ภาพที่นำมาใช้ในการทดสอบคือภาพที่ได้จากการสังเคราะห์ชุดเดียวกับที่ได้นำมาทดสอบในข้างต้น โดยมีทั้งหมด 44 รูปภาพ ผลการเปรียบเทียบพบว่ามีความคลาดเคลื่อนเฉลี่ยร้อยละ 76.77 และมีความคลาดเคลื่อนสูงสุดอยู่ที่ร้อยละ 206.45 สำหรับความคลาดเคลื่อนน้อยที่สุดอยู่ที่ร้อยละ 36.48 ดังที่แสดงในตารางที่ 4.6 ซึ่งผลการเปรียบเทียบ


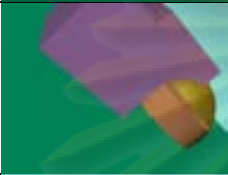




บางส่วนแสดงได้ตามตารางที่ 4.7 ทั้งนี้ค่าที่นำมาเปรียบเทียบนั้นเป็นค่าความลึกที่ได้จากการคำนวณโดยวิธี MRF ซึ่งวิธีนี้ทางเซ็กชีน่าได้ค่าความแปรปรวนสำหรับแต่ละจุดภาพด้วย แต่ในการทดสอบนี้ไม่ได้นำค่านี้มาใช้ในการคำนวณ

ตารางที่ 4.6 ตารางแสดงการเปรียบเทียบร้อยละความคลาดเคลื่อนในการประมาณค่าความลึกจากรูปภาพสังเคราะห์ 44 รูปภาพ

ขั้นตอนวิธี	ร้อยละความคลาดเคลื่อน		
	ต่ำสุด	สูงสุด	เฉลี่ย
วิธีของเซ็กชีน่า	31.12	206.45	76.77
วิธีของงานวิจัยนี้	14.65	195.31	38.99

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.7 ตารางแสดงการเปรียบเทียบร้อยละความคลาดเคลื่อนในการประมาณค่าความลึกบางส่วน

รหัสรูปที่ใช้	รูปตัวอย่าง	วิธีของแซ็กชีนา	วิธีของงานวิจัยนี้
s3-1		55.58	14.65
s3-2		68.41	195.31
s3-4		83.17	23.09
s5-7		48.07	77.46
s7		43.91	33.98
s8-3		144.92	31.40
s8-10		31.12	30.11

จากตารางที่ 4.6 จะเห็นว่าวิธีที่นำเสนอโดยงานวิจัยนี้มีความแม่นยำโดยเฉลี่ยมากกว่าร้อยละ 37.78 แสดงให้เห็นว่าวิธีที่ได้นำเสนอสามารถประมาณตำแหน่งในพิกัดสามมิติได้แม่นยำกว่าวิธีที่ผ่านมาในระดับหนึ่ง

ตารางที่ 4.7 สำหรับตารางที่ 4.7 จะเห็นได้ว่าวิธีที่นำเสนอสามารถประเมินรูป s3-1 และ s3-4 ได้แม่นยำกว่าวิธีการของแซ็กซีนา สำหรับรูปรหัส s3-1 สาเหตุที่วิธีของงานวิจัยนี้สามารถประมาณได้แม่นยำกว่าสืบเนื่องมาจากการประมาณค่าตำแหน่งโดยใช้ลักษณะของค่าความลึกของวัตถุแต่ละชั้น ในเมื่อสามารถระบุประเภทของความสัมพันธ์ได้ใกล้เคียงกับความเป็นจริง วิธีของงานวิจัยนี้จึงสามารถประมาณความลึกได้อย่างแม่นยำ ในขณะที่วิธีของแซ็กซีนา ใช้การหาค่าความลึกที่ละจุดพิกเซลด้วยการใช้ค่าจุดเด่นจากลอว์มาคส์ จึงทำให้ค่าความลึกมีความแปรปรวนมาก สำหรับ s3-4 นั้นถือเป็นจุดเด่นของงานวิจัยนี้ที่สามารถประมาณความลึกได้แม้จะไม่มีพื้นอ้างอิง ด้วยเหตุนี้ภาพที่มีลักษณะเช่นนี้งานวิจัยนี้จึงสามารถทำงานได้ดีกว่า แต่พอมาเปรียบเทียบในรูป s3-2 ซึ่งงานวิจัยนี้ทำได้ไม่ดีนั้น พบว่าวิธีของแซ็กซีนาสามารถประมาณได้แม่นยำกว่า แม้ว่าจะมีจะมีพื้นส่วนด้านบนที่วิธีใช้วิธีการประมาณแบบท้องฟ้า แต่วิธีของแซ็กซีนาได้ข้อดีของการประเมินความลึกที่ละจุดภาพมาช่วยลดการผิดพลาดส่วนใหญ่ลงไป

สำหรับผลการทดลองในหัวข้อนี้สามารถสรุปได้ว่าวิธีการที่นำเสนอสามารถประมาณค่าพิกัดสามมิติได้อย่างแม่นยำโดยมีความคลาดเคลื่อนเฉลี่ยอยู่ที่ร้อยละ 38.99 โดยสามารถลดความคลาดเคลื่อนในการประมาณสามมิติจากวิธีที่เสนอโดยแซ็กซีนาเฉลี่ยร้อยละ 37.78 ทั้งนี้รูปที่สามารถประมาณตำแหน่งได้แม่นยำคือรูปเป็นไปตามข้อสมมติฐานของงานวิจัยนี้ นั่นคือทุกวัตถุในภาพนั้นเป็นส่วนต่อชิ้นมาจากพื้นกล่าวคือทุกวัตถุในภาพไม่ได้ลอยอยู่ และพื้นดังกล่าวมีแนวระนาบขนานไปกับแนวสายตาของกล้อง

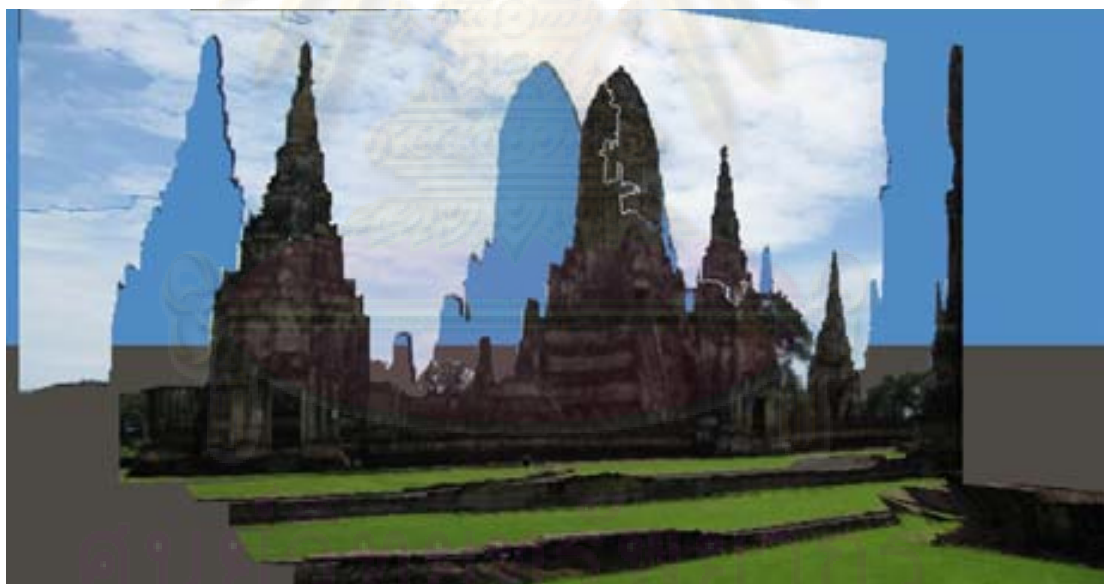
#### 4.4 วิเคราะห์รูปร่างของแบบจำลองสามมิติ

ในด้านของลักษณะรูปร่างแบบจำลอง ความแตกต่างระหว่างรูปแบบของฉากจำลองที่ได้จากวิธีการของโฮเอ็ม แซ็กซีนาและงานวิจัยนี้สามารถแสดงได้ในตารางที่ 4.8 โดยงานวิจัยนี้เมื่อเทียบกับของโฮเอ็ม จะสามารถสร้างแบบจำลองฉากสามมิติมีความซับซ้อนมากกว่า คือสามารถสร้างวัตถุได้ในหลายระดับของแนวระนาบและสามารถใช้ความสัมพันธ์ในการเติมส่วนที่ขาดหายไปได้ และถ้าเทียบกับแซ็กซีนาจะเห็นว่ามีความซับซ้อนน้อยกว่าเพราะการประมาณฉากสามมิติของแซ็กซีนาจะสามารถทำได้ในระนาบใด ๆ ทั้งนี้การสร้างส่วนที่มองไม่เห็นจะเป็นสร้างโดยเชื่อมต่อทั้งหมด น้อยจากนี้ถ้าวิธีการสร้างแบบจำลองสามมิติด้วยวิธีการวิเคราะห์ความสัมพันธ์นี้ จะสามารถสร้างฉากจำลองที่ใช้พื้นฐานเป็นวัตถุ (Object based) ซึ่งจะสามารถมองเห็นวัตถุเป็นชิ้น ๆ ลอยออกมาได้ ตัวอย่างแบบจำลองสามมิติที่ได้ถูกแสดงในรูปที่ 4.5



ตารางที่ 4.8 สรุปลักษณะของแบบจำลองสามมิติเปรียบเทียบกับวิธีของโฮเอ็มและแซ็กซีนา

	โฮเอ็ม	แซ็กซีนา	งานวิจัยนี้
ลักษณะแบบจำลอง	3 ประเภท ท้องฟ้า พื้น วัตถุตั้ง แยกเป็นแค่สามประเภท	ระนาบเล็ก ๆ จำนวนมาก วางตัวแบบอิสระ	ตามความสัมพันธ์ ทั้ง 6 แบบ แบบจำลองแยกชิ้นส่วนตามรูปร่างวัตถุ
การเชื่อมต่อ	แยกกันทุกวัตถุ ในลักษณะด้นขึ้นมาจากพื้น	ทุกระนาบต่อกันเป็นแบบจำลองเดี่ยวไม่มีช่องว่าง	พิจารณาการเชื่อมต่อด้วยความสัมพันธ์
มองเห็นส่วนที่ขาด	เห็น	ไม่เห็น เนื่องจากทุกส่วนเชื่อมกันหมด	เห็น แต่น้อยกว่าโฮเอ็ม เนื่องจากสามารถนำความสัมพันธ์มาประมาณได้
ขนาดข้อมูลแบบจำลอง	เล็ก ใช้เพียงสองระนาบคือ พื้น และวัตถุที่ตั้งขึ้น ท้องฟ้าถูกตัดออก	ใหญ่ เนื่องจากจำนวนระนาบบนแบบจำลองที่แทนที่ทุก ๆ ส่วนย่อย	ใหญ่ เนื่องจากจำนวนจุดยอดที่กำหนดตามจุดวิกฤต



รูปที่ 4.5 แสดงผลการสร้างแบบจำลองจากสามมิติ

จุฬาลงกรณ์มหาวิทยาลัย

## บทที่ 5

### สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

งานวิจัยนี้ได้นำเสนอวิธีการสร้างแบบจำลองสามมิติโดยอัตโนมัติ ซึ่งใช้ความสัมพันธ์ระหว่างวัตถุมาเป็นตัวตัดสินลักษณะและตำแหน่งของแบบจำลองสามมิติ ทั้งนี้ยังได้เสนอขั้นตอนวิธีในการวิเคราะห์จำแนกประเภทของความสัมพันธ์และการนำความสัมพันธ์นี้ไปใช้ในการสร้างแบบจำลองสามมิติ โดยวิธีที่นำเสนอขึ้นจัดอยู่ในการสร้างแบบจำลองจากภาพ โดยมีหลักพื้นฐานอยู่บนการวิเคราะห์ความสัมพันธ์ทางเรขาคณิต เช่นเดียวกับงานวิจัยของโฮเอ็่มและแซ็กซีนา แต่แตกต่างกันตรงที่งานวิจัยของโฮเอ็่มเป็นระบุประเภทส่วนย่อยของรูปภาพเป็นพื้นดิน ท้องฟ้า หรือวัตถุตั้งตรง เพื่อใช้ความสัมพันธ์ระหว่างพื้นดิน ท้องฟ้าและวัตถุตั้งตรงในการระบุตำแหน่ง สำหรับงานวิจัยของแซ็กซีนาใช้แบบจำลองความน่าจะเป็นโดยประมาณจากค่าสีและตำแหน่ง ในการระบุพิกัดสามมิติ ในขณะที่งานวิจัยนี้ได้เสนอให้การหาความสัมพันธ์ระหว่างวัตถุในภาพ และใช้ความสัมพันธ์นี้ในการประมาณพิกัดสามมิติและสร้างแบบจำลองสามมิติขึ้นมาตามความสัมพันธ์นั้น ๆ ซึ่งวิธีที่ได้นำเสนอประสบความสำเร็จโดยสามารถสรุปได้จากการทดลองดังนี้

จากการทดลองการกระจายตัวของค่าจุดเด่นตามที่ได้รายงานในหัวข้อ 4.1 สามารถสรุปได้ว่าถ้าซูปเปอร์พิกเซล  $S_i$  และ  $S_j$  มีความสัมพันธ์  $R_{ij}$  ที่ต่างกัน ค่าจุดเด่นที่ได้ก็จะมีลักษณะแตกต่างกัน โดยคลาสที่เด่นชัดที่สุดคือคลาส  $R_{ij} = PT$  รองลงมาคือ  $R_{ij} = SO$  ซึ่งในการทดลองความแม่นยำในการจำแนกประเภทซึ่งได้รายงานในหัวข้อที่ 4.2 สามารถช่วยยืนยันได้ว่าสามารถระบุความสัมพันธ์  $R_{ij}$  ระหว่างซูปเปอร์พิกเซล  $S_i$  และ  $S_j$  เป็นหกคลาสได้โดยการใช้ค่าจุดเด่นดังกล่าว โดยมีความแม่นยำเฉลี่ยอยู่ที่ร้อยละ 65.74 นอกจากนี้ในหัวข้อที่ 4.3 ได้ทำการทดสอบความแม่นยำในการประมาณพิกัดสามมิติ โดยวิธีที่งานวิจัยนี้ได้นำเสนอสามารถประมาณตำแหน่งได้ดีและมีค่าความคลาดเคลื่อนเฉลี่ยร้อยละ 38.99 ซึ่งเมื่อเปรียบเทียบกับวิธีจากงานวิจัยของแซ็กซีนาสามารถลดความคลาดเคลื่อนลงไปได้โดยเฉลี่ยร้อยละ 37.78

นอกจากนี้จากผลการทดลองจากหัวข้อที่ 4.1 4.2 และ 4.3 สามารถสรุปลักษณะรูปภาพนำเข้าที่วิธีการนี้สามารถทำงานได้ดี รูปที่เหมาะสมจะมีลักษณะสำคัญสองประการดังนี้

1. รูปภาพที่ใช้เป็นไปตามข้อสมมติฐานของงานวิจัยนี้ นั่นคือทุกวัตถุในภาพนั้นเป็นส่วนต่อขึ้นมาจากพื้น กล่าวคือทุกวัตถุในภาพไม่ได้ลอยอยู่ และพื้นดังกล่าวมีแนวระนาบขนานไปกับแนวสายตาของกล้อง
2. รูปภาพที่สามารถวิเคราะห์ความสัมพันธ์ได้ง่าย โดยรูปภาพที่วิเคราะห์ได้ง่ายนั้นคือรูปภาพที่มีลักษณะส่วนย่อยที่ชัดเจนทั้งเส้นขอบและสภาพสี และภาพที่มีการวางตัวเรียงตัวตามลักษณะที่ว่ายังอยู่ข้างล่างยังอยู่ใกล้

สำหรับลักษณะของแบบจำลองที่ได้จากนำความสัมพันธ์ระหว่างวัตถุนี้ไปใช้ จะมีลักษณะที่แตกต่างจากงานวิจัยของโฮเอ็่มและงานวิจัยของแซ็กซีนา โดยเมื่อเทียบกับงานของโฮเอ็่มจะมีสามารถสร้างฉากที่ซับซ้อนได้มากกว่าเนื่องจากไม่ได้จำกัดการวางตัวของวัตถุเป็นแค่สามแบบดังที่โฮเอ็่มได้กำหนดไว้ได้แก่ พื้นดิน ท้องฟ้า และวัตถุตั้งตรง แต่จะไม่สามารถสร้างฉากที่ซับซ้อนได้เท่ากับของแซ็กซีนาที่สามารถกำหนดการวางตัวได้อิสระตามระนาบ

## 5.1 ข้อเสนอแนะ

เนื่องจากงานวิจัยนี้เป็นการนำเสนอแนวทางใหม่ซึ่งเทคโนโลยีที่นำมาใช้ส่วนใหญ่จึงยังไม่ใช้เทคโนโลยีที่เหมาะสมที่สุด ดังเช่น ขั้นตอนวิธีในการจำแนกประเภท จะเห็นได้ว่ายังใช้ค่าจุดเด่นมาพิจารณาได้ไม่เต็มที่ ซึ่งอาจจะพัฒนาเป็นการใช้เงื่อนไขต่อเนื่อง เช่นถ้าค่าจุดเด่น  $A$  เป็น คลาส  $k$  แล้วค่าจุดเด่น  $B$  จะมีค่าเป็นคลาส  $k$  ไม่ได้เป็นต้น นอกจากนี้การนำค่าความน่าจะเป็นมาช่วยถ่วงน้ำหนักในการเลือกคลาสนี้ก็จะช่วยเพิ่มความแม่นยำได้มากขึ้น สำหรับส่วนที่เพิ่มขึ้นมากจากการใช้ความสัมพันธ์ประมาณนั้นลายผิวยังเป็นลายที่ใช้จากบริเวณขอบของวัตถุ ซึ่งสามารถทำให้มีความสวยงามสมจริงมากขึ้นได้ด้วยการนำเอาหลักการการสังเคราะห์ลายผิวมาช่วยประมาณส่วนดังกล่าวได้ นอกจากนี้ตามที่ได้กล่าวไปในบทนำ คืองานวิจัยนี้ใช้รูปภาพนำเข้าเพียงแค่รูปเดี่ยวดังนั้นความแม่นยำและส่วนที่ขาดหายไปอาจเพิ่มเติมได้จากกรุปจากมุมมองอื่น ๆ ทั้งนี้จำเป็นต้องสร้างส่วนเชื่อมต่อระหว่างตำแหน่งของวัตถุในแต่ละภาพ

งานนี้สามารถนำไปประยุกต์ใช้กับงานอื่น ๆ ได้หลายทางยกตัวอย่างเช่น นำวิเคราะห์สิ่งที่หุ่นยนต์เห็น (Robot vision) ซึ่งสามารถบอกว่าสิ่งใดอยู่ด้านหน้าด้านหลัง และการวางตัวอย่างคร่าว ๆ ได้ แต่ทั้งนี้การจะนำไปประยุกต์ใช้กับงานทางด้านนี้จำเป็นต้องพัฒนาเรื่องความเร็วในการประมวลผลซึ่งอาจสามารถแลกได้กับความแม่นยำ นอกจากนี้งานวิจัยนี้ยังสามารถใช้เป็นเครื่องมือในการขยายภาพได้ แต่จะเป็นการขยายภาพแบบทัศนมิติ คือผู้ใช้จะรู้สึกเหมือนตัวเองเดินเข้าไปมากกว่าที่ภาพใหญ่ขึ้น ซึ่งถ้าสามารถสร้างฉากสามมิติได้อัตโนมัติอยู่แล้วก็เพียงแค่นำตำแหน่งของกล้องเข้าหาแบบจำลองของวัตถุเท่านั้นเอง งานวิจัยนี้ยังสามารถนำไปช่วยสร้างโลกเสมือน ซึ่งใช้ข้อมูลนำเข้าเป็นรูปภาพจากอินเทอร์เน็ตซึ่งมีจำนวนมหาศาลได้ ซึ่งสิ่งที่ต้องเพิ่มเติมคือการกำหนดส่วนเชื่อมต่อระหว่างรูปภาพ ซึ่งยังคงเป็นปัญหาที่ทำหาย

จุฬาลงกรณ์มหาวิทยาลัย

## รายการอ้างอิง

- [1] Youichi, H., Ken-Ichi, A., and Kiyoshi, A. (1997). Tour into the picture: using a spidery mesh interface to make animation from a single image. Proceeding of 24th annual conference on Computer graphics and interactive techniques.
- [2] Saxena, A., CHung, S.H., and Ng, A.Y. (2005). Learning Depth from Single Monocular Image. Proceeding of Neural Information Preccessing Systems Conference 2005.
- [3] Saxena, A., Sun, M., and Ng, A.Y. (2007). Learning 3-D Scene Structure from a Single Still Image. Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on, pp. 1-8.
- [4] Prados, E., and Faugeras, O. (2005). Shape from shading: a well-posed problem? Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, pp. 870-877 vol. 872.
- [5] Derek, H., Alexei, A.E., and Martial, H. (2005). Automatic photo pop-up. ACM SIGGRAPH 2005 Papers, Los Angeles, California.
- [6] Shimshoni, I., Moses, Y., and Lindenbaumpr, M. (1999). Shape reconstruction of 3D bilaterally symmetric surfaces. Proceedings of International Conference on Image Analysis and Processing, 1999., pp. 76-81.
- [7] Levin, A., Fergus, R., Durand, F., and Freeman, W.T. (2007). Image and Depth from a Conventional Camera with a Coded Aperture. ACM Transactions on Graphics.
- [8] Tam, G.K.L., and Lau, R.W.H. (2007). Deformable Model Retrieval Based on Topological and Geometric Signatures. Transactions on Visualization and Computer Graphics 13: pp. 470-482.
- [9] Hoiem, D., Efros, A.A., and Hebert, M. (2006). Putting Objects in Perspective. Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, pp. 2137-2144.
- [10] Oliveira, M.M. (2002). Image-Based Modeling and Rendering Techniques: A Survey. RITA IX: pp. 37-66.
- [11] Nahmias, D. (2000). Image Based Modeling Technique <http://www.cs.ucl.ac.uk/staff/D.Nahmias/log.html>. [2007, November]

- [12] Kutulakos, K.N., and Seitz, S.M. (1999). A theory of shape by space carving. Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, pp. 307-314 vol.301.
- [13] Yi, X., and Daniel, G.A. (2007). Robust pixel classification for 3D modeling with structured light. Proceedings of Graphics Interface 2007, Montreal, Canada.
- [14] Takeo, I., Satoshi, M., and Hidehiko, T. (2007). Teddy: a sketching interface for 3D freeform design. ACM SIGGRAPH 2007 courses, San Diego, California.
- [15] Basri, R., Basri, R., and Moses, Y. (1998). When is it possible to identify 3D objects from single images using class constraints? Computer Vision, 1998. Sixth International Conference on, pp. 541-548.
- [16] G.Lowe, D. (2004). Distinctive Image features form Scale-Invariant Keypoints. International Journal of Computer Vision
- [17] Funkhouser, T., Min, P., Kazhdan, M., Chen, J., Halderman, A., and Dobkin, D. (2003). A Search Engine for 3D Model. ACM Transactions on Graphics 22: pp. 83-105.
- [18] Volker, B., and Thomas, V. (1999). A morphable model for the synthesis of 3D faces. Proceedings of the 26th annual conference on Computer graphics and interactive techniques.
- [19] Criminisi, A., Reid, I., and Zisserman, A. (1999). Single view metrology. International Journal of Computer Vision 1: pp. 434-441 vol.431.
- [20] Shapiro, L.G., and Stockman, G.C. (2001). Computer Vision. New Jersey: Prentice-Hall. pp 279-325.
- [21] Christoudias, C.M., Georgescu, B., and Merr, P. (2002). Synergism in low-level vision. 16th International Conference on Pattern Recognition, pp. 150-155. Quebec City, Canada.
- [22] Fernand, M. (1994). Topographic distance and watershed lines. International Journal of Signal Process. 38: pp. 113-125.
- [23] Pedro, F.F., and Daniel, P.H. (2004). Efficient Graph-Based Image Segmentation. International Journal of Computer Vision 59: pp. 167-181.
- [24] Jain, R., Kasturi, R., and SChunck, B.G. (1995). Machine Vision. Singapore: McGraw-Hill Book Co.
- [25] Schneider, P.J., and Eberly, D.H. (2003). Geometric tools for computer graphics. Morgan Kaufmann Publisher.



- [26] Davies, E.R. (1997). Laws' testure energy in TEXTURE. Laws' testure energy in TEXTURE. ed. San Diego: Academi Press.
- [27] Hoiem, D., Stein, A.N., Efros, A.A., and Hebert, M. (2007). Recovering Occlusion Boundaries from a Single Image. Proceeding of Eleventh IEEE International Conference on Computer Vision.
- [28] Hartman, J., and Wernecke, J. (1996). The VRML 2.0 Handbook. Addison-Wesley Publishing Company.



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

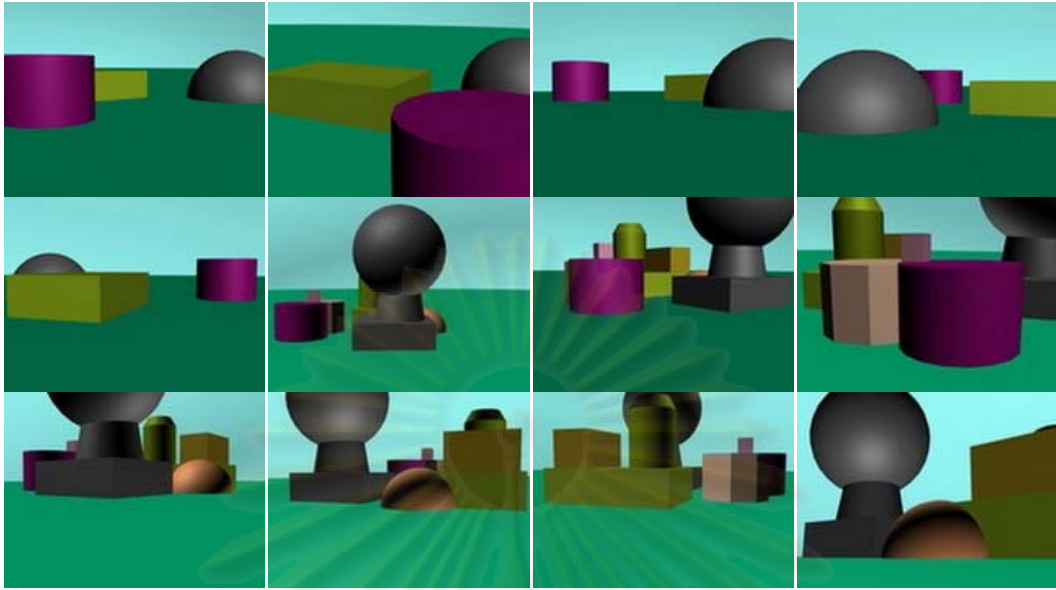


ภาคผนวก

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่เป็นภาพสังเคราะห์จากโปรแกรมสามมิติมีจำนวนทั้งสิ้น 44 รูป รูปเหล่านี้  
 ได้จากการสร้างแบบจำลองสามมิติแล้วจึงนำไปผ่านกระบวนการให้แสงและเงา โดยแบบจำลอง  
 นี้ถูกสร้างโดยโปรแกรม Autodesk 3ds Max 2009 แบบจำลองที่นำมาใช้มีทั้งที่สร้างขึ้นด้วย  
 ผู้วิจัยเองและที่มาจากอินเทอร์เน็ต





รูปที่เป็นรูปถ่ายมีจำนวนทั้งสิ้น 46 รูปภาพ เป็นรูปถ่ายทิวทัศน์ทั่วไป ทั้งที่ได้มาจากอินเทอร์เน็ต และรูปที่ผู้วิจัยถ่ายด้วยตัวเอง







ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

## ประวัติผู้เขียนวิทยานิพนธ์

นายณัฏพล ฟุตระกูล เกิดวันที่ 18 มีนาคม พ.ศ. 2528 สำเร็จการศึกษาระดับมัธยมศึกษาที่โรงเรียนมงฟอร์ตวิทยาลัย จังหวัดเชียงใหม่ จากนั้นจึงเข้าศึกษาต่อที่คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษาที่ 2546 และสำเร็จการศึกษาในระดับปริญญาตรีภาควิชาคอมพิวเตอร์ในปีการศึกษา 2549 และเข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิตต่อในปีการศึกษา 2550 ในภาควิชาคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย