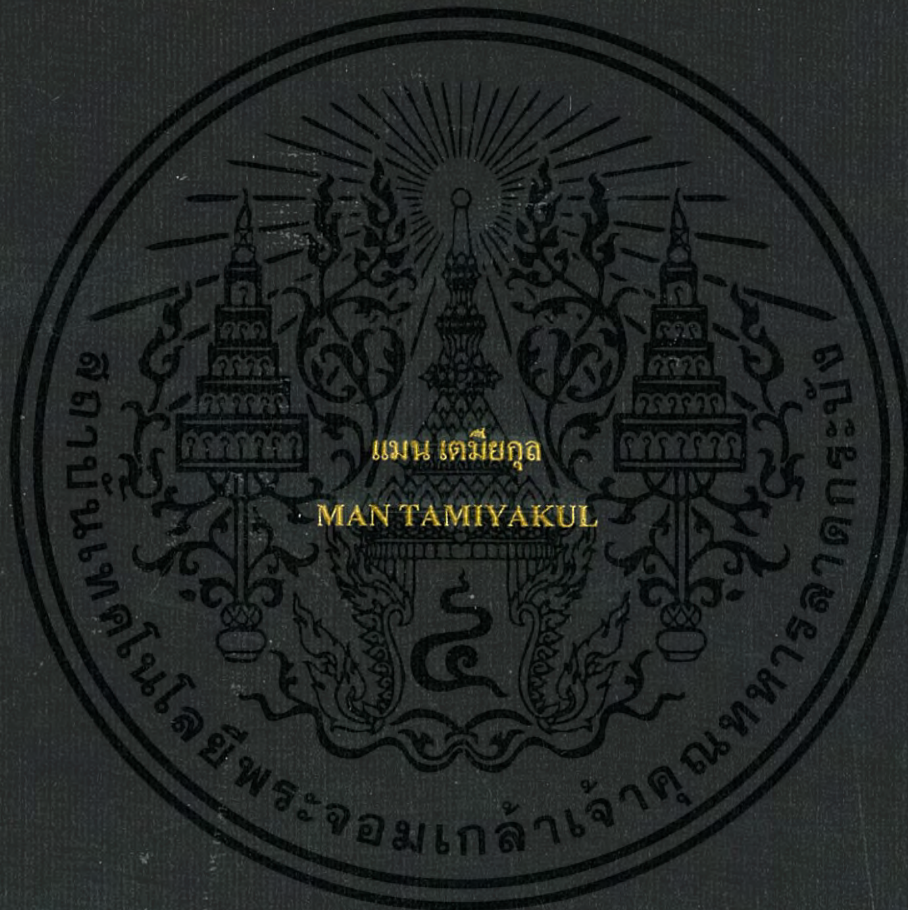


การสร้างภาพไตรแมบแบบอัตโนมัติด้วยข้อมูลภาพสีและ  
ข้อมูลภาพระยะลึกร่วมกัน

AUTOMATIC TRIMAP IMAGE GENERATION  
USING COLOR AND DEPTH IMAGES



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีระบบสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2559

KMITL-2016-IT-M-001-003

การสร้างภาพไตรแมบแบบอัตโนมัติด้วยข้อมูลภาพสีและ  
ข้อมูลภาพระยะลึกร่วมกัน

AUTOMATIC TRIMAP IMAGE GENERATION  
USING COLOR AND DEPTH IMAGES



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีระบบสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2559

KMITL-2016-IT-M-001-003

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**AUTOMATIC TRIMAP IMAGE GENERATION  
USING COLOR AND DEPTH IMAGES**



**A THESIS SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENT FOR THE DEGREE OF  
MASTER OF SCIENCE IN INFORMATION SYSTEM TECHNOLOGY  
FACULTY OF INFORMATION TECHNOLOGY  
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

**2016**

**KMITL-2016-IT-M-001-003**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



**COPYRIGHT 2016**

**FACULTY OF INFORMATION TECHNOLOGY**

**KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์ของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คณะเทคโนโลยีสารสนเทศ  
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง  
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ การสร้างภาพไตรแมปแบบอัตโนมัติด้วยข้อมูลภาพสี และข้อมูลภาพระยะลึกร่วมกัน  
Automatic trimap image generation using color and depth images

นักศึกษา นายแมน เตมียกุล

รหัสประจำตัว 54660413

ปริญญา วิทยาศาสตรมหาบัณฑิต

สาขาวิชา เทคโนโลยีสารสนเทศ

อาจารย์ที่ปรึกษาวิทยานิพนธ์ รองศาสตราจารย์ ดร.นพพร โชติกกำจร

คณะกรรมการสอบวิทยานิพนธ์	ลายมือชื่อ
ผู้ช่วยศาสตราจารย์ ดร.กัณฑ์พงษ์ วรรณปัญญา	
รองศาสตราจารย์ ดร.ชม กิมปาน	
รองศาสตราจารย์ ดร.นพพร โชติกกำจร	
ดร.สุภวรรณ อันนันทน์	

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง  
KING MONKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

วัน / เดือน / ปี ที่สอบ วันจันทร์ที่ 13 มิถุนายน 2559 เวลา 10.00 น. เป็นต้นไป

สถานที่สอบ ณ ห้อง M04 คณะเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศรับรองแล้ว



(รองศาสตราจารย์ ดร.นพพร โชติกกำจร)

คณบดีคณะเทคโนโลยีสารสนเทศ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษา วันที่ 21 เดือน พฤษภาคม พ.ศ. 2559

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การสร้างภาพไตรแมมแบบอัตโนมัติจากข้อมูลภาพสีและข้อมูลภาพระยะลึกร่วมกัน
นักศึกษา	นายแมน เตมียกุล
รหัสนักศึกษา	54660413
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
แขนงวิชา	เทคโนโลยีระบบสารสนเทศ
พ.ศ.	2559
อาจารย์ที่ปรึกษา	รศ. ดร. นพพร โชติกกำจร

### บทคัดย่อ

รายงานวิทยานิพนธ์ฉบับนี้นำเสนอเทคนิคการสร้างภาพไตรแมมแบบอัตโนมัติจากข้อมูลภาพสีและข้อมูลภาพระยะลึก เพื่อนำมาใช้แทนการสร้างภาพไตรแมมด้วยมือที่ต้องใช้เวลาและความชำนาญ ภาพไตรแมมเป็นข้อมูลภาพตั้งต้นที่ใช้ในการสร้างภาพแมท ซึ่งใช้ในการแยกภาพวัตถุกับภาพพื้นหลังเพื่อนำไปซ้อนทับกับภาพพื้นหลังใหม่ วิธีหนึ่งในการสร้างภาพไตรแมม คือ การสร้างภาพไตรแมมจากข้อมูลภาพระยะลึกที่อาศัยความแตกต่างของระยะลึกระหว่างวัตถุกับพื้นหลังเป็นตัวกำหนด แต่การสร้างภาพไตรแมมจากข้อมูลภาพระยะลึกนั้นพบปัญหาความไม่ถูกต้องของข้อมูลภาพระยะลึก ซึ่งส่งผลกระทบต่อประสิทธิภาพในการสร้างภาพไตรแมม และเพื่อแก้ปัญหาดังกล่าว วิทยานิพนธ์ฉบับนี้เสนอวิธีการสร้างภาพไตรแมมจากข้อมูลภาพสีและข้อมูลภาพระยะลึกแบบอัตโนมัติ โดยกระบวนการทำงานแบ่งออกได้เป็น 3 ขั้นตอนคือ 1) การปรับปรุงข้อมูลภาพระยะลึก, 2) การสร้างเส้นขอบวัตถุจากข้อมูลภาพสีและข้อมูลภาพระยะลึก และ 3) การกำหนดบริเวณรอยต่อระหว่างวัตถุเบื้องหน้าและพื้นหลัง โดยในการทดลองกับภาพวัตถุที่สีแตกต่างกับพื้นหลัง การใช้ข้อมูลภาพสีสามารถสร้างเส้นขอบที่มีความใกล้เคียงกับขอบวัตถุมากกว่าการใช้ข้อมูลภาพระยะลึกสำหรับกรณีภาพวัตถุมีสีคล้ายคลึงกับพื้นหลัง ผลการทดลองพบว่า วิธีการที่นำเสนอสามารถสร้างเส้นขอบวัตถุได้ถูกต้องมากกว่าการใช้ข้อมูลภาพสีเพียงอย่างเดียวในการสร้างภาพไตรแมม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<b>Thesis</b>	Automatic trimap image generation using color and depth images
<b>Student</b>	Mr. Man Tamiyakul
<b>Student ID.</b>	54660413
<b>Degree</b>	Master of Science
<b>Program</b>	Information Technology
<b>Major</b>	Information System Technology
<b>Year</b>	2016
<b>Thesis Advisor</b>	Assoc. Prof. Dr. Nopporn Chotikakamthorn

## ABSTRACT

This thesis proposes a technique for automatic trimap image generation using color and depth images. The technique is used to replace a manual process, which takes time and requires expertise. A trimap image is an initial data used for generating a matte image. A matte image is needed for separating a foreground object from a background image, and adding a new background image. One method is to generate a trimap image from a depth image information by using difference of depth between a foreground object and background. However, measurement error of depth information affects a trimap image generation. To solve the problem, this thesis proposes automatic trimap image generation using color and depth images. The process can be divided into three steps: 1) Depth image enhancement, 2) Object boundary detection from color and depth images information and 3) Determination of object boundary region for generating a trimap image. From the experiment where a foreground object has different color to the background, the result shows improvement over a method using depth information alone. If a foreground object image has similar color to the background, the proposed method can generate more accurate object edge than a method using only color information for generating a trimap image.

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จได้ด้วยดีเพราะความกรุณาจากอาจารย์ที่ปรึกษา  
รองศาสตราจารย์ ดร.นพพร โชติกกำจร ที่ให้ความช่วยเหลือ ให้คำชี้แนะช่วยแก้ปัญหาตลอดจนให้  
ความรู้และประสบการณ์ที่ดีแก่ข้าพเจ้า

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.กนต์พงษ์ วรรณปัญญา, ดร. สุภวรรณ อันนันทน์  
และรองศาสตราจารย์ ดร.ชม กิมปาน คณะกรรมการสอบวิทยานิพนธ์ที่ได้ให้คำแนะนำที่เป็น  
ประโยชน์อย่างยิ่งต่อการทำวิทยานิพนธ์ฉบับนี้

สุดท้ายนี้ขอขอบพระคุณกำลังใจจากคุณพ่อ คุณแม่ และเพื่อนๆ ที่ให้กำลังใจต่อข้าพเจ้าให้มี  
ความมุ่งมั่นที่จะในการทำวิทยานิพนธ์ฉบับนี้ให้สำเร็จลุล่วงไปได้ด้วยดี



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญ

	หน้า
บทคัดย่อ .....	I
ABSTRACT .....	II
กิตติกรรมประกาศ .....	III
สารบัญ .....	IV
สารบัญรูป .....	VII
สารบัญตาราง .....	X
บทที่ 1 บทนำ .....	1
1.1 ความเป็นมาและความสำคัญของปัญหา .....	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา .....	2
1.3 สมมติฐานของการวิจัย .....	2
1.4 นิยามศัพท์ .....	2
1.5 ทฤษฎีและแนวคิดที่ใช้ในการวิจัย .....	3
1.6 ขอบเขตการวิจัย .....	4
1.7 ขั้นตอนการวิจัย .....	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง .....	5
2.1 ทฤษฎีการซ้อนภาพ (Matting) .....	5
2.1.1 หลักการทั่วไปของการซ้อนภาพพื้นหลัง .....	6
2.1.2 เทคนิคการซ้อนภาพฉากหลังแบบดิจิทัล .....	8
2.1.3 วิธีการแยกภาพฉากหลังธรรมชาติ .....	10
2.2 กล้องถ่ายภาพแบบ RGB-D .....	12
2.2.1 การปรับตั้งค่าอุปกรณ์ Kinect .....	13
2.2.2 การคำนวณค่าความคลาดเคลื่อนของตัวอุปกรณ์ .....	14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญ (ต่อ)

	หน้า
2.3 การคัดแยกภาพวัตถุด้วยข้อมูลภาพระยะลึก.....	15
2.4. การคำนวณภาพแมทจากข้อมูลภาพระยะลึก.....	16
2.4.1 การใช้ภาพระยะลึกกับเทคนิค Bayesian-based matting .....	16
2.4.2 การใช้ภาพระยะลึกกับเทคนิค Closed-form matting.....	21
2.5 การสร้างแผนภาพไทรแมบโดยอัตโนมัติจากข้อมูลระยะลึก.....	24
บทที่ 3 การสร้างภาพไทรแมบบนแบบอัตโนมัติจากข้อมูลภาพสีและข้อมูลภาพระยะลึกร่วมกัน .....	25
3.1 การสร้างภาพไทรแมบบนแบบอัตโนมัติจากข้อมูลภาพ RGBD .....	25
3.1.1 การปรับปรุงภาพระยะลึก (Depth image enhancement).....	25
3.1.2 การตรวจสอบเส้นขอบภาพระยะลึก (Depth edge detection).....	27
3.1.3 การตรวจสอบเส้นขอบภาพสี (Color edge detection).....	27
3.1.4 การลบเส้นขอบลวดลาย (Texture edge removal) .....	28
3.1.5 การตรวจสอบขอบวัตถุ (Object boundary detection) .....	28
3.1.6 การสร้างภาพไทรแมบ (Trimap generation).....	29
3.2 การปรับปรุงความไม่สมบูรณ์ของข้อมูลภาพระยะลึก .....	30
3.3 การตรวจสอบหาขอบวัตถุจากข้อมูลระยะลึก .....	36
3.4 การตรวจสอบหาขอบวัตถุจากเส้นขอบสี.....	36
3.5 การคัดกรองเส้นขอบสีที่เกิดจากลวดลาย.....	36
3.5.1 การแบ่งเส้นขอบสีออกเป็นช่วง.....	40
3.5.2 การตรวจสอบค่าสีรอบเส้นขอบ .....	41
3.5.3 การตรวจสอบคัดแยกค่าสีโดยรอบเส้นขอบ .....	42
3.5.4 การคัดกรองเส้นขอบสีที่เกิดจากลวดลาย .....	43
3.5.5 การคัดกรองความยาวของเส้นขอบสี.....	44

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญ (ต่อ)

	หน้า
3.6 การสร้างเส้นขอบวัตถุ.....	44
3.7 การสร้างภาพไครแมบจากเส้นขอบวัตถุ.....	48
บทที่ 4 ผลการทดลอง.....	50
4.1 วัตถุประสงค์ของการทดลอง.....	50
4.2 อุปกรณ์ที่ใช้ในการทดลอง.....	51
4.3 รายละเอียดและผลการทดลอง.....	51
4.3.1 การทดลองเปรียบเทียบความถูกต้องของการตรวจจับเส้นขอบวัตถุ.....	51
4.3.2 การทดสอบประสิทธิภาพในการกำหนดขอบเขตของภาพไครแมบ.....	59
4.3.3 การวัดประสิทธิภาพความถูกต้องของภาพแมท.....	62
บทที่ 5 สรุปผลการทดลอง.....	65
5.1 สรุปผลการทดลอง.....	65
5.2 การติดตั้งอุปกรณ์และการนำไปใช้งาน.....	66
5.3 ข้อจำกัดและปัญหาที่พบ.....	67
5.4 แนวทางการทำวิจัยต่อ.....	67
เอกสารอ้างอิง.....	66
ภาคผนวก.....	69
บทความและผลงานวิจัยที่ได้รับการตีพิมพ์.....	70

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญรูป

รูปที่	หน้า
2.1 ตัวอย่างการใช้การซ้อนภาพแบบบลูสกรีน เพื่อการซ้อนภาพฉากหลัง ในรายการโทรทัศน์ และอุตสาหกรรมภาพยนตร์.....	5
2.2 แสดงขั้นตอนการซ้อนภาพพื้นหลังโดยทั่วไป.....	7
2.3 การใช้ภาพเมทในการตัดส่วนของภาพที่ไม่ต้องการออก.....	8
2.4 แสดงการจัดเก็บค่าสีในแต่ละพิกเซลของภาพดิจิทัล.....	8
2.5 ขั้นตอนการซ้อนภาพแบบดิจิทัล.....	9
2.6 ตัวอย่างภาพไตรแมป ที่แสดงขอบเขตทั้งสามส่วนของภาพ.....	12
2.7 องค์ประกอบภายในของ Kinect.....	13
2.8 ตัวอย่างภาพ Checkerboard.....	13
2.9 ภาพตัวอย่างความคลาดเคลื่อนเมื่อซ้อนภาพระยะลึกกับภาพสี.....	14
2.10 แสดงกราฟของ $H_\alpha$ เมื่อเทียบกับค่า $\alpha$ ที่แตกต่างกัน.....	19
3.1 ภาพแสดงลำดับขั้นตอนการสร้างภาพไตรแมปที่เหมาะสมจากข้อมูล RGBD.....	26
3.2 (ก) ตัวอย่างระยะลึกก่อนการปรับปรุงแก้ไข (ข) ภาพระยะลึกหลังการปรับปรุง.....	25
3.3 (ก) ภาพระยะลึกหลังการปรับปรุง (ข) ภาพเส้นขอบระยะลึกที่ได้.....	28
3.4 (ก) ภาพตัวอย่างเส้นขอบสีก่อนการคัดกรองด้วยเส้นขอบระยะลึก (ข) ตัวอย่างเส้นขอบสีจากการคัดกรองด้วยเส้นขอบระยะลึกกับข้อมูลความคลาดเคลื่อนของ Kinect.....	28
3.5 (ก) ภาพสีเริ่มต้น (ข) ตัวอย่างภาพเส้นขอบสีก่อนทำการคัดแยกและหลังคัดแยก.....	28
3.6 (ก) ภาพสีต้นฉบับ (ข) ภาพไตรแมปแบบอัตโนมัติจากข้อมูล RGBD.....	29
3.7 (ก) จุด A คือจุดที่ไม่ได้รับแสงอินฟราเรดเพราะถูกบดบังโดยวัตถุ FG (ข) ตัวอย่างภาพระยะลึก เมื่อซ้อนกับภาพสีซึ่ง A คือบริเวณที่เกิดปัญหาการบดบังแสง.....	30
3.8 วัตถุที่มีความหนาน้อยโดยบริเวณเงาที่เกิดจากบังแสงของวัตถุเป็นส่วนของพื้นหลัง.....	32
3.9 วัตถุที่มีความหนามากโดยบริเวณบดบังแสงของวัตถุ.....	32

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญรูป (ต่อ)

รูปที่		หน้า
3.10	ภาพตัวอย่างลักษณะการค้นหาบริเวณที่ไม่ทราบค่า.....	33
3.11	ภาพตัวอย่างลักษณะการแบ่งเส้นระหว่าง $\bar{P}_w(m_j, n_j)$ และ $\bar{P}_R$ ออกเป็นช่วงๆ.....	34
3.12	ภาพการตรวจสอบแนวแสงอินฟราเรดที่จุด A พบว่าที่จุด B บนแนวแสงกับจุด C ที่เป็นขอบของวัตถุที่ทำให้เกิดเงามีระยะห่างระหว่างจุดน้อยและมีระยะลึกใกล้เคียงกัน .....	35
3.13	ตัวอย่างภาพไทรแมบเริ่มต้นที่สร้างจากค่าความเคลื่อนไหวของตัวอุปกรณ์ Kinect กับข้อมูลเส้นขอบระยะลึก .....	37
3.14	(ก) ส่วนข้อมูลสีพื้นหลังจากภาพไทรแมบเริ่มต้น (ข) ส่วนข้อมูลสีวัตถุจากภาพไทรแมบเริ่มต้น (ค) ส่วนข้อมูลสีบริเวณไม่ชัดเจนจากภาพไทรแมบเริ่มต้น .....	37
3.15	(ก) จุดบริเวณที่เป็นรอยต่อเกิดจากการรวมตัวของกลุ่มเส้นขอบสีมากกว่า 1 กลุ่ม (ข) ตัวอย่างภาพพิกเซลบริเวณดังกล่าวซึ่งเกิดจากการรวมตัวของกลุ่มเส้นขอบสีมากกว่า 1 กลุ่ม (ค) ภาพหลังการแยกกลุ่มของเส้นขอบสี .....	40
3.16	(ก) บริเวณเส้นขอบสีที่ทำการครอบ $W$ ที่ตำแหน่ง $D_j$ , (ข) ตัวอย่างการอ้างอิงตำแหน่งเส้นขอบสีที่พบในกรอบ.....	41
3.17	(ก) ภาพส่วนขยายของกลุ่ม $C_{i,g}$ ที่ขนาด $15 \times 15$ (ข) ภาพส่วนขยายของกลุ่ม $C_{i,g}$ ที่ขนาด $5 \times 5$ (ค) บริเวณส่วนต่างที่นำไปใช้ในการคำนวณ .....	42
3.18	ตัวอย่างช่วงการจับคู่ของ $C_{k,s}$ กับ $D_j$ ที่เกิดช่วงที่ไม่มีคู่ .....	46
3.19	ตัวอย่าง Pseudo code การสร้างเส้นขอบวัตถุ.....	47
3.20	ตัวอย่าง Pseudo code ขั้นตอนการสร้างภาพไทรแมบ .....	49
4.1	ตัวอย่างเส้นขอบวัตถุที่กำหนดโดยผู้ใช้งาน .....	50
4.2	ตัวอย่างภาพวัตถุที่มีสีแตกต่างกับภาพพื้นหลังชัดเจน.....	52
4.3	ตัวอย่างภาพวัตถุที่มีสีแตกต่างกับภาพพื้นหลังชัดเจนและไม่มีลวดลาย.....	52
4.4	ภาพที่สีของวัตถุและพื้นหลังใกล้เคียงกันทำให้วิธี OED-BF เกิดความคลาดเคลื่อน.....	57

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญรูป (ต่อ)

รูปที่	หน้า
4.5 ตัวอย่างปัญหาการใช้ Bilateral filter ตรวจจับเส้นขอบวัตถุ (ก) ภาพสี (ข) ภาพระยะลึก ปรับปรุงด้วย Bilateral filter (ค) เส้นขอบวัตถุ.....	60
4.6 กราฟผลการทดลองส่วนความถูกต้องในการกำหนดขอบเขตของภาพไทรแมบ .....	62
4.7 ผลการทดลองเปรียบเทียบค่า RMSE ของภาพวัตถุกับพื้นหลังแบบลวดลายทั้ง 3 วิธี.....	63
5.1 การจัดตำแหน่งของกล้อง DSLR และกล้อง Kinect.....	66



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญตาราง

ตารางที่	หน้า
4.1 ตัวอย่างการทดลองเพื่อหาเส้นขอบวัตถุ.....	53
4.2 ผลการทดลองระยะระหว่างเส้นขอบเมื่อสีวัตถุแตกต่างกับสีพื้นหลัง.....	54
4.3 ตัวอย่างภาพไทรแมบและภาพแมทที่ภาพวัตถุมีสีแตกต่างกับภาพพื้นหลัง.....	55
4.4 ส่วนของภาพแมทแต่ละวิธีที่แตกต่างกับภาพแมทจริงเมื่อสีวัตถุแตกต่างกับพื้นหลัง .....	56
4.5 ผลการทดลองระยะระหว่างเส้นขอบเมื่อสีวัตถุกับสีพื้นหลังมีความคล้ายคลึงกัน .....	57
4.6 ตัวอย่างภาพไทรแมบและภาพแมทที่ภาพวัตถุมีสีคล้ายคลึงกับภาพพื้นหลัง.....	58
4.7 ส่วนของภาพแมทแต่ละวิธีที่แตกต่างกับภาพแมทจริงเมื่อสีวัตถุคล้ายคลึงกับพื้นหลัง .....	59
4.8 แสดงลักษณะผลการคำนวณทดสอบประสิทธิภาพในการกำหนดขอบเขต.....	59
4.9 ข้อมูลรูปภาพวัตถุและภาพพื้นหลังที่ใช้ในการทดลอง .....	60
4.10 ผลการทดลองภาพพื้นหลังแบบ ไม่มีลวดลาย.....	63

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหา

จากความก้าวหน้าทางเทคโนโลยีที่ได้ถูกพัฒนาอย่างต่อเนื่อง เพื่อใช้อำนวยความสะดวกในด้านต่าง ๆ และตอบสนองความต้องการที่เพิ่มขึ้น โดยเฉพาะด้านอุตสาหกรรมภาพยนตร์และสื่อโทรทัศน์ ทำให้เทคโนโลยีการซ้อนภาพในระบบดิจิทัล (Digital matting) ได้เข้ามามีบทบาทเพื่อตอบสนองต่อความต้องการดังกล่าว

สำหรับปัญหาสำคัญของการซ้อนภาพในระบบดิจิทัล คือ การแยกภาพวัตถุออกจากภาพพื้นหลัง เพื่อนำภาพวัตถุนั้นมาซ้อนทับกับภาพพื้นหลังใหม่ โดยการแยกภาพหรือการซ้อนภาพวัตถุนั้นจำเป็นต้องใช้ข้อมูลที่เรียกว่า ภาพแมท (Matte image) ในการตัดบริเวณพื้นหลังในภาพเดิมออกให้เหลือเฉพาะส่วนของภาพที่เป็นวัตถุด้านหน้าและด้านหลัง จากนั้นภาพแมทถูกใช้อีกครั้งในการซ้อนภาพพื้นหลังใหม่เข้ากับภาพที่ตัดพื้นหลังเดิมออกแล้ว เพื่อให้ได้ภาพใหม่ที่ประกอบไปด้วยวัตถุที่ปรากฏอยู่ในภาพต้นฉบับและภาพพื้นหลังใหม่ โดยภาพแมทสามารถคำนวณได้จากวิธีการต่าง ๆ เช่น การซ้อนภาพแบบเบย์เซียน (Bayesian matting) และการซ้อนภาพแบบโคลสฟอร์ม (Closed form matting) เป็นต้น แต่ในการสร้างภาพแมทไม่ว่าด้วยวิธีใดจำเป็นต้องมีข้อมูลตั้งต้นที่เป็นตัวระบุบริเวณของวัตถุด้านหน้าและพื้นหลังในภาพต้นฉบับแบบคร่าว ๆ โดยภาพที่ใช้ในการระบุตำแหน่งของวัตถุและพื้นหลังเพื่อเป็นข้อมูลเบื้องต้นนั้นเรียกว่า ภาพไตรแมป (Trimap image) ซึ่งใช้ในการระบุบริเวณต่าง ๆ ในภาพต้นฉบับอันได้แก่ 1) บริเวณวัตถุเบื้องหน้า (Foreground object region) , 2) บริเวณพื้นหลัง (Background region) และ 3) บริเวณไม่ชัดเจน (Unknown region) ซึ่งอยู่ในช่วงรอยต่อระหว่างบริเวณวัตถุเบื้องหน้าและพื้นหลัง

การสร้างภาพไตรแมปสามารถทำได้โดยการวาดด้วยมือเพื่อกำหนดขอบเขตของภาพวัตถุและพื้นหลัง อย่างไรก็ตามการสร้างภาพไตรแมปด้วยวิธีนี้ต้องการอาศัยเวลาและความชำนาญ การแก้ปัญหาดังกล่าวกระทำได้โดยอาศัยวิธีต่าง ๆ ในการสร้างภาพไตรแมปแบบอัตโนมัติ ซึ่งแต่ละวิธีอาศัยข้อมูลที่แตกต่างกันไปในการระบุบริเวณต่าง ๆ ในภาพตามที่กล่าวมาแล้วข้างต้น เช่น อาศัยความแตกต่างระหว่างค่าสีของวัตถุกับพื้นหลังที่ปรากฏในภาพต้นฉบับ หรือ อาศัยค่าความแตกต่างระยะลึกของวัตถุกับพื้นหลังจากข้อมูลภาพระยะลึก (Depth image) เป็นต้น แต่การใช้ข้อมูลภาพสีหรือข้อมูลระยะลึกเพียงอย่างเดียวอาจไม่เพียงพอ เพราะการใช้ข้อมูลภาพสีเพียงอย่างเดียวเกิดปัญหาในการระบุตำแหน่งของภาพวัตถุเมื่อข้อมูลสีของวัตถุและพื้นหลังมีความใกล้เคียงกัน หรือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตเห็นไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การใช้ข้อมูลระยะลึกเพียงอย่างเดียว ทำให้เกิดปัญหาความคลาดเคลื่อนของตำแหน่งเมื่อเทียบกับภาพวัตถุจริง เพราะความผิดพลาดในการวัดระยะลึกของตัวอุปกรณ์ หรือ การที่ข้อมูลระยะลึกมีความไม่สมบูรณ์ ซึ่งล้วนแล้วแต่ส่งผลกระทบต่อความถูกต้องในการสร้างภาพไตรแมบ อันมีผล ไปถึงการสร้างภาพเมทในที่สุดด้วย งานวิจัยฉบับนี้มีเป้าหมายในการพัฒนาวิธีการสร้างภาพไตรแมบบแบบอัตโนมัติ โดยอาศัยข้อมูลภาพสีและข้อมูลภาพระยะลึกประกอบกัน

## 1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

วิทยานิพนธ์นี้มีวัตถุประสงค์ของการศึกษาเพื่อพัฒนาเทคนิคสร้างภาพไตรแมบที่เหมาะสมแบบอัตโนมัติจากข้อมูลภาพสีและภาพระยะลึก ซึ่งข้อมูลภาพประกอบไปด้วยค่าความสว่างของแม่สีทั้ง 3 และค่าข้อมูลระยะลึก (สีแดง สีเขียว สีน้ำเงิน และ ระยะลึก) (Red Green Blue and Depth หรือเรียกย่อ ๆ ว่า RGBD Image) โดยทำการศึกษาในกรณีของภาพนิ่ง ที่มีภาพพื้นหลังในลักษณะต่าง ๆ ทั้งพื้นหลังสีพื้น และที่มีลวดลายและความหลากหลายของสี

## 1.3 สมมติฐานของการวิจัย

สมมติฐานของการศึกษาสำหรับวิทยานิพนธ์ฉบับนี้คือ การสร้างภาพไตรแมบบแบบอัตโนมัติจากภาพสีและระยะลึก (RGBD Image) ช่วยเพิ่มประสิทธิภาพและคุณภาพของภาพไตรแมบบมากกว่าการสร้างภาพไตรแมบบจากข้อมูลระยะลึกเพียงอย่างเดียว

## 1.4 นิยามศัพท์

1. Image matting หรือ การซ้อนภาพ คือ การแยกภาพวัตถุออกจากภาพพื้นหลัง เพื่อนำภาพวัตถุที่ได้มาซ้อนทับกับภาพพื้นหลังใหม่
2. Color image หรือ ภาพสี คือ รูปภาพสีที่ได้จากการถ่ายภาพ
3. Color edge หรือ เส้นขอบสี คือ เส้นขอบของรูปภาพสีที่ได้จากการคำนวณ
4. Depth image หรือ ภาพระยะลึก คือ รูปภาพที่ได้จากการวัดค่าระยะลึกด้วยอุปกรณ์ Kinect
5. Depth edge หรือ เส้นขอบระยะลึก คือ เส้นขอบที่ได้จากภาพระยะลึก
6. Matte image หรือ ภาพเมท คือ ภาพที่บอกตำแหน่งของวัตถุและพื้นหลัง ใช้สำหรับการคัดแยกภาพวัตถุกับภาพพื้นหลังและการซ้อนภาพวัตถุกับภาพพื้นหลังใหม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7. Trimap image หรือ ภาพไตรเมบ คือ ภาพที่ใช้ระบุตำแหน่งของภาพวัตถุและภาพพื้นหลังที่ชัดเจน เพื่อเป็นข้อมูลเบื้องต้นในการคำนวณเพื่อสร้างภาพเมท ประกอบไปด้วย 3 ส่วนหลักคือ ส่วนของวัตถุเบื้องหน้า, (Foreground object) ส่วนของพื้นหลัง (Background object) และส่วนของบริเวณไม่ชัดเจน (Unknown region)
8. Unknown region หรือ บริเวณไม่ชัดเจน คือ บริเวณรอยต่อระหว่างภาพวัตถุและพื้นหลังที่ไม่สามารถระบุได้แน่ชัดจนว่าเป็นของภาพวัตถุหรือภาพพื้นหลัง เพราะเกิดจากการผสมกันระหว่างสีของภาพวัตถุและสีของภาพพื้นหลัง
9. RGBD Image คือ ข้อมูลภาพสี และข้อมูลภาพระยะลึก ซึ่งข้อมูลภาพ RGBD ประกอบไปด้วยค่าความสว่างของแม่สีทั้ง 3 (สีแดง สีเขียว สีน้ำเงิน และข้อมูลระยะลึก)

### 1.5 ทฤษฎีและแนวคิดที่ใช้ในการวิจัย

1. ทฤษฎีการซ้อนภาพในระบบดิจิทัล (Digital matting) คือ ทฤษฎีเรื่องการแยกภาพวัตถุออกจากภาพพื้นหลังในระบบดิจิทัล เพื่อนำภาพวัตถุที่ได้มาซ้อนรวมกับภาพพื้นหลังใหม่ที่ได้จัดเตรียมไว้ในระบบดิจิทัลจนเกิดเป็นภาพใหม่ที่มีภาพวัตถุเดิมกับภาพพื้นหลังใหม่
2. ทฤษฎีเรื่องรูปภาพสีและรูปภาพระยะลึก (RGBD Image) คือ ทฤษฎีเรื่องข้อมูลรูปภาพ ซึ่งประกอบไปด้วยค่าความสว่างของแม่สีทั้ง 3 และค่าข้อมูลระยะลึก (สีแดง สีเขียว สีน้ำเงิน และ ระยะลึก)
3. ทฤษฎีรูปแบบแสง (Structured-light) คือ ทฤษฎีเรื่องการคำนวณเพื่อหาข้อมูลระยะลึกของวัตถุจากในรูปแบบแสงที่สะท้อน
4. ทฤษฎีเกี่ยวกับการประมวลผลรูปภาพ (Image processing techniques) คือ ทฤษฎีเรื่องการประมวลผลของรูปภาพเพื่อข้อมูลที่ต้องการ เช่น การค้นหาเส้นขอบของภาพ เป็นต้น
5. ทฤษฎีการจัดกลุ่ม (K -mean) คือ ทฤษฎีเรื่องการจัดกลุ่มของข้อมูลที่มีความเหมือนหรือคล้ายคลึงกัน

## 1.6 ขอบเขตการวิจัย

ขอบเขตของการวิจัยเป็นการพัฒนาการเทคนิคการสร้างภาพไตรแมบที่เหมาะสมแบบอัตโนมัติโดยอาศัยข้อมูล RGBD จากกล้องถ่ายภาพประเภท structured-light depth camera (Kinect) พิจารณาในประเด็นที่ภาพพื้นหลังมีลักษณะเป็นลวดลายแตกต่างกัน โดยที่ภาพวัตถุมีระยะลึกห่างกับกล้องอย่างน้อย 1 เมตรและมีระยะห่างกับภาพพื้นหลังมากกว่า 50 ซม. ขึ้นไป โดยทำการวิจัยกับข้อมูลภาพถ่ายจริงและภาพสังเคราะห์ ในประเด็นหลักต่อไปนี้

1. การปรับปรุงคุณภาพของภาพถ่ายระยะลึกที่ได้จากกล้อง Kinect
2. การพัฒนาเทคนิคการสร้างภาพไตรแมบจากภาพ RGBD ที่ทำงานได้กับพื้นหลังในลักษณะต่าง ๆ
3. การศึกษาผลกระทบในแง่ความถูกต้องแม่นยำของการกำหนดขอบเขตบริเวณไม่ชัดเจนที่มีผลต่อคุณภาพของการซ้อนภาพ

โดยในการวิจัยพิจารณาประสิทธิภาพของการซ้อนภาพด้วยวิธีการซ้อนภาพแบบ โคลสฟอร์ม (Closed form matting) เปรียบเทียบกับการสร้างภาพไตรแมบด้วยมือ

## 1.7 ขั้นตอนการวิจัย

1. กำหนดหัวข้อ, เป้าหมาย, วัตถุประสงค์ และขอบเขตการทำวิทยานิพนธ์
2. ศึกษาทฤษฎีและหลักการพื้นฐานที่ใช้ในการวิจัย
3. ศึกษาเทคนิคต่าง ๆ ที่มีอยู่รวมถึงแนวคิด หลักการ ข้อดี และข้อบกพร่องของแต่ละวิธี
4. ศึกษาทำการทดลองตามสมมุติฐาน
5. ปรับปรุงข้อผิดพลาด และสรุปผล
6. จัดทำเอกสารประกอบวิทยานิพนธ์

## บทที่ 2

# ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 ทฤษฎีการซ้อนภาพ (Matting)

เทคนิคการซ้อนภาพ (Matting) เป็นวิธีการที่ใช้กันอย่างแพร่หลายในการถ่ายทำภาพยนตร์สื่อวิดีโอหรือโทรทัศน์ บางครั้งมีความจำเป็นต้องปรับเปลี่ยนพื้นหลัง (Background) ที่ปรากฏในภาพที่ถ่ายทำในครั้งแรกด้วยพื้นหรือฉากหลังใหม่โดยไม่ให้กระทบกับภาพของผู้แสดงหรือวัตถุที่อยู่ด้านหน้า (Foreground object) ทั้งนี้อาจด้วยเหตุผลของความจำเป็นในการลดค่าใช้จ่ายในการผลิตคิวนัดหมายนักแสดงในการถ่ายทำที่จำกัดหรือเพื่อประกอบการสร้างเทคนิคพิเศษด้านภาพ (Visual/Special effects) ที่ไม่สามารถหาสร้างหรือจำลองฉากหลังจริงขึ้นมาได้ เป็นต้น ซึ่งเทคนิคในการซ้อนภาพฉากหลังดังกล่าว เป็นที่รู้จักและใช้งานกันอย่างแพร่หลายมาหลายสิบปี แต่อาจเป็นที่รู้จักกันในชื่อที่แตกต่างกัน เช่น เทคนิค แบบบลูสกรีน (Blue screen matting) หรือเทคนิค โครมาคีย์ (Chroma key)

ในอดีตกาลซ้อนภาพลงบนแผ่นฟิล์มที่ใช้ในการถ่ายทำภาพยนตร์ ต้องอาศัยกระบวนการในการล้างฟิล์มและถ่ายทำซ้ำหลายครั้ง เพื่อให้ได้ผลที่ต้องการ ในขณะที่กระบวนการในระบบการบันทึกภาพด้วยเทปวีดิทัศน์ (Video tape) ที่ใช้กันแพร่หลายในการผลิตรายการโทรทัศน์ จะอาศัยการประมวลผลภาพด้วยอุปกรณ์อิเล็กทรอนิกส์ในระบบอนาล็อก ซึ่งเทคโนโลยีเหล่านี้ต้องอาศัยการจัดฉากหลังและแสงที่เหมาะสม จึงจะได้คุณภาพของภาพที่ต้องการ นอกจากนี้ยังมีข้อจำกัดที่สำคัญบางประการ กล่าวคือให้ผลไม่ดีนักกับบริเวณของภาพที่ปรากฏวัตถุด้านหน้าที่มีขนาดเล็ก เช่น เส้นผม หรือวัตถุที่มีความโปร่งแสงบางส่วน รวมทั้งความจำเป็นที่ต้องจัดฉากหลังในการถ่ายทำให้เป็นพื้นสีเดียว เช่น สีน้ำเงิน หรือ สีเขียว ไม่สามารถใช้กับภาพที่ถ่ายทำโดยติดพื้นหลังใด ๆ ที่ปรากฏตามธรรมชาติได้ [1-2]



รูปที่ 2.1 ตัวอย่างการใช้การซ้อนภาพแบบบลูสกรีนเพื่อการซ้อนภาพฉากหลัง ในรายการโทรทัศน์

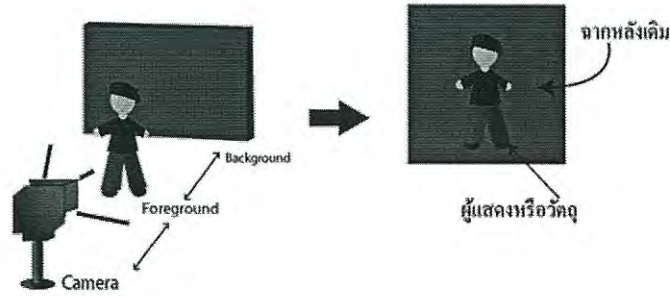
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าและอุตสาหกรรมภาพยนตร์  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด้วยความก้าวหน้าของเทคโนโลยีดิจิทัลและคอมพิวเตอร์ ทำให้ในปัจจุบัน เทคนิคการซ้อนภาพได้ถูกพัฒนาเพื่อให้สามารถใช้ประโยชน์จากประสิทธิภาพที่สูงขึ้นของคอมพิวเตอร์ได้ ซึ่งเทคนิคการซ้อนภาพแบบดิจิทัลใหม่ ๆ ที่มีการพัฒนาขึ้นช่วยลดข้อจำกัดหลายประการของเทคนิคการซ้อนภาพในแบบเดิมได้ บทความนี้ได้นำเสนอหลักการโดยทั่วไปของเทคนิคการซ้อนภาพ ทั้งในระบบเดิมและในระบบดิจิทัล จากนั้นอธิบายถึงหลักการพื้นฐานของเทคนิคแบบดิจิทัล พร้อมทั้งระบุข้อดีของเทคนิคดังกล่าวเมื่อเปรียบเทียบกับเทคนิคแบบเดิมในระบบอนาล็อก โดยเฉพาะอย่างยิ่งความสามารถของเทคนิคการซ้อนภาพแบบดิจิทัลใหม่ ๆ ที่สามารถทำงานได้กับภาพที่ถ่ายโดยติดพื้นหลังใด ๆ ซึ่งเรียกรวม ๆ กันว่าเป็นพื้นหลังธรรมชาติได้

### 2.1.1 หลักการทั่วไปของการซ้อนภาพพื้นหลัง

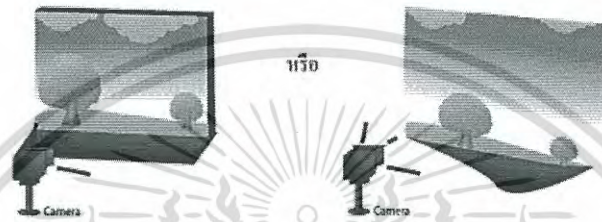
การซ้อนภาพพื้นหลังมีหลักการและขั้นตอน โดยทั่วไป แสดงดังในรูปที่ 2.2 จากภาพสามารถอธิบายขั้นตอนต่าง ๆ ได้โดยสังเขป ดังนี้

1. ทำการถ่ายภาพนักแสดงหรือวัตถุที่ปรากฏอยู่ด้านหน้ากล้องในภาพ โดยในขั้นตอนนี้ สำหรับเทคนิคการซ้อนภาพโดยทั่วไป จำเป็นต้องจัดฉาก (Screen) ที่เป็นพื้นหลัง สีเดียว เช่น สีน้ำเงิน หรือสีเขียว ในขณะที่เทคนิคการซ้อนภาพแบบดิจิทัลสมัยใหม่ อาจเป็นพื้นหลังใด ๆ เช่น พื้นหลังที่ปรากฏตามธรรมชาติ ณ สถานที่ที่ถ่ายทำนั้น ๆ ในภาพตัวอย่าง แสดงเฉพาะกรณีที่พื้นหลังเป็นฉากสีน้ำเงิน เพื่อง่ายต่อการนำเสนอ
2. ทำการจัดเตรียมฉากหลังใหม่ ซึ่งอาจจัดทำโดยการวาดหรือสร้างด้วยคอมพิวเตอร์ หรือการถ่ายภาพวัตถุกับสถานที่จริงก็ได้
3. ทำการตัดส่วนที่เป็นฉากหลังออกไปจากภาพที่ถ่ายได้จากขั้นตอนที่ 1 ซึ่งมีผลทำให้บริเวณดังกล่าวเป็นสีดำสนิท
4. จากภาพที่ได้ในขั้นตอนที่ 2 ทำกระบวนการเดียวกับในขั้นตอนที่ 3 หากแต่ในขั้นตอนนี้ ตัดภาพในบริเวณที่ตรงกับตำแหน่งของนักแสดงในภาพที่ถ่ายจากขั้นตอนที่ 1 แทน
5. นำภาพที่ได้จากขั้นตอนที่ 3 และ 4 มาผสมเข้าด้วยกัน โดยส่วนที่เป็นสีดำในแต่ละภาพ ไม่มีผลต่อสีหรือความสว่างของภาพบริเวณนั้นหลังการผสมเข้าด้วยกัน ทำให้ในภาพสุดท้าย บริเวณที่เป็นนักแสดงหรือวัตถุที่อยู่ด้านหน้ายังคงเหมือนภาพต้นฉบับที่ถ่ายได้จากในขั้นตอนที่ 1 มีการเปลี่ยนแปลงเฉพาะบริเวณฉากหลังเดิมที่ถูกแทนที่ด้วยฉากหลังใหม่



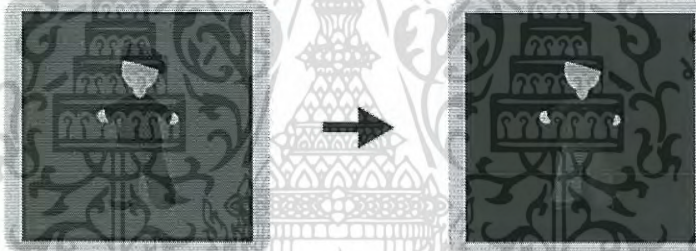
ก

1. ถ่ายภาพนักแสดงที่มีพื้นหลังเป็นสีเดียว



ข

2. ถ่ายหรือสร้างภาพฉากหลังใหม่ที่ต้องการ



ค

3. ตัดฉากหลังออกจากภาพต้นฉบับ



ง

4. ตัดบริเวณที่ซ้อนทับตำแหน่งนักแสดงออกจากภาพฉากหลังใหม่



จ

5. นำภาพที่ได้ขึ้นตอนที่จาก 3 และ 4 มารวมกัน

รูปที่ 2.2 แสดงขั้นตอนการซ้อนภาพพื้นหลังโดยทั่วไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

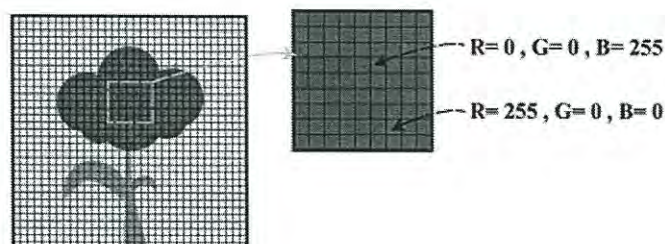
ในทางปฏิบัติ การตัดภาพส่วนที่ไม่ต้องการออกจากภาพต้นฉบับในขั้นตอนที่ 3 (หรือขั้นตอนที่ 4 ที่กระทำกับภาพฉากหลังใหม่) จำเป็นต้องทำการสร้างภาพแมท (Matte image) ซึ่งเป็นภาพที่มีขนาดเท่ากับภาพในขั้นตอนที่ 1 โดยภาพแมทส่วนที่ตรงกับบริเวณของภาพที่ไม่ต้องการให้ถูกตัดออกไป จะมีลักษณะโปร่งแสงเหมือนแผ่นใส ในขณะที่บริเวณที่ต้องการตัดออก มีลักษณะทึบแสง เมื่อนำภาพแมทมาวางไว้เหนือภาพต้นฉบับที่เตรียมไว้ (จากขั้นตอนที่ 1 หรือ 2) ก็จะได้ภาพสุดท้ายตามต้องการ (ดูรูปที่ 2.3 ประกอบ) กระบวนการในรายละเอียดสำหรับแต่ละขั้นตอน อาจแตกต่างกันออกไป ขึ้นอยู่กับเทคโนโลยีที่ใช้ เช่น ในกรณีของการใช้ฟิล์มภาพยนตร์ ภาพแมทถูกสร้างโดยกระบวนการถ่ายภาพลงบนแผ่นฟิล์มที่มีลักษณะเฉพาะ และในการรวมภาพเข้าด้วยกัน กระทำโดยการถ่ายภาพ [2-3] ที่ได้จากการนำฟิล์มของภาพแมทและภาพต้นฉบับมาซ้อนทับกัน



รูปที่ 2.3 การใช้ภาพแมทในการตัดส่วนของภาพที่ไม่ต้องการออก

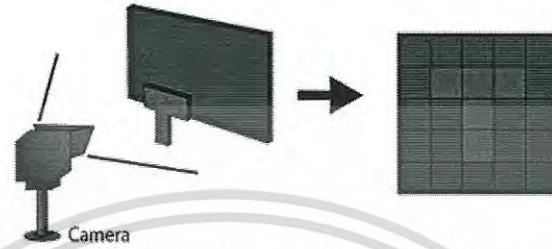
### 2.1.2 เทคนิคการซ้อนภาพฉากหลังแบบดิจิทัล

ภาพดิจิทัล ไม่ว่าจะเป็นภาพนิ่ง หรือภาพเคลื่อนไหว มีหลักการพื้นฐานเหมือนกัน กล่าวคือ ข้อมูลภาพถูกตัดแบ่งออกเป็นบริเวณหรือจุดย่อย ๆ เป็นจำนวนมาก แต่ละจุดเรียกว่า พิกเซล (pixel) โดยการจัดเก็บภาพในรูปแบบดิจิทัล ทำการบันทึกค่าสีหรือความสว่างในแต่ละพิกเซลไว้ (ดูรูปที่ 2.4 ประกอบ)

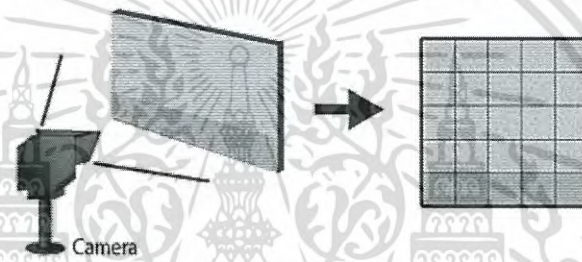


รูปที่ 2.4 แสดงการจัดเก็บค่าสีในแต่ละพิกเซลของภาพดิจิทัล

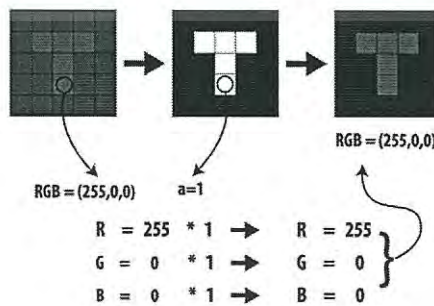
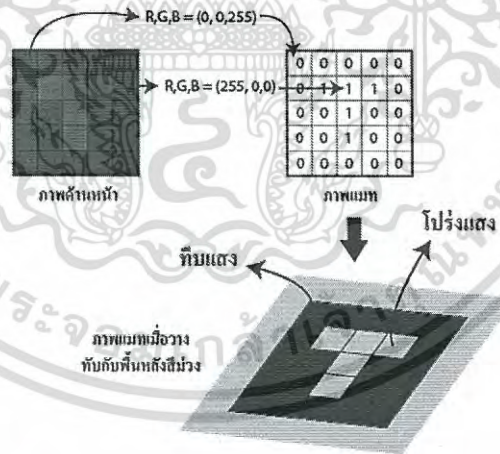
สำหรับภาพสี แต่ละพิกเซลประกอบไปด้วยค่าความสว่างของแม่สีทั้ง 3 อันได้แก่ สีแดง สีเขียว และสีน้ำเงิน (Red Green Blue หรือเรียกย่อๆว่า RGB) โดยแต่ละสีมีค่าความสว่างระหว่าง 0 ถึง 255 เป็นต้น เมื่อพิจารณากระบวนการซ้อนภาพดังที่ได้อธิบายไว้ในหัวข้อ 2.2 และ รูปที่ 2 ในกรณีของภาพดิจิทัล มีรายละเอียดดังแสดงในรูปที่ 2.5



1. ภาพดิจิทัล (ขาว) ที่ได้จากการถ่ายภาพวัตถุ โดยใช้พื้นหลังสีเดียว



2. ภาพดิจิทัล (ขาว) ที่ได้จากการถ่ายภาพจากหลังใหม่ที่ต้องการ



3. การสร้างภาพแมสก์ และกระบวนการในการรวมภาพเข้าด้วยกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับ... **รูปที่ 2.5 ขั้นตอนการซ้อนภาพแบบดิจิทัล** ...  
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 2.5 หลังผ่านขั้นตอนที่ 1 และ 2 ได้ภาพคิจิตอลของนักแสดงหรือวัตถุที่อยู่ด้านหน้า และภาพฉากหลังใหม่ ตามลำดับ เมื่อเข้าสู่ขั้นตอนที่ 3 เริ่มจากการสร้างภาพแมท โดยภาพแมทเป็นภาพคิจิตอลที่มีขนาดเท่ากับภาพต้นฉบับ แต่มีชื่อแตกต่างกันตรงที่ในแต่ละพิกเซลของภาพแมท ทำการจัดเก็บค่าระดับความโปร่งแสงเป็นตัวเลขตัวเดียวที่มีค่าอยู่ระหว่าง 0 ถึง 1 แทนการเก็บค่าสี RGB จำนวน 3 ค่าเหมือนในกรณีของภาพต้นฉบับ โดยการสร้างภาพแมท ขึ้นอยู่กับลักษณะของฉากหลังเดิม ในหัวข้อนี้ ขอยกตัวอย่างกรณีที่ฉากหลังเป็นพื้นหลังสีเดียว ที่ทราบค่าสีล่วงหน้า ในกรณีนี้ โปรแกรมที่ใช้ในการประมวลผลภาพเพื่อการซ้อนภาพฉากหลัง ทำการอ่านค่าสีในแต่ละพิกเซลของภาพที่ได้จากขั้นตอนที่ 1 หาก ณ ตำแหน่งพิกเซลใดมีค่าสีตรงหรือใกล้เคียงกับค่าสีของฉากหลัง (ที่ทราบค่าอยู่ก่อนแล้ว) ให้ทำการกำหนดค่าสีของภาพแมท ณ ตำแหน่งพิกเซลที่ตรงกันให้เป็นสีดำหรือทึบแสง (ค่าของพิกเซลเท่ากับ 0) หากสีไม่ตรงหรือไม่ใกล้เคียงกับสีฉากหลัง กำหนดให้พิกเซลดังกล่าวของภาพแมทเป็นจุดโปร่งแสงแทน (ค่าของพิกเซลเท่ากับ 1) หลังจากได้ภาพแมทแล้ว การตัดภาพในส่วนที่ไม่ต้องการ กระทำโดยการคูณค่าความโปร่งแสงที่เก็บไว้ในแต่ละพิกเซลของภาพแมทเข้ากับค่าสีในพิกเซลที่ตำแหน่งตรงกันกับภาพต้น เป็นผลทำให้ภาพที่ได้ที่อยู่ในส่วนที่ทึบแสงของภาพแมท ปรากฏเป็นสีดำเสมอ ในขณะที่ส่วนที่ตรงกับบริเวณโปร่งแสง จะไม่มีการเปลี่ยนแปลงของค่าสีจากภาพต้นฉบับแต่อย่างใด อนึ่งภาพแมทดังกล่าวซึ่งบ่งบอกระดับความโปร่งแสงของแต่ละพิกเซล โดยบางครั้งจะทำการบันทึกรวมไปกับข้อมูลค่าสี RGB ของรูปภาพ เรียกว่าค่าแอลฟา (Alpha value)

สำหรับกรณีของภาพฉากหลังใหม่ ภาพแมทที่ใช้เพื่อตัดส่วนของภาพที่ซ้อนกับตำแหน่งวัตถุในภาพต้นฉบับ สามารถสร้างได้จากภาพแมทสำหรับรูปภาพต้นฉบับข้างต้น โดยการสลับค่าความโปร่งแสงในแต่ละพิกเซล เช่น ในกรณีที่ค่าเดิมเป็น 0 ให้สลับเป็น 1 ในขณะที่หากเดิมค่าความโปร่งแสงในภาพแมทเป็น 1 ให้สลับเป็น 0 ได้ภาพแมทที่ต้องใช้สำหรับภาพฉากหลังใหม่ตามขั้นตอนที่ 4 ดังที่ได้กล่าวไปในหัวข้อก่อนหน้านี้

หลังจากได้ภาพที่ผ่านกระบวนการตัดส่วนภาพออกแล้ว การผสมภาพ ทำได้โดยการนำค่าสีของแต่ละพิกเซลที่ตรงกันของภาพต้นฉบับและภาพฉากหลังใหม่ มาบวกเข้าด้วยกัน

### 2.1.3 วิธีการแยกภาพฉากหลังธรรมชาติ

ในกระบวนการซ้อนภาพ ขั้นตอนที่สำคัญ คือการแยกภาพวัตถุด้านหน้าออกจากภาพฉากหลัง กระบวนการดังกล่าวมีความยุ่งยากมากขึ้น เมื่อฉากหลังเป็นวัตถุใด ๆ ที่ไม่สามารถควบคุมสีและความสม่ำเสมอของสีหรือความสว่างได้ ตลอดจนภาพที่พิกเซลบริเวณขอบวัตถุด้านหน้ากับฉากหลังมีการผสมผสานกัน ทำให้กระบวนการแยกฉากหลังออกมีความซับซ้อนมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อให้สามารถทำการซ้อนภาพในกรณีดังกล่าวข้างต้น จำเป็นต้องกำหนดให้พิกเซลบริเวณรอยต่อระหว่างวัตถุด้านหน้ากับพื้นหลัง เป็นการผสมกันระหว่างสีของวัตถุด้านหน้ากับสีพื้นหลัง ตามสมการต่อไปนี้

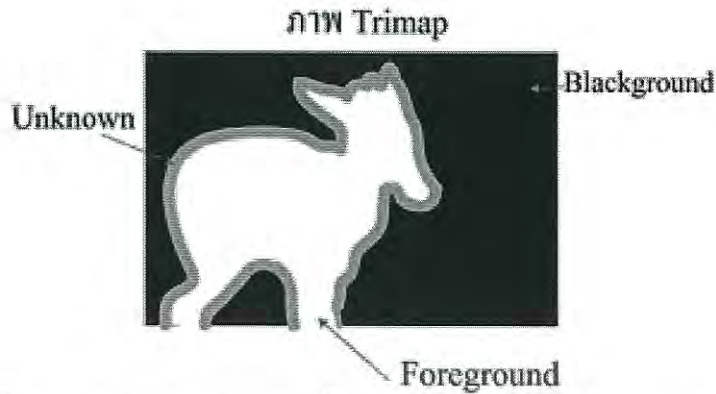
$$I_k = \alpha_k F_k + (1 - \alpha_k) B_k \quad (2.1)$$

โดยที่ จากสมการข้างต้น  $F_k$  คือค่าสี ณ พิกเซลที่  $k$  อันเกิดจากวัตถุที่อยู่ด้านหน้า (สำหรับภาพสี  $F_k$  มี 3 ค่าคือค่า สีแดง เขียว และน้ำเงิน แต่ในที่นี้ละไว้เพื่อให้การนำเสนอมีความกระชับ) ส่วน  $B_k$  คือค่าสี ณ พิกเซลที่  $k$  อันเกิดจากวัตถุในฉากหลัง ค่า  $\alpha_k$  คือค่าความโปร่งแสงของวัตถุที่อยู่ด้านหน้าในส่วนที่ตรงกับตำแหน่งพิกเซลที่  $k$  ส่วน  $I_k$  คือค่าสี ณ พิกเซลที่  $k$  ของภาพต้นฉบับที่ถ่ายได้

จากสมการที่ 2.1 หากสามารถคำนวณหรือหาค่า  $\alpha_k$  และ  $F_k$  จากภาพที่ถ่าย ( $I_k$ ) ได้ ก็สามารถสร้างภาพที่ซ้อนจากหลังใหม่ได้ด้วยสมการต่อไปนี้

$$J_k = \alpha_k F_k + (1 - \alpha_k) B'_k \quad (2.2)$$

โดยที่จากสมการข้างต้น  $B'_k$  คือค่าสี ณ พิกเซลที่  $k$  ของภาพฉากหลังใหม่ และ  $J_k$  คือค่าสี ณ พิกเซลที่  $k$  ของภาพใหม่ที่ได้ ปัญหาคือการคำนวณหาค่า  $\alpha_k$  และ  $F_k$  จาก  $I_k$  ด้วยสมการที่ 2.1 โดยทั่วไป ไม่สามารถหาค่าดังกล่าวได้ เนื่องจากมีตัวแปรที่ไม่ทราบค่าถึง 3 ตัว อย่างไรก็ตาม ในทางปฏิบัติอาศัยการเลือกค่าสีของวัตถุที่อยู่ด้านหน้า และค่าสีของวัตถุในฉากหลัง ตรงบริเวณที่ใกล้เคียงกับขอบวัตถุที่ทราบค่าอยู่ก่อนแล้ว ในการประมาณค่า  $F_k$  และ  $I_k$  เพื่อให้สามารถแก้สมการหาค่าของ  $\alpha_k$  ได้ ด้วยวิธีการดังกล่าว สิ่งสำคัญที่ต้องกำหนด (โดยผู้ใช้) จากภาพต้นฉบับ คือบริเวณในภาพที่เป็นวัตถุที่อยู่ด้านหน้า และบริเวณในภาพที่เป็นวัตถุถัดไปที่อยู่ในฉากหลัง ข้อมูลดังกล่าว ซึ่งโดยทั่วไปกำหนดให้อยู่ในรูปของแผนภาพไตรแมบ (ดูรูปที่ 2.6 ประกอบ) ถูกนำไปใช้ประกอบกับภาพสีต้นฉบับในการคำนวณหาภาพเมท เพื่อใช้แยกฉากหลังธรรมชาติ ซึ่งโดยทั่วไปมีขั้นตอนดังต่อไปนี้



รูปที่ 2.6 ตัวอย่างภาพไตรแมป ที่แสดงขอบเขตทั้งสามส่วนของภาพ

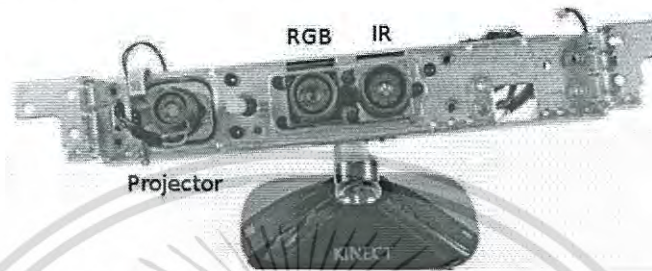
1. สำหรับแต่ละจุดพิกเซล  $k$  ใด ๆ ที่อยู่ในบริเวณไม่ชัดเจน ( Unknown region ) ทำการเลือกจุดพิกเซลที่ใกล้เคียงจากบริเวณพื้นหลัง ( Background region ) และใช้ค่าสีของพิกเซลดังกล่าวแทนค่า  $B_k$  ที่แท้จริง และในทำนองเดียวกัน ทำการเลือกจุดพิกเซลที่ใกล้เคียงจากบริเวณเบื้องหน้า ( Foreground region ) และใช้ค่าสีของพิกเซลดังกล่าวแทนค่า  $F_k$  ที่แท้จริง เช่น อาจเลือกใช้จุดในบริเวณเบื้องหน้าหรือบริเวณพื้นหลัง ที่อยู่ใกล้กับจุดที่ต้องการคำนวณมากที่สุด เป็นต้น
2. ทำการคำนวณหาค่าความโปร่งแสง ( $\alpha_k$ ) จากสมการที่ (2.1)
3. สำหรับตำแหน่งของพิกเซลที่อยู่ในบริเวณพื้นหลังมากำหนดค่า  $\alpha_k$  เป็น 0 และตำแหน่งที่อยู่ใน Foreground region กำหนดค่า  $\alpha_k$  เป็น 1
4. ทำการซ้อนภาพฉากหลังใหม่ โดยใช้สมการที่ (2.1) โดยอาศัยค่า  $\alpha_k$  ที่คำนวณได้จากขั้นตอนที่แล้ว และค่าประมาณการของ  $F_k$  จากในขั้นตอนที่ 2
5. คุณภาพของการซ้อนภาพอาจปรับปรุงให้ดีขึ้น ด้วยการคำนวณค่าของ  $B_k$  และ  $F_k$  ให้ใกล้เคียงกับค่าสีที่แท้จริงมากขึ้น ส่งผลให้รายละเอียดภาพมีความถูกต้องมากขึ้น

## 2.2 กล้องถ่ายภาพแบบ RGB-D

กล้องถ่ายภาพแบบ RGB-D คือกล้องที่มีคุณสมบัติในการถ่ายภาพสีและบันทึกข้อมูลระยะความลึกของภาพ ซึ่งปัจจุบันสามารถแบ่งออกได้เป็น 3 ประเภทหลักตามลักษณะการคำนวณระยะความลึกดังนี้ 1) แบบ Time-of-flight sensor (TOF sensor) ที่คำนวณระยะลึกจากระยะเวลาการเดินทางของแสง, 2) แบบ Stereo camera ที่อาศัยภาพหลายมุมมองในการคำนวณหาระยะความลึก และ 3) แบบ Structured light ที่คำนวณระยะลึกจากแสงอินฟราเรดที่สะท้อนกลับ ซึ่งในแต่ละแบบนี้มีข้อดีและข้อเสียที่แตกต่างกันไป สำหรับรายงานวิจัยฉบับนี้ใช้อุปกรณ์ Kinect ที่ใช้หลักการ Structured light ในการคำนวณค้นหาข้อมูลระยะลึกมาใช้

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Kinect คือ อุปกรณ์ที่มีความสามารถในการถ่ายภาพระยะลึก ซึ่งภายในประกอบไปด้วย 3 ส่วนหลักคือ กล้องอินฟราเรด (Infrared-IR camera), ส่วนฉายแสงอินฟราเรด (IR projector) และกล้องสี (RGB camera) ดังรูปที่ 2.7 โดย Kinect สามารถคำนวณระยะลึกจากการฉายแสงอินฟราเรดซึ่งเป็นแสงที่มองไม่เห็นด้วยตาเปล่าไปยังภาพแบบมีแบบแผน โดยกล้องอินฟราเรดทำหน้าที่รับแสงสะท้อนกลับมาทำการคำนวณหาระยะความลึก



รูปที่ 2.7 องค์ประกอบภายในของ Kinect

### 2.2.1 การปรับตั้งค่าอุปกรณ์ Kinect

สิ่งหนึ่งที่เป็นต่อการคำนวณเพื่อให้ผลที่ถูกต้องและเกิดประสิทธิภาพ คือ การปรับตั้งค่า (Calibration) เพราะในกล้องแต่ละตัวที่ผลิตออกมามีสิ่งที่เรียกว่าความยาวโฟกัส (Focal length) และขนาดของเซนเซอร์ในกล้องที่มีความแตกต่างกัน ซึ่งสามารถเกิดได้ความคลาดเคลื่อนของจากสายการผลิต ซึ่งประโยชน์ในการปรับตั้งค่าเพื่อให้ได้ประสิทธิภาพและความถูกต้องสูงที่สุด สำหรับการปรับตั้งค่าปกติทั่วไปสามารถทำได้จากการถ่ายภาพลายกระดานหมากรุก (Checkerboard) มาใช้ในการคำนวณหาค่าตำแหน่งที่แท้จริงของตัวอุปกรณ์ ในงานวิจัยฉบับนี้ใช้ซอฟต์แวร์ MATLAB Toolbox ที่เผยแพร่ใน [8] เพื่อการปรับตั้งค่าดังกล่าว



รูปที่ 2.8 ตัวอย่างภาพ Checkerboard

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.2.2 การคำนวณค่าความคลาดเคลื่อนของตัวอุปกรณ์

สิ่งหนึ่งที่ส่งผลต่อประสิทธิภาพและความถูกต้องการนำข้อมูลภาพระยะลึกไปใช้ คือ ความคลาดเคลื่อนของภาพระยะลึกที่เกิดจากตัวอุปกรณ์ในการวัดระยะลึก ซึ่งมีผลต่อความถูกต้องคำนวณหาตำแหน่งที่ระหว่างภาพระยะลึกกับภาพสี กล่าวคือ บริเวณที่เป็นวัตถุแต่ข้อมูลระยะลึกถูกตรวจสอบเป็นระยะลึกของพื้นหลัง หรือ บริเวณที่เป็นพื้นหลังแต่ข้อมูลระยะลึกตรวจสอบเป็นระยะลึกของวัตถุ ดังรูปที่ 2.9 ซึ่งผลความคลาดเคลื่อนที่เกิดขึ้นส่งผลต่อใช้ข้อมูลระยะลึกในการค้นหาขอบของภาพวัตถุหรือการใช้ข้อมูลระยะลึกในการสร้างภาพ ไตรแมบ



รูปที่ 2.9 ภาพตัวอย่างความคลาดเคลื่อนเมื่อซ้อนภาพระยะลึกกับภาพสี

สำหรับการคำนวณค่าความคลาดเคลื่อนทำได้โดยการวัดความถูกต้องของข้อมูลระยะลึกของตัวอุปกรณ์เทียบกับข้อมูลระยะลึกจริงของวัตถุ ซึ่งกล้อง Kinect มีการวัดข้อมูลระยะลึกคลาดเคลื่อนไปจากความเป็นจริงอยู่ที่ประมาณ 20 ซม. โดยสามารถทำการแปลงค่าความคลาดเคลื่อนที่เกิดขึ้นเป็นหน่วยพิกเซลได้จากการ Projection ไปยังตำแหน่งพิกเซลของภาพสีเมื่อข้อมูลระยะลึกของภาพมีการเปลี่ยนแปลงสามารถทำได้ตามขั้นตอนต่อไปนี้

1. ปกติภาพที่ได้จากข้อมูลระยะลึกนั้นสามารถแบ่งออกได้เป็น 3 ส่วนคือ 1) ส่วนบริเวณระยะลึกวัตถุ , 2) บริเวณระยะลึกพื้นหลัง และ 3) บริเวณที่ไม่ใช่ข้อมูลระยะลึก โดยการคำนวณความคลาดเคลื่อนของตัวอุปกรณ์ใช้ข้อมูลในส่วนขอบของภาพระยะลึกวัตถุในการตรวจสอบ กำหนดให้  $Q^t$  คือ ความคลาดเคลื่อนซึ่งเกิดจากการ Projection ของ Kinect ไปยังรูปภาพ เมื่อเกิดการเปลี่ยนแปลงของข้อมูลระยะลึกในแนวแกน z และ กำหนดให้  $I_d(x, y)$  คือข้อมูลภาพระยะลึก โดยที่  $x, y$  คือตำแหน่งพิกเซลของวัตถุในภาพระยะลึก และ  $I_d$  คือภาพระยะลึก

2. ค้นหาตำแหน่งระบบพิกัดจริงของวัตถุ (World coordinate) จากที่ตรงกับตำแหน่ง  $(x, y)$  กำหนดให้คือ  $\vec{P}_d = [p_x, p_y, p_z]^T$  โดยที่  $p_x, p_y$  และ  $p_z$  คือ ตำแหน่งที่พบระบบพิกัดจริงของวัตถุ
3. หาค่าตำแหน่งของภาพวัตถุเมื่อค่าระยะลึก  $p_z$  จากตำแหน่งในระบบพิกัดจริงเกิดการเปลี่ยนแปลงของข้อมูลระยะลึก กำหนดให้  $x', y'$  และ  $x'', y''$  คือตำแหน่งการโปรเจกชันมายังภาพของ  $\vec{P}_d$  เมื่อค่าระยะลึก  $p_z$  มีการเปลี่ยนแปลงไปจากเดิมที่  $p_z + 0.2$  และ  $p_z - 0.2$  ตามลำดับ
4. คำนวณหาค่าความคลาดเคลื่อนที่เกิดขึ้นจากสมการต่อไปนี้ และโดยกำหนดให้  $f$  คือ ปัจจัยอื่นๆที่อาจส่งผลกระทบต่อความคลาดเคลื่อน

$$pt' = \sqrt{(y - y')^2 + (x - x')^2} \quad (2.3)$$

$$pt'' = \sqrt{(y - y'')^2 + (x - x'')^2} \quad (2.4)$$

$$Q^t = pt' + pt'' + f \quad (2.5)$$

5. ทำขั้นตอนที่ 2 ถึง 4 ซ้ำจนครบทุกตำแหน่งของภาพวัตถุเพื่อนำค่า  $E^t$  แต่ละตำแหน่งมาหาเฉลี่ย กำหนดให้  $Q$  คือ ความคลาดเคลื่อนเฉลี่ยจาก  $Q^t$  ซึ่งเกิดจากการ Projection ของ Kinect

### 2.3 การคัดแยกภาพวัตถุด้วยข้อมูลภาพระยะลึก

ข้อมูลระยะลึกสามารถช่วยเป็นตัวบ่งบอกตำแหน่งของวัตถุบนภาพได้ โดยอาศัยค่าความแตกต่างของระยะลึกที่เกิดขึ้นเป็นตัวช่วยในการคัดแยกวัตถุกับพื้นหลัง ซึ่งในงานวิจัยเรื่องการคัดแยกภาพวัตถุกับพื้นหลัง [6][10] จะใช้ K-means ในการแบ่งกลุ่มข้อมูลระยะลึกเพื่อค้นหาตำแหน่งตัววัตถุจากสมมติฐานที่ว่าตำแหน่งของวัตถุคือ กลุ่มของระยะลึกที่น้อยที่สุด แต่อย่างไรก็ตามข้อมูลระยะลึกที่ได้มาในแต่ละวิธีการนั้นจะประสบกับปัญหาเรื่องความไม่สมบูรณ์ของข้อมูลระยะลึก ในงานวิจัยของ [12][13] ได้ใช้วิธีการที่เรียกว่า Bilateral filter ในการปรับปรุงข้อมูลระยะลึกที่ขาดหายไปโดยอาศัยข้อมูลภาพสีและข้อมูลภาพระยะลึก ซึ่งวิธีการดังกล่าวให้ผลดีกับภาพวัตถุและภาพพื้นหลังที่มีสีแตกต่างและไม่หลากหลาย เพราะถ้าวัตถุและพื้นหลังมีสีใกล้เคียงทำให้ข้อมูลระยะลึกคำนวณได้มีความใกล้เคียงกัน ทำให้ไม่สามารถแยกความแตกต่างระหว่างวัตถุกับพื้นหลัง ซึ่งงานวิจัยฉบับนี้เสนอแนวทางการปรับปรุงข้อมูลภาพระยะลึกที่ได้จาก Kinect โดยใช้อาศัยหลักการตรวจสอบตามแนวแสงซึ่งกล่าวไว้ในบทถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับเพื่อการเรียนการสอนเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.4. การคำนวณภาพแมทจากข้อมูลภาพระยะลึก

ปัญหาหลักที่ส่งผลต่อการแยกภาพวัตถุกับภาพพื้นหลังเพื่อนำมาซ้อนกับภาพพื้นหลังใหม่คือการที่ภาพวัตถุและภาพพื้นหลังบริเวณดังกล่าวมีสีพื้นที่ใกล้เคียงกันทำให้ยากต่อการกำหนดขอบของบริเวณดังกล่าวส่วนใดเป็นของภาพพื้นหลังหรือของวัตถุ ตัวอย่างเช่น ภาพสีพื้นหลังบริเวณดังกล่าวมีค่าสี  $R=200$ ,  $G=180$ ,  $B=190$  และค่าสีของวัตถุคือ  $R=195$ ,  $G=185$ ,  $B=192$  ซึ่งเห็นได้ว่าค่าของภาพพื้นหลังและภาพวัตถุบริเวณดังกล่าวมีความใกล้เคียงกันมาก ทำให้ส่งผลต่อความถูกต้องในการทำ matting และการคำนวณภาพแมทหรือค่า  $\alpha$  จากปัญหาดังกล่าวข้อมูลระยะลึกถูกเพิ่มเข้ามาเพื่อช่วยเพิ่มความถูกต้องของการคำนวณภาพแมท โดยข้อมูลระยะลึกสามารถช่วยในการคัดกรองภาพพื้นหลังและวัตถุโดยอาศัยจากการที่วัตถุโดยทั่วไปมีระยะห่างจากภาพพื้นหลัง เช่นในตัวอย่างที่กล่าวมาในข้างต้นที่ ถ้าภาพพื้นหลังมีค่าระยะลึกคือ 3.5 เมตร และภาพวัตถุอยู่ที่ 1.2 เมตร เห็นได้ว่าแม้ว่าสีพื้นหลังกับวัตถุมีความใกล้เคียงกันแต่พื้นหลังกับภาพวัตถุนั้นมีค่าระยะลึกที่แตกต่างกันถึง 2.3 เมตร ทำให้สามารถแยกบริเวณดังกล่าวได้ สำหรับเทคนิคการคำนวณภาพแมทโดยใช้ข้อมูลระยะลึกซึ่งนิยมใช้ในปัจจุบันมีด้วยกัน 2 วิธีคือ

1. การใช้ภาพระยะลึกกับเทคนิค Bayesian-based matting
2. การใช้ภาพระยะลึกกับเทคนิค Closed-form matting

### 2.4.1 การใช้ภาพระยะลึกกับเทคนิค Bayesian-based matting

Bayesian-based matting เป็นเทคนิคการคำนวณภาพแมทดิจิทัล ที่ได้มีการนำเสนอมาตั้งแต่ในช่วงแรก ๆ ที่ได้มีการวิจัยและพัฒนาเทคนิคในกลุ่มดังกล่าวขึ้น หลักการคือการประยุกต์ใช้วิธีการทางสถิติของเบย์ (Bayesian estimation techniques) ในการคำนวณค่า  $\alpha$  (alpha) ในภาพแมท โดยการหาคำตอบของสมการต่อไปนี้

$$\begin{aligned}
 (F, B, \alpha) &= \arg \max_{F, B, \alpha} P(F, B, \alpha | C) \\
 &= \arg \max_{F, B, \alpha} P(C|F, B, \alpha)P(F)P(B)P(\alpha)/P(C) \\
 &= \arg \max_{F, B, \alpha} L(C|F, B, \alpha) + L(F) + L(B) + L(\alpha)
 \end{aligned}
 \tag{2.6}$$

จากสมการข้างต้น  $C$ ,  $F$ , และ  $B$  เป็นเวกเตอร์ขนาด  $3 \times 1$  ที่แทนข้อมูลค่าสี แดง เขียว และ น้ำเงิน ของ ค่าสี ณ จุดใด ๆ ( $C$ ) ที่เกิดจากการผสมกันระหว่างสีด้านหน้า ( $F$  - Foreground color) และสีพื้นหลัง ( $B$  - Background color) ซึ่งการหาคำตอบจากสมการข้างต้น กระทำได้

โดยการแก้สมการหาค่า  $F$ , และ  $B$  สลับไปมากับการหาค่า  $\alpha$  ตามสมการต่อไปนี้ [2]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{bmatrix} \Sigma_F^{-1} + \alpha^2/\sigma_c^2 & \alpha(1-\alpha)/\sigma_c^2 \\ \alpha(1-\alpha)/\sigma_c^2 & \Sigma_F^{-1} + \alpha(1-\alpha)^2/\sigma_c^2 \end{bmatrix} \begin{bmatrix} [F] \\ [B] \end{bmatrix} = \begin{bmatrix} \Sigma_F^{-1}\bar{F} + \alpha/\sigma_c^2 \\ \Sigma_B^{-1}\bar{B} + \alpha(1-\alpha)/\sigma_c^2 \end{bmatrix} \quad (2.7)$$

และ

$$\alpha = \frac{(C - B)(F - B)}{\|F - B\|^2} \quad (2.8)$$

ในการแก้สมการแรกข้างต้น อาศัยค่าเริ่มต้นของ  $\alpha$  ที่กำหนดโดยผู้ใช้หรืออาจเป็นค่าที่ได้จากการคำนวณในรอบที่แล้ว (Previous iteration) มาใช้ โดยปกติการคำนวณกลับไปมาข้างต้น กระทำจนกระทั่งการเปลี่ยนแปลงของค่า  $\max_{F,B,\alpha} P(F, B, \alpha | C)$  ที่ได้จากสมการ (2.6) น้อยกว่าระดับค่าที่ตั้งไว้

แนวทางโดยทั่วไปที่ได้มีการเสนอเพื่อใช้ภาพระยะลึกที่ได้จากกล้องถ่ายภาพประเภท RGB-D ให้เป็นประโยชน์ต่อการทำภาพแมทด้วยวิธีการดังกล่าวข้างต้น คือการเพิ่มข้อมูลระยะลึกเข้าไปในเวกเตอร์ C, F, และ B เป็นค่าที่ 4 ต่อจากค่าสี R, G, และ B จากนั้นจึงใช้หลักการคำนวณในลักษณะเดียวกับวิธีการดั้งเดิมข้างต้น โดยไม่ต้องปรับแก้สมการใด ๆ อย่งไรก็ดีในทางปฏิบัติ การดัดแปลงวิธีการเดิมในลักษณะดังกล่าวให้ใช้ประโยชน์จากข้อมูลระยะลึกได้นั้น มีข้อเสียที่สำคัญอย่างน้อย 2 ประการคือ

1. ข้อมูลระยะลึกกับข้อมูลค่าสีเป็นข้อมูลคนละประเภทที่มีสเกลที่แตกต่างกัน ทำให้ยากต่อการกำหนดความสำคัญหรือให้น้ำหนักที่เหมาะสมระหว่างข้อมูล 2 ประเภทดังกล่าว
2. ในความเป็นจริง ข้อมูลระยะลึกไม่ได้มีความสัมพันธ์โดยตรงกับค่าแมท ( $\alpha$ ) เนื่องจากค่า  $\alpha$  มีความสัมพันธ์กันโดยตรงกับส่วนที่เป็นค่าสีในเวกเตอร์ C, F, และ B ไม่ใช่ค่าระยะลึก และในขณะที่ค่าสีใน C เกิดจากการผสมกันระหว่างค่าสีของ F และ B ในขณะที่ค่าระยะลึกที่จุดเดียวกันนั้น ไม่ได้มีความสัมพันธ์ในทางกายภาพกับค่าระยะลึกใน F และ B แต่อย่างใด (และหากมี ก็ไม่ได้มีความสัมพันธ์โดยตรงกับค่า  $\alpha$  ที่ต้องการทราบ)

ด้วยเหตุผลดังกล่าวข้างต้น ทำให้จำเป็นต้องมีการกำหนดโมเดลที่เป็นสมมติฐานที่ใช้ในการระบุความสัมพันธ์ระหว่างค่าระยะลึกกับการคำนวณค่า  $\alpha$  ซึ่งโมเดลหรือสมมติฐานที่แตกต่างกัน นำไปสู่วิธีการที่แตกต่างกันด้วย ซึ่งได้กล่าวถึงแต่ละวิธีการและสมมติฐาน

เอกสารนี้เป็นความลับอันเป็นที่มาของวิธีการต่าง ๆ ดังกล่าว ดังต่อไปนี้ ภาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### ก) วิธีของ Wang [3]

โดยหลักการแล้ว หากวัตถุที่อยู่ด้านหน้าและพื้นหลังที่ต้องการแยกออกจากกัน อยู่ห่างกันมากพอตามแนวทิศทางการมองของกล้องที่ใช้บันทึกภาพ (แกน z) และข้อมูลระยะลึกที่ได้ มีความถูกต้องเพียงพอ เราสามารถแยกวัตถุด้านหน้าออกจากพื้นหลังได้ โดยทำการแบ่งกลุ่มของข้อมูลระยะลึกออกกลุ่ม ๆ โดยใช้เทคนิคการแบ่งกลุ่มข้อมูล (Clustering techniques) เช่น K-means เป็นต้น พิกเซลที่ค่าระยะลึกอยู่ในกลุ่มที่ใกล้กับกล้องมากกว่า จะเป็นพิกเซลที่ประกอบกันขึ้นเป็นภาพวัตถุด้านหน้า ส่วนพิกเซลอื่น ๆ ที่อยู่ในกลุ่มที่อยู่ห่างออกเป็นภาพที่เกิดจากวัตถุที่เป็นส่วนหนึ่งของพื้นหลัง วิธีการดังกล่าวซึ่งจัดเป็นลักษณะของการแบ่งแยกวัตถุในภาพแบบ Hard segmentation (คือจุดใด ๆ ในภาพ เป็นสมาชิกในกลุ่มวัตถุกลุ่มใดกลุ่มหนึ่งเท่านั้น) ไม่เหมาะกับการใช้งานในลักษณะของภาพที่จำเป็นต้องใช้เทคนิคการคำนวณภาพแบบดิจิทัลด้วยเหตุผลหลัก 2 ประการ คือ

1. กรณีขอบหรือรอยต่อระหว่างวัตถุด้านหน้าและพื้นหลังที่ไม่ชัดเจน เช่น บริเวณเส้นผมหรือขน หรือวัตถุที่โปร่งแสงบางส่วน พิกเซลในบริเวณดังกล่าว ไม่สามารถระบุว่าเป็นของวัตถุด้านหน้า หรือพื้นหลัง หากต้องการแยกส่วนของวัตถุด้านหน้าออกจากสีของพื้นหลังที่ผสมในพิกเซลเหล่านั้น จำเป็นต้องคำนวณค่า  $\alpha$  และสีที่แท้จริงของวัตถุด้านหน้าที่ตรงกับพิกเซลดังกล่าว แต่เนื่องจากค่าระยะลึกไม่ได้มีความสัมพันธ์โดยตรงกับค่า  $\alpha$  เราจึงไม่สามารถคำนวณค่า  $\alpha$  จากค่าระยะลึกได้
2. ค่าระยะลึกที่วัดได้ด้วยกล้อง RGB-D มีความคลาดเคลื่อนจากความเป็นจริง มากน้อยแตกต่างกันออกไป โดยเฉพาะสำหรับ consumer-grade camera เช่น กล้อง Kinect มีความคลาดเคลื่อนที่สามารถสังเกตเห็นได้ด้วยตา เมื่อเปรียบเทียบขอบวัตถุที่ได้จากค่าระยะลึก กับขอบวัตถุที่แท้จริง ดังนั้น แม้แต่การแบ่งวัตถุด้านหน้าออกจากฉากหลัง กรณีที่เป็นวัตถุทึบแสง ก็ไม่สามารถใช้ข้อมูลระยะลึกในการแบ่งได้อย่างถูกต้องโดยลำพัง

ด้วยเหตุผลและข้อจำกัดดังกล่าวข้างต้น ทำให้แนวทางในการใช้ข้อมูลระยะลึกสำหรับการคำนวณภาพเมทโดยทั่วไป เป็นการใช้ข้อมูลดังกล่าว (โดยตรงหรือโดยอ้อมก็ตาม) ในการเสริมความถูกต้องให้กับการแบ่งขอบเขตภาพอย่างหยาบ (Coarse segmentation) เพื่อระบุส่วนของวัตถุด้านหน้า และส่วนของพื้นหลัง ที่ไม่อยู่ในบริเวณใกล้เคียงหรือตรงรอยต่อระหว่างวัตถุสองส่วนดังกล่าว (คือส่วน Foreground region และ Background region ในภาพไทรแมป) โดยวิธีทางตรงคือการใช้ข้อมูลระยะลึกในการคำนวณหรือสร้างภาพไทรแมป

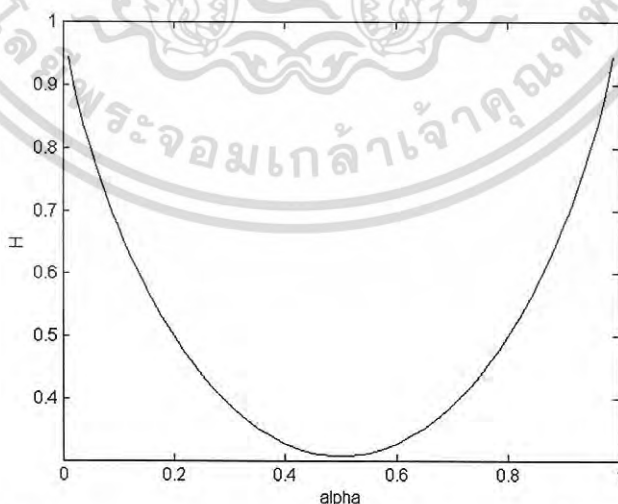
ใน [3] ใช้ข้อมูลระยะลึกช่วยในการคำนวณภาพเมท โดยอาศัยสมมติฐานที่ว่ามีความเชื่อมโยงสัมพันธ์กันระหว่างค่า  $\alpha$  กับความถูกต้องของกลุ่มข้อมูลระยะลึก กล่าวคือ สำหรับไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งยังมีให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พิกเซลที่มีค่า  $\alpha$  ที่สูงมาก หรือต่ำมาก (ใกล้ 1 หรือ 0) การแบ่งส่วนภาพโดยอาศัยข้อมูลระยะลึกมีความถูกต้องสูง เนื่องจากวัตถุในบริเวณดังกล่าวทึบแสงและไม่ได้อยู่บริเวณขอบหรือรอยต่อของวัตถุ ที่การวัดค่าระยะลึกมีโอกาสคลาดเคลื่อนสูง (ไม่ว่าจะใช้กล้องที่อาศัยการวัดค่าระยะลึกด้วยวิธีหรือเทคโนโลยีใดก็ตาม เช่น Structured light stereoscopic sensor หรือ Time-of-flight sensor) ในทางกลับกันค่า  $\alpha$  ที่มีค่ากลาง ๆ (ใกล้เคียง 0.5) มักเกิดจากบริเวณผิววัตถุที่โปร่งแสงบางส่วน หรือบริเวณขอบวัตถุ ที่การวัดค่าระยะลึกมักมีความคลาดเคลื่อนสูงกว่ากรณีแรก จากสมมติฐานดังกล่าว ข้อมูลภาพในแต่ละพิกเซล ที่ประกอบไปด้วยค่าสีและค่าระยะลึก ก่อนถูกนำไปใช้ในการแก้สมการ (3.3) ส่วนของข้อมูลที่เป็นระยะลึกถูกนำมาคูณกับค่าน้ำหนักที่สะท้อนความน่าเชื่อถือของค่าระยะลึกนั้น ๆ ตามสมการต่อไปนี้

$$H(\alpha) = 1 + \alpha \log(\alpha) + (1 - \alpha) \log(1 - \alpha) \tag{2.10}$$

โดยที่  $\alpha$  ในสมการข้างต้น คือค่า  $\alpha$  ที่จุดหรือพิกเซลเดียวกับค่าระยะลึกที่ต้องการกำหนดค่าน้ำหนัก

รูปที่ 2.10 แสดงกราฟของค่าน้ำหนัก  $H(\alpha)$  ที่ค่า  $\alpha$  ต่าง ๆ สังเกตว่าค่าน้ำหนักสูงเมื่อ  $\alpha$  ใกล้ 0 หรือ 1 และต่ำสุดเมื่อ  $\alpha$  เข้าหาค่า 0.5 โดยการใช้ค่าน้ำหนักดังกล่าว ทำให้ค่าระยะลึกในส่วนที่มีความถูกต้องต่ำ เมื่อนำไปคูณกับค่าน้ำหนักแล้ว มีผลต่อการคำนวณค่า  $\alpha$  น้อยตามไปด้วย ทั้งนี้ เพื่อไม่ให้ความผิดพลาดที่มีอยู่ในค่าระยะลึกส่งผลในทางลบต่อการคำนวณภาพ



รูปที่ 2.10 แสดงกราฟของ  $H(\alpha)$  เมื่อเทียบกับค่า  $\alpha$  ที่แตกต่างกัน

ปัญหาที่สำคัญของวิธีการนี้ ในการนำไปใช้งาน คือ สมการที่ 2.10 ที่ใช้ในการคำนวณค่าน้ำหนัก ต้องอาศัยค่า  $\alpha$  ซึ่งสำหรับพิกเซลบริเวณขอบวัตถุ (ซึ่งอยู่ในบริเวณไม่ชัดเจนของภาพไทรแมบ) ไม่ทราบมาก่อน ดังนั้นในทางปฏิบัติ การคำนวณในรอบแรก (1st iteration) สำหรับพิกเซลเหล่านี้จำเป็นต้องอาศัยค่าเริ่มต้น เช่น  $\alpha = 0.5$  ซึ่งมีผลทำให้ผลการคำนวณในรอบแรกนี้ ค่าระยะลึกในบริเวณดังกล่าวจะมีผลน้อยต่อผลลัพธ์ที่ได้ หลังจากนั้นในรอบการประมวลผลถัด ๆ มา สามารถใช้ค่า  $\alpha$  ที่ได้จากการคำนวณในรอบก่อนหน้านั้น สำหรับใช้แทนลงในสมการ 2.10 ได้

#### ข) วิธีการของ Lu [13]

เช่นเดียวกับวิธีการของ Wang วิธีการที่ 2 นี้ อาศัยค่าระยะลึกในการช่วยแบ่งขอบเขตระหว่างวัตถุด้านหน้ากับพื้นหลังร่วมกับข้อมูลค่าสี โดยอาศัยสมมติฐานต่อไปนี้

1. บริเวณพิกเซลในภาพที่ค่าสี Foreground (F) มีค่าใกล้เคียงกับค่าสีที่วัดได้ (C) มีความเป็นไปได้สูงที่เป็นบริเวณของวัตถุด้านหน้า ในทำนองเดียวกัน บริเวณพิกเซลในภาพที่ค่าสี Background (B) มีค่าใกล้เคียงกับค่าสีที่วัดได้ (C) มีความเป็นไปได้สูงที่เป็นบริเวณของวัตถุพื้นหลัง (หมายถึงทั้งสองกรณีมีความเป็นไปได้น้อยที่เป็นบริเวณรอยต่อระหว่างวัตถุด้านหน้าและพื้นหลัง) ดังนั้น ค่าระยะลึกที่อ่านได้จากบริเวณดังกล่าว น่าจะมีความถูกต้องหรือน่าเชื่อถือสูง จึงควรให้น้ำหนักมากกับค่าระยะลึกของพิกเซลในบริเวณดังกล่าว
2. บริเวณพิกเซล ที่ค่าสี Foreground (F) และค่าสี Background (B) มีค่าใกล้เคียงกัน ควรให้น้ำหนักกับค่าระยะลึกของพิกเซลในบริเวณเดียวกันนั้นให้มาก เนื่องจากในกรณีนี้ ข้อมูลค่าสีมีความน่าเชื่อถือหรือความแม่นยำต่ำในการแยกหรือจำแนกว่าบริเวณดังกล่าวเป็นวัตถุด้านหน้าหรือด้านหลัง และการคำนวณค่า  $\alpha$  ก็มีแนวโน้มความคลาดเคลื่อนสูง การใช้ข้อมูลระยะลึกในบริเวณดังกล่าวจึงควรให้น้ำหนักมาก

หากพิจารณาสมมติฐานข้อแรก พบว่ามีความเหมือนหรือสอดคล้องกับสมมติฐานที่ใช้ในวิธีการแรกของ Wang กล่าวคือบริเวณของภาพที่เป็นส่วนของวัตถุด้านหน้าหรือพื้นหลังที่ทึบแสง และไม่ใช่อบริเวณรอยต่อ ค่าระยะลึกมีความถูกต้องสูงและควรให้น้ำหนักมาก หากแต่การระบุบริเวณดังกล่าว ใช้วิธีที่แตกต่างกัน (แต่สัมพันธ์กัน) กล่าวคือ ในขณะที่วิธีการของ Wang ใช้ค่า  $\alpha$  มาเป็นตัววัด วิธีการของ Lu ใช้ผลต่างหรือความคล้ายคลึงกันระหว่าง  $F(B)$  กับ  $C$  แทน ซึ่งสัมพันธ์กับค่า  $\alpha$  ในแง่ที่ว่า ค่า  $F(B)$  ที่ใกล้เคียงกับ  $C$  มีผลทำให้ค่า  $\alpha$  เข้าใกล้ 1 (0 ในกรณีของ B) ซึ่งทั้งสองวิธี ให้น้ำหนักกับค่าระยะลึกมากกว่าบริเวณอื่น ๆ ที่ค่า  $F(B)$  แตกต่างจาก  $C$  ซึ่งในกรณีหลัง ค่าของ  $\alpha$  เข้าใกล้ 0.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมมติฐานข้อ 2 ที่ใช้ใน [8] เป็นส่วนที่เพิ่มเติมเข้ามา ซึ่งไม่มีปรากฏใน [3] และจากสมมติฐานทั้ง 2 ข้อดังกล่าว Lu เสนอการกำหนดค่าน้ำหนักที่ให้กับข้อมูลระยะลึกโดยใช้สมการต่อไปนี้

$$w_d = \exp\left(-\frac{w_F \cdot w_B}{\sigma_d^2}\right) \quad (2.11)$$

โดยที่  $w_F$  (และ  $w_B$ ) คำนวณจากสมการต่อไปนี้

$$w_F = 1 - \exp(-\|F^* - C^*\|^2) \quad (2.12)$$

$$w_B = 1 - \exp(-\|B^* - C^*\|^2) \quad (2.13)$$

จากสมการข้างต้น ในที่นี้  $F^*$  และเวกเตอร์อื่น ๆ ในสมการ คือเวกเตอร์ที่มีเฉพาะข้อมูลสีเท่านั้น ไม่รวมข้อมูลระยะลึก ใน [13] ได้เสนอการกำหนดค่า  $\sigma_d = 0.1$  และเช่นเดียวกับวิธีการของ Wang สำหรับพิกเซลบริเวณไม่ชัดเจน ค่า  $F$  ( $B$ ) เป็นค่าที่ไม่ทราบและต้องการคำนวณหา ซึ่งวิธีการของ Lu หลีกเลี่ยงปัญหาดังกล่าว โดยในขั้นตอนของการคำนวณค่า  $F$  ( $B$ ) ในสมการที่ (2.12) ดำเนินการ โดยไม่มีการคูณค่าน้ำหนักให้กับข้อมูลระยะลึกแต่อย่างใด (คือคำนวณเหมือนวิธีการ Bayesian-based matting ยกเว้นแต่เพิ่มค่าระยะลึกเข้ามา นอกเหนือจากค่าสี) ต่อเมื่อทำการคำนวณค่า  $\alpha$  ด้วยสมการ (2.13) จึงคำนวณค่าน้ำหนักด้วยสมการข้างต้น แล้วนำไปคูณเข้ากับข้อมูลระยะลึก (ค่าที่ 4 ใน  $F$ ,  $B$ , และ  $C$ ) ก่อนที่นำไปใช้คำนวณหาค่า  $\alpha$  ตามสมการที่ (2.12) ต่อไป

#### 2.4.2 การใช้ภาพระยะลึกกับเทคนิค Closed-form matting

Closed-form matting เป็นเทคนิคการคำนวณภาพแมทอีกวิธีการหนึ่งที่มีการอ้างอิงบ่อยครั้ง ส่วนหนึ่งมาจากการที่เทคนิคดังกล่าว มีการคำนวณและการใช้งานที่ไม่ซับซ้อน และได้ผลดี โดยเฉพาะอย่างยิ่งสำหรับภาพที่ไม่มีลวดลาย ทั้งในส่วนวัตถุด้านหน้าและพื้นหลัง เนื่องจากเทคนิคดังกล่าวพัฒนาอยู่บนสมมติฐานที่ว่า ค่าสีไม่เปลี่ยนแปลงในบริเวณใกล้เคียงกันบนภาพ จากสมมติฐานดังกล่าว ค่า  $\alpha$  สามารถคำนวณได้โดยการแก้สมการต่อไปนี้ [4]

$$\alpha = \arg \min \alpha^T L \alpha + \lambda (\alpha^T - b_s^T) D_s (\alpha - b_s) \quad (2.14)$$

โดยที่  $L$  คือเมทริกซ์ขนาด  $N \times N$  โดยมีค่า ณ ตำแหน่งที่  $(i, j)^{th}$  เท่ากับ

$$L_{i,j} = \sum_{k|(i,j) \in w_k} \left\{ \delta_{i,k} - \frac{1}{|w_k|} (1 + (I_i - \mu_k) \left( R_k + \frac{\epsilon}{|w_k|} I_3 \right)^{-1} (I_j - \mu_k)) \right\} \quad (2.15)$$

จากสมการข้างต้น  $R_k$  คือ Covariance matrix สำหรับจุดหรือพิกเซลในภาพที่อยู่ภายในขอบเขตตามที่กำหนดในเซต  $w_k$  ซึ่งกำหนดขอบเขตโดยใช้กรอบหน้าต่างต่างเลื่อนไปตามจุดต่าง ๆ บนภาพต้นฉบับ (หากไม่นับจุดบริเวณขอบของภาพ จำนวนเซต  $w_k$  เท่ากับจำนวนพิกเซล  $N$  ที่ต้องการคำนวณหาค่าเมท) ในสมการเดียวกัน ค่า  $\epsilon$  คือค่าคงที่ที่เอาไว้ควบคุมหรือกำหนดอัตราการเปลี่ยนแปลงของค่า  $\alpha$  ในบริเวณที่อยู่ข้างเคียงกัน (Smoothing factor) ตลอดจนใช้เพื่อให้การประมวลผลมีเสถียรภาพ ส่วน  $I_3$  ในสมการเดียวกัน คือเมทริกซ์เอกลักษณ์ ขนาด  $3 \times 3$  ในสมการ 2.14 ข้างต้น  $\alpha$  คือเวกเตอร์ขนาด  $N \times 1$  ที่ประกอบไปด้วยค่า  $\alpha$  ของจุดต่าง ๆ ในภาพ

หลังจากคำนวณค่า  $\alpha$  ได้แล้ว ค่าสีของวัตถุด้านหน้า และสีพื้นหลัง สามารถหาได้โดยการแก้สมการต่อไปนี้ [4]

$$(F^c, B^c) = \arg \min \sum_{i \in I} \left( \alpha_i F_i^c + (1 - \alpha_i) B_i^c - I_i^c \right)^2 + |\nabla_x \alpha_i| \left( (\nabla_x F_i^c)^2 + (\nabla_x B_i^c)^2 \right) + |\nabla_y \alpha_i| \left( (\nabla_y F_i^c)^2 + (\nabla_y B_i^c)^2 \right) \quad (2.16)$$

จากสมการข้างต้น  $F^c$  คือเวกเตอร์ที่ประกอบไปด้วยค่าสี Foreground ที่จุดที่  $i$  ต่าง ๆ ( $F_i^c$ ) โดย  $c$  แทนชานเนลสีใด ๆ (Red, Green, Blue)  $B^c$  ในสมการข้างต้นมีความหมายในการทำงานเดียวกัน นอกจากนี้ สัญลักษณ์  $\nabla_x(\cdot)$  แทน Derivative operator ตามแนวแกน  $x$  ส่วน  $\nabla_y(\cdot)$  แทน Derivative operator ตามแนวแกน  $y$  เพื่อให้ได้ค่าสีที่ครบถ้วน สมการข้างต้น ต้องกระทำ 3 ครั้งกับชานเนลสีที่แตกต่างกัน (Red, Green, Blue)

สำหรับแนวทางการใช้ค่าระยะสีกร่วมกับวิธีการดังกล่าว ในที่นี้ กล่าวถึงอยู่ 2 วิธีดังต่อไปนี้

#### ก) วิธีการของ Pitie [15]

วิธีการนี้ใช้แนวทางเดียวกับการปรับใช้ค่าระยะสีกับ Bayesian-based matting [2] กล่าวคือ เพิ่มข้อมูลระยะสีเข้าไปในสมการเพื่อคำนวณค่า  $\alpha$  ซึ่งในกรณีนี้ คือการเพิ่มค่าที่ 4 เข้าไปในเวกเตอร์  $I_i$  ที่เดิมมีขนาด  $3 \times 1$ . ทำให้สมการ (2.15) เปลี่ยนเป็น

$$L_{i,j} = \sum_{k|(i,j) \in w_k} \left\{ \delta_{i,k} - \frac{1}{|w_k|} (1 + (I_i - \mu_k)(R_k + E_k)^{-1}(I_j - \mu_k)) \right\} \quad (2.17)$$

โดยที่

$$E_k = \frac{1}{|w_k|} \begin{bmatrix} \epsilon & 0 & 0 & 0 \\ 0 & \epsilon & 0 & 0 \\ 0 & 0 & \epsilon & 0 \\ 0 & 0 & 0 & \epsilon_D \end{bmatrix} \quad (2.18)$$

และเช่นเดียวกับวิธีการที่ผ่านมา คือจำเป็นต้องมีการกำหนดค่าน้ำหนักให้กับข้อมูลค่าระยะลึก โดย [5] เสนอการกำหนดค่าน้ำหนักด้วยการปรับเปลี่ยนค่า  $\epsilon_D$  ในสมการข้างต้น โดยค่ามากมีผลทำให้น้ำหนักของข้อมูลระยะลึกมีผลต่อการคำนวณน้อย และในทางกลับกัน หาก  $\epsilon_D$  มีค่าน้อย ทำให้อำนาจของน้ำหนักต่อการคำนวณมาก อย่างไรก็ตาม [5] ไม่ได้กำหนดหรือแนะนำวิธีในการคำนวณหรือหาค่า  $\epsilon_D$  ที่เหมาะสมแต่อย่างใด

#### ข) วิธีการของ Sun [7]

แทนการเพิ่มข้อมูลระยะลึกเข้าไปในสมการเพื่อการคำนวณ เช่น ในกรณีของ [5] ในบทความ [7] อาศัยค่าข้อมูลระยะลึกในการปรับค่าพารามิเตอร์  $\epsilon$  แทน โดยกำหนดให้ค่า  $\epsilon$  ไม่คงที่ตลอดทั้งภาพ สามารถแตกต่างกันได้ในแต่ละพิกเซล โดยมีค่าขึ้นอยู่กับระดับการผันแปรของค่าระยะลึกในแต่ละบริเวณของภาพที่ทำการคำนวณ ตามสมการต่อไปนี้

$$\epsilon_i = k_1 e^{-k_2 \cdot \text{var}(d_i)} \quad (2.19)$$

โดยที่  $\text{var}(d_i)$  คือค่าความผันแปร (Variance) ของค่าระยะลึก ณ พิกเซลที่อยู่รอบ ๆ จุด  $i$  ค่า ระยะลึกดังกล่าวถูกปรับให้มีค่าอยู่ระหว่าง 0 ถึง 1 โดยในบทความ [7] กำหนดให้ค่า  $k_1 = 0.1$ , และ  $k_2 = 10^6$

จากสมการข้างต้น แนวคิดสมมติฐานอันเป็นที่มาของสมการคือ หากค่าความผันแปรของระยะลึกสูง มีแนวโน้มว่าบริเวณดังกล่าวเป็นบริเวณขอบวัตถุ ดังนั้น จึงกำหนดค่า  $\epsilon_i$  ให้ต่ำ และหากค่าความผันแปรของระยะลึกต่ำ ค่า  $\epsilon_i$  สูง ซึ่งค่า  $\epsilon_i$  เป็นตัวควบคุมระดับความผันแปรของค่า  $\alpha_i$  ที่คำนวณได้ หาก  $\epsilon_i$  ต่ำทำให้ค่า  $\alpha_i$  ที่คำนวณได้มีการเปลี่ยนแปลงได้มากในระยะใกล้ ๆ แต่หาก  $\epsilon_i$  มีค่ามาก ทำให้  $\alpha_i$  ที่คำนวณได้มีแนวโน้มค่อนข้างคงที่หรือเปลี่ยนแปลงน้อย หรืออีกนัยหนึ่ง  $\epsilon_i$  เป็นพารามิเตอร์ที่มีลักษณะที่เรียกว่าเป็น Smoothing

เอกสารนี้เป็นเอกสารลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

factor ความสัมพันธ์ข้างต้น ถูกกำหนดจากสมมติฐานที่ว่า ความผันแปรของระยะลึกในบริเวณ Foreground หรือ Background ก่อนข้างต่ำ เมื่อเปรียบเทียบกับความผันแปรในบริเวณของวัตถุ เช่นเดียวกันกับการผันแปรของค่า  $\alpha_i$  ในบริเวณที่เป็น Foreground หรือ Background ที่ค่อนข้างต่ำหรือคงที่ ในขณะที่ค่า  $\alpha_i$  ในบริเวณของวัตถุมีแนวโน้มที่สูง นำไปสู่ข้อสรุปว่าความผันแปรของค่าระยะลึกและค่า  $\alpha_i$  มีแนวโน้มไปในทิศทางเดียวกัน อันเป็นพื้นฐานแนวความคิดของการกำหนดค่า  $\epsilon_i$  ตามสมการข้างต้น อนึ่งพึงสังเกตว่าสมมติฐานดังกล่าวมีความคล้ายคลึงสอดคล้องกับที่ใช้ในวิธีการอื่น ๆ ดังที่ได้กล่าวมาแล้วข้างต้น

## 2.5 การสร้างแผนภาพไทรแมบโดยอัตโนมัติจากข้อมูลระยะลึก

สำหรับภาพไทรแมบ คือ ภาพที่เป็นตัวบอกรอบเขตของภาพวัตถุ, ภาพพื้นหลัง และบริเวณที่ไม่ชัดเจน เบื้องต้นก่อนนำไปคำนวณเพื่อสร้างภาพแมท ดังนั้นภาพไทรแมบจึงส่งผลกระทบต่อประสิทธิภาพและคุณภาพในการซ้อนภาพ เพราะการกำหนดขอบเขตของวัตถุที่ผิดพลาดส่งผลผิดพลาดในการซ้อนภาพเช่น ส่วนของภาพวัตถุถูกกำหนดให้เป็นพื้นหลัง เป็นต้น ด้วยเหตุนี้ทำให้การสร้างภาพไทรแมบที่ถูกต้องเหมาะสมจึงเป็นความสำคัญ ทำให้โดยปกติการสร้างภาพไทรแมบที่ถูกกำหนดขอบเขตจากผู้ใช้งาน ด้วยเหตุนี้เพื่อสร้างภาพไทรแมบที่เหมาะสมแบบอัตโนมัติข้อมูลระยะลึกจึงถูกนำมาใช้ ดังเช่นงานวิจัยของ [7] ที่ใช้ภาพข้อมูลระยะลึกที่ได้จากภาพสองมุมมองหรือกล้อง สเตอริโอ (Stereo camera) มาสร้างภาพไทรแมบ จากการคัดแยกด้วยวิธีการ K-means เพื่อหาขอบเขตของภาพวัตถุกับพื้นหลังเบื้องต้นออกเป็น 2 กลุ่ม และใช้ระยะห่างในแต่ละกลุ่มที่เกินกว่า 95 % จากค่าเฉลี่ยเป็นตัวกำหนดบริเวณไม่ชัดเจน ซึ่งวิธีการดังกล่าวประสบปัญหาเมื่อพื้นหลังมีหลายระดับหรือการที่วัตถุกับพื้นหลังมีสีที่ไม่แตกต่างกันทำให้ไม่สามารถคำนวณระยะลึกได้เพราะ การคำนวณระยะลึกของกล้องสเตอริโอต้องหาจุดที่ภาพทั้งสองมีความเหมือนกันเพื่อใช้เป็นตำแหน่งอ้างอิงในการคำนวณระยะลึก หรืองานวิจัยของ [13] ที่ใช้ภาพจากกล้อง Kinect ที่ภาพการปรับปรุงด้วยวิธีการ Bilateral filter มาใช้ในการสร้างภาพไทรแมบ แต่จากที่กล่าวมาแล้ววิธีการปรับปรุงภาพด้วย Bilateral filter ได้ผลดีก็ต่อเมื่อสีของภาพวัตถุและพื้นหลังมีความแตกต่างและไม่มีความหลากหลายของสีหรือวัตถุ ซึ่งการที่ภาพข้อมูลระยะลึกมีความผิดพลาดส่งผลต่อการคัดแยกและการนำไปใช้ในการคำนวณภาพแมทอีกด้วย

### บทที่ 3

## การสร้างภาพไทรแมบแบบอัตโนมัติจากข้อมูลภาพสีและ ข้อมูลภาพระยะลึกร่วมกัน

ตามที่ได้กล่าวไปแล้วในบทที่ 1 ปัญหาสำคัญของการสร้างภาพไทรแมบ เพื่อนำไปมาใช้ในการซ้อนภาพ คือ การกำหนดขนาดบริเวณไม่ชัดเจนที่เหมาะสมกับภาพวัตถุเพราะขนาดของบริเวณไม่ชัดเจนส่งผลต่อความถูกต้อง , คุณภาพและประสิทธิภาพในการซ้อนภาพ โดยปกติทั่วไปจำเป็นต้องอาศัยผู้ใช้งานกำหนดบริเวณไม่ชัดเจน สำหรับบทที่ 3 นี้ผู้วิจัยได้นำเสนอแนวทางการสร้างภาพไทรแมบแบบอัตโนมัติ โดยอาศัยข้อมูลภาพสีและข้อมูลภาพระยะลึกที่ได้จากกล้อง Kinect

### 3.1 การสร้างภาพไทรแมบแบบอัตโนมัติจากข้อมูลภาพ RGBD

ในหัวข้อนี้ได้กล่าวถึงขั้นตอนหลักการสร้างภาพไทรแมบแบบอัตโนมัติจากข้อมูลภาพสีและข้อมูลภาพระยะลึกจากกล้องถ่ายภาพประเภท RGBD Camera (MS Kinect) ดังต่อไปนี้ (รูปที่ 3.1 ประกอบ)

#### 3.1.1 การปรับปรุงภาพระยะลึก (Depth image enhancement)

ขั้นตอนนี้เป็นการปรับปรุงแก้ไขความไม่สมบูรณ์ของภาพระยะลึกที่ได้จากกล้อง Kinect โดยบริเวณของภาพที่ข้อมูลดังกล่าวไม่สมบูรณ์หรือขาดหายไป ปรากฏค่าเป็น NaN (Not a number) ซึ่งเกิดขึ้นได้จาก 2 สาเหตุ คือ จุดดังกล่าวเกิดจากบริเวณผิววัตถุที่ไม่สามารถสะท้อนแสงอินฟราเรดที่ส่องออกมาจากตัวกล้องได้ และอีกหนึ่งสาเหตุเกิดจากการที่วัตถุที่อยู่ด้านหน้าบดบังแสงอินฟราเรด พื้นหลังบางส่วนที่ไม่ได้รับแสงปรากฏเป็นค่า NaN บนภาพระยะลึก ขั้นตอนนี้ทำการตรวจสอบว่าแต่ละจุดที่ข้อมูลระยะลึกขาดหายไปเกิดจากสาเหตุใด และทำการแก้ไขให้เหมาะสมกับแต่ละสาเหตุ

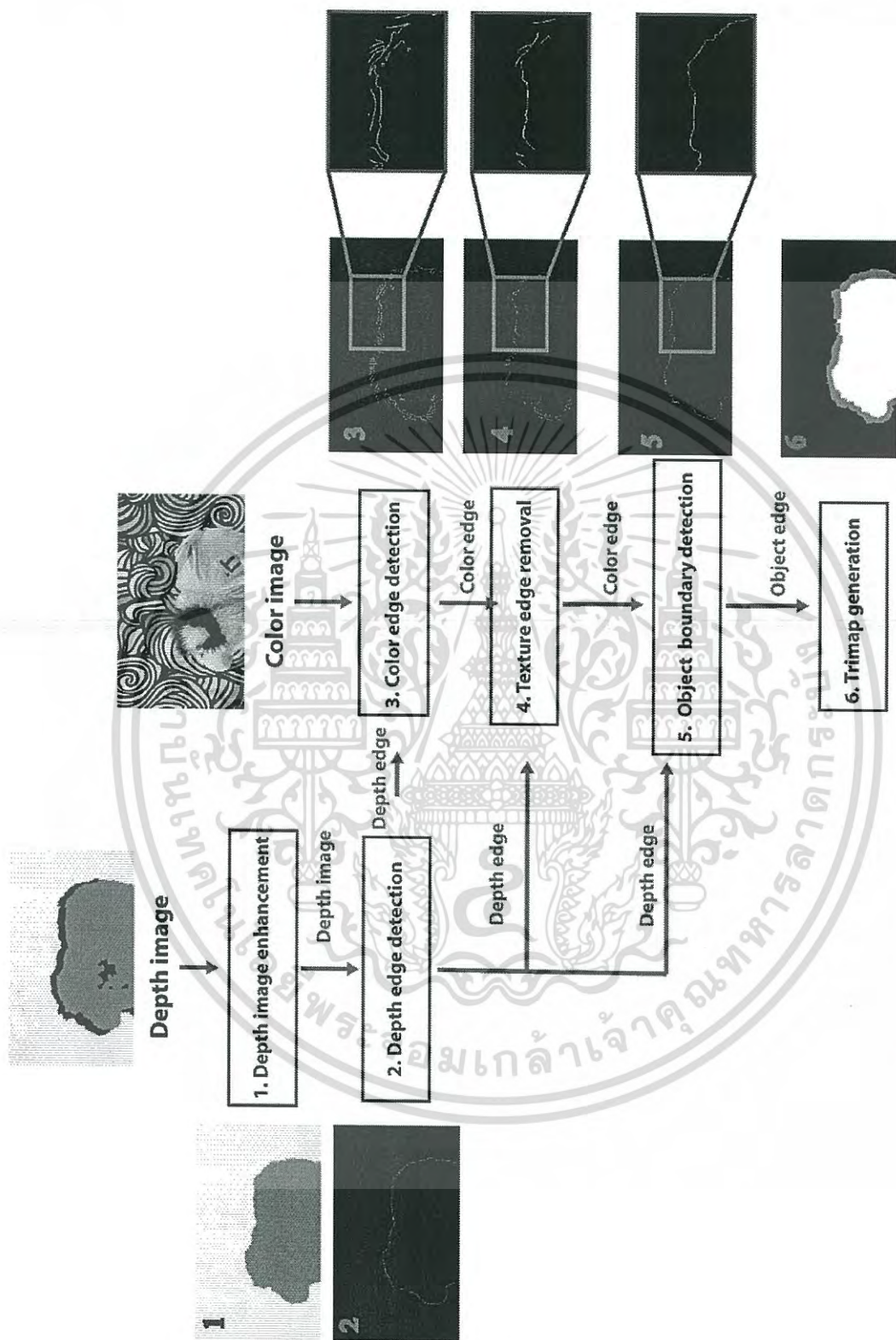


(ก)



(ข)

เอกสารนี้เป็นรูปที่ 3.2 (ก) ตัวอย่างระยะลึกก่อนการปรับปรุงแก้ไข (ข) ภาพระยะลึกหลังการปรับปรุงการคำนวณว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.1 ภาพแสดงลำดับขั้นตอนการสร้างภาพไตรแมปที่เหมาะสมจากข้อมูล RGBD

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.1.2 การตรวจสอบเส้นขอบภาพระยะลึก (Depth edge detection)

ขั้นตอนนี้เป็นการตรวจสอบเส้นขอบวัตถุ (Depth edge) จากข้อมูลภาพระยะลึกที่ผ่านการปรับปรุง เพื่อนำเส้นขอบวัตถุที่ได้ในการระบุขอบเขตตำแหน่งของวัตถุหรือคัดกรองข้อมูล



รูปที่ 3.3 (ก) ภาพระยะลึกหลังการปรับปรุงแก้ไขข้อมูลระยะลึก (ข) ภาพเส้นขอบระยะลึกที่ได้

### 3.1.3 การตรวจสอบเส้นขอบภาพสี (Color edge detection)

จุดประสงค์หลักของขั้นตอนนี้คือการค้นหาเส้นขอบสี (Color edge) ที่มีความชัดเจน โดยอาศัยข้อมูลเส้นขอบระยะลึกกับข้อมูลความคลาดเคลื่อนของตัวอุปกรณ์ Kinect [หัวข้อที่ 2.2.2] เป็นตัวช่วยในการคัดกรองเส้นขอบสี ของภาพวัตถุ เพื่อตัดส่วนของเส้นขอบสีที่ไม่เกี่ยวข้องกับขอบวัตถุออกไป

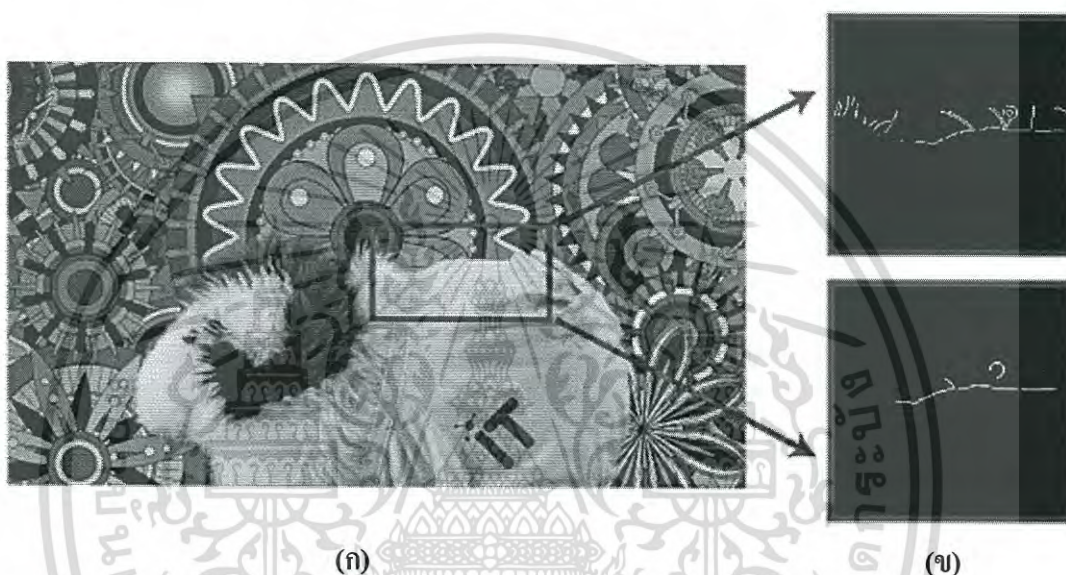


รูปที่ 3.4 (ก) ภาพตัวอย่างเส้นขอบสีก่อนการคัดกรองด้วยเส้นขอบระยะลึก (ข) ภาพตัวอย่างเส้นขอบสีจากการคัดกรองด้วยเส้นขอบระยะลึกกับข้อมูลความคลาดเคลื่อนของ Kinect

### 3.1.4 การลบเส้นขอบลวดลาย (Texture edge removal)

ขั้นตอนนี้เป็น การคัดกรองเส้นขอบที่เกิดจากลวดลายของพื้นหลังออกจากเส้นขอบที่เกิดจากรอยต่อระหว่างขอบวัตถุกับพื้นหลัง

การคัดแยกดังกล่าว กระทำเพื่อตัดเส้นขอบที่เกิดจากลวดลายพื้นหลังออกไป เพื่อให้เหลือเฉพาะเส้นขอบวัตถุ ที่ต้องการนำไปใช้ประกอบในการระบุขอบเขตของวัตถุ สำหรับคำนวณภาพไทรแมปในลำดับต่อไป



รูปที่ 3.5 (ก) ภาพสีเริ่มต้น (ข) ตัวอย่างภาพเส้นขอบสีก่อนทำการคัดแยกและหลังคัดแยก

### 3.1.5 การตรวจสอบขอบวัตถุ (Object boundary detection)

ขั้นตอนนี้เป็น การสร้างเส้นขอบวัตถุจากเส้นขอบสีและเส้นขอบระยะลึก เพื่อนำไปใช้สร้างภาพไทรแมป โดยทั่วไป ในการหาเส้นขอบวัตถุ สามารถแบ่งออกเป็นวิธีต่าง ๆ ได้ดังต่อไปนี้คือ

- 1) การอ้างอิงตำแหน่งภาพวัตถุจากข้อมูลเส้นขอบระยะลึก โดยข้อดีของวิธีการนี้คือสามารถทำได้สะดวกและรวดเร็ว เพราะใช้ความแตกต่างของค่าระยะลึกระหว่างภาพวัตถุกับพื้นหลังจึงไม่ต้องคำนึงถึงเรื่องสีของวัตถุและพื้นหลัง
- 2) การอ้างอิงตำแหน่งภาพวัตถุจากเส้นขอบสี โดยข้อดีของวิธีการนี้คือ ให้ความถูกต้องในการอ้างอิงสูงถ้าสีของวัตถุกับพื้นหลังมีความแตกต่างอย่างชัดเจน

การอ้างอิงตำแหน่งของภาพวัตถุทั้งสองวิธีนั้นพบปัญหาที่แตกต่างกันดังนี้

- 1) การอ้างอิงด้วยข้อมูลเส้นขอบระยะลึกพบปัญหาความคลาดเคลื่อนของข้อมูลภาพระยะลึกกับข้อมูลภาพสีที่ไม่ตรงกันเกิดจากตัวอุปกรณ์ ทำให้ต้องกำหนดขนาดบริเวณ ไม่ชัดเจนให้ครอบคลุมความคลาดเคลื่อนที่เกิดขึ้นของอุปกรณ์
- 2) การอ้างอิงด้วยข้อมูลเส้นขอบสีมีปัญหการตรวจสอบและความไม่สมบูรณ์ของเส้นขอบสี หากสีของภาพวัตถุกับพื้นหลังไม่แตกต่างกัน

วิธีการสร้างเส้นขอบวัตถุที่เสนอในที่นี้อาศัยข้อดีของเส้นขอบสีที่ให้ความถูกต้องสูงเมื่อวัตถุกับพื้นหลังมีสีแตกต่างกันและเส้นขอบระยะลึกที่มีความสมบูรณ์ของเส้นมาใช้ในการอ้างอิงตำแหน่งวัตถุร่วมกัน หรือกล่าวได้ว่า บริเวณเส้นขอบสีที่เป็นขอบวัตถุซึ่งมีความชัดเจนผ่านการตรวจสอบถูกเลือกในบริเวณดังกล่าว โดยที่บริเวณที่ไม่มีเส้นขอบสีหรือไม่ผ่านเกณฑ์ที่กำหนดเลือกใช้เส้นขอบระยะลึกทดแทน ทำให้ได้เส้นขอบที่ใช้ในการอ้างอิงตำแหน่งของวัตถุใหม่ที่อาศัยข้อดีของเส้นขอบสีที่มีความถูกต้องสูงและเส้นขอบระยะลึกที่มีความสมบูรณ์ของเส้นมาใช้ในการอ้างอิงตำแหน่งวัตถุ

### 3.1.6 การสร้างภาพไตรแมป (Trimap generation)

ขั้นตอนนี้เป็นกรสร้างภาพไตรแมปที่เหมาะสมจากเส้นขอบวัตถุที่ได้จากขั้นตอนที่แล้ว โดยทำการสร้างขอบเขตของบริเวณไม่ชัดเจนขึ้นจากเส้นขอบวัตถุ



(ก)



(ข)

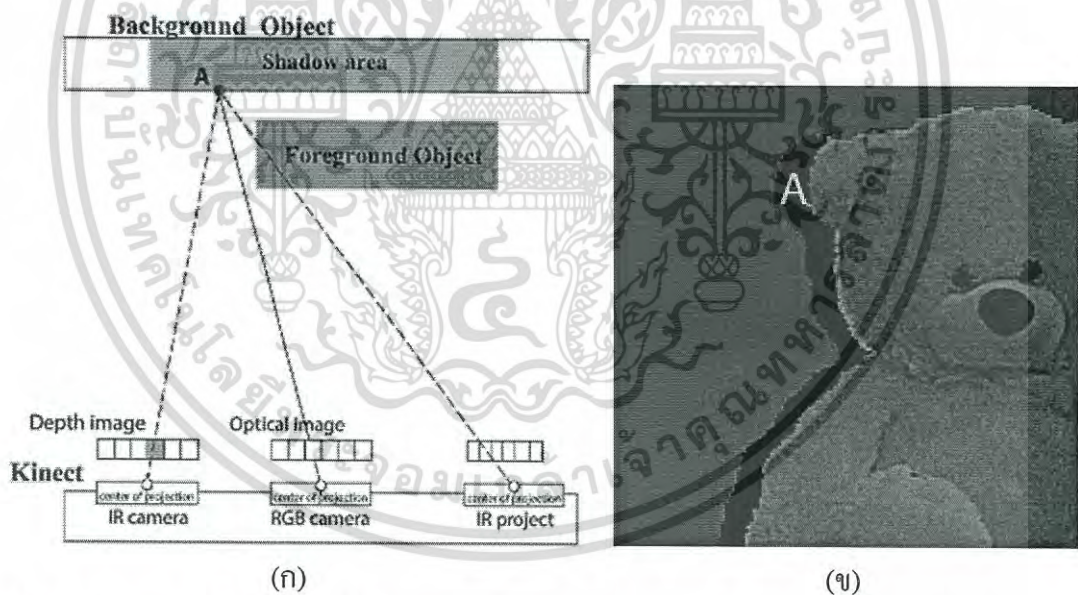
รูปที่ 3.6 (ก) ภาพสีต้นฉบับ (ข) ภาพไตรแมปแบบอัต โนมัตติจากข้อมูล RGBD

สำหรับรายละเอียดของแต่ละขั้นตอนข้างต้น ได้กล่าวถึงในหัวข้อต่อไปตามลำดับ

### 3.2 การปรับปรุงความไม่สมบูรณ์ของข้อมูลภาพระยะลึก

ความไม่สมบูรณ์ของข้อมูลภาพระยะลึกที่ได้รับจากกล้อง Kinect สามารถเกิดได้จาก 2 สาเหตุดังต่อไปนี้

1. ความไม่สมบูรณ์ของข้อมูลระยะลึกที่เกิดจากพื้นผิววัตถุไม่สะท้อนแสงอินฟราเรด (Infrared light) ทำให้ไม่สามารถคำนวณระยะลึกของบริเวณดังกล่าวได้ เช่น วัตถุที่มีพื้นผิวโปร่งใส, วัตถุที่เป็นเส้นขน เป็นต้น
2. ความไม่สมบูรณ์ของข้อมูลระยะลึกที่เกิดจากวัตถุบดบังแสงอินฟราเรด (Occlusion problem) หรืออาจเรียกได้ว่าบริเวณที่เป็นเงาของวัตถุเบื้องหน้า ด้วยเหตุนี้ทำให้บริเวณเงาของวัตถุกล้องไม่สามารถคำนวณข้อมูลระยะลึกได้ดังรูปที่ 3.7 (ก) แสดงการถูกบดบังแสงอินฟราเรดโดยที่จุด A คือจุดที่ถูกบดบังแสงอินฟราเรดจากวัตถุ FG ทำให้จุด A ไม่สามารถคำนวณหาค่าความลึกได้และรูปที่ 3.7 (ข) คือ ภาพระยะลึกที่เกิดปัญหาการบดบังแสงของวัตถุเบื้องหน้าเมื่อนำมาซ้อนทับภาพสีแสดงให้เห็นถึงบริเวณที่เกิดการบดบังแสงของวัตถุเบื้องหน้าที่จุด A ดังกล่าว



รูปที่ 3.7 (ก) จุด A คือจุดที่ไม่ได้รับแสงอินฟราเรดเพราะถูกบดบังโดยวัตถุ FG (ข) ตัวอย่างภาพระยะลึก เมื่อซ้อนทับภาพสีซึ่ง A คือบริเวณที่เกิดปัญหาการบดบังแสง

จากปัญหาความไม่สมบูรณ์ของข้อมูลระยะลึกทั้ง 2 สาเหตุที่เกิดขึ้นส่งผลกระทบต่อการนำข้อมูลระยะลึกไปใช้ดังต่อไปนี้คือ

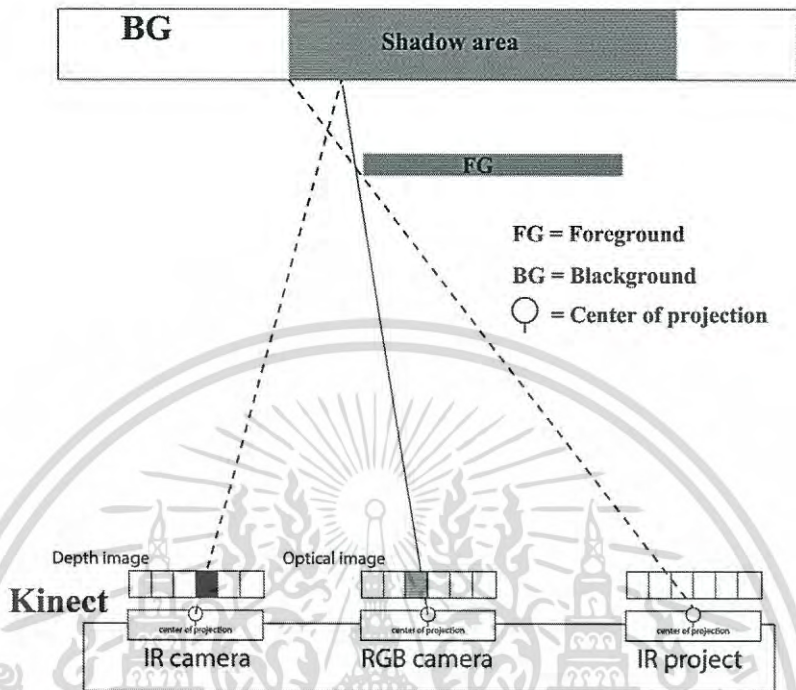
1. การใช้ข้อมูลระยะลึกในการระบุตำแหน่งของวัตถุบนภาพพบปัญหาเรื่องความถูกต้องในการอ้างอิงตำแหน่งกล่าวคือ บริเวณที่ไม่มีข้อมูลระยะลึกไม่สามารถแยกหรือระบุได้แน่ชัดว่าเป็นวัตถุหรือพื้นหลัง
2. ปัญหาการนำข้อมูลภาพระยะลึกมาใช้คำนวณสร้างภาพเมทเพื่อนำไปใช้ในการซ้อนภาพนั้นไม่สามารถคำนวณได้เพราะระบุไม่ได้ว่าเป็นระยะลึกของภาพวัตถุหรือพื้นหลัง

สำหรับแนวทางการแก้ไขปัญหาความไม่สมบูรณ์ของข้อมูลระยะลึกแก้ไขตามสาเหตุการเกิดที่กล่าวมาข้างต้น เพราะการแก้ไขความไม่สมบูรณ์ของข้อมูลระยะลึกด้วยวิธีการหาค่าเฉลี่ยจากระยะลึกข้างเคียงไม่สามารถคัดแยกวัตถุกับพื้นหลังได้ เนื่องจากข้อมูลระยะลึกที่เกิดจากเฉลี่ยค่าระหว่างวัตถุและพื้นหลังนั้น ทำให้ไม่สามารถระบุบริเวณดังกล่าวได้แน่ชัดว่าเป็นของวัตถุหรือพื้นหลัง ด้วยเหตุนี้การแก้ไขตามสาเหตุให้ผลที่ถูกต้องมากกว่าโดยสามารถแบ่งออกเป็น 2 ขั้นตอนหลักได้ดังนี้

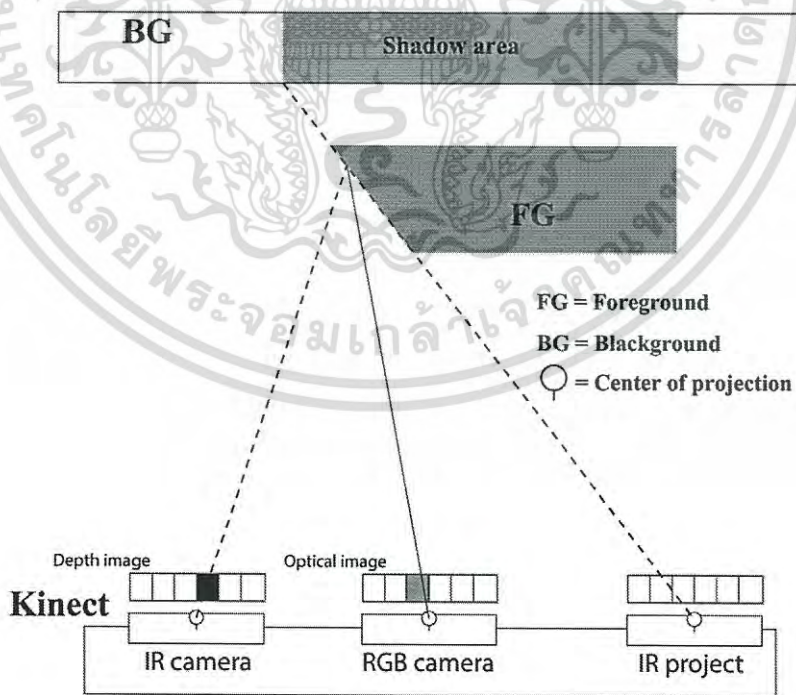
1. ขั้นตอนการระบุประเภทหรือสาเหตุของปัญหาความไม่สมบูรณ์ของข้อมูลระยะลึกที่เกิดขึ้นได้แก่กรณีที่วัตถุไม่สะท้อนแสงหรือถูกบังแสง
2. ขั้นตอนการคำนวณค่าระยะความลึกที่ขาดหายไป โดยวิธีการที่ใช้ขึ้นอยู่กับประเภทหรือสาเหตุของปัญหาที่เกิดขึ้น โดยสามารถแบ่งออกเป็น 2 กรณีได้แก่
  - กรณีที่จุดไม่ทราบค่าระยะลึกเกิดจากผิววัตถุไม่สะท้อนแสงอินฟราเรด สามารถคำนวณระยะลึกที่ขาดหายไปได้โดยอาศัยข้อมูลระยะลึกข้างเคียงแบบถ่วงน้ำหนัก
  - กรณีที่จุดไม่ทราบค่าระยะลึกเกิดเนื่องมาจากการบังแสงอินฟราเรด โดยวัตถุเบื้องหน้าสามารถคำนวณข้อมูลระยะลึกที่ขาดหายไปด้วยข้อมูลระยะลึกของพื้นหลังข้างเคียงแบบถ่วงน้ำหนักมาใช้

สำหรับการประมาณการนี้ใกล้เคียงกับความเป็นจริงมากขึ้นถ้าวัตถุมีความหนาน้อย หากวัตถุมีความหนาทำให้เกิดความผิดพลาดขึ้นได้ ดังรูปที่ 3.8 เมื่อวัตถุมีความหนาน้อยส่งผลทำให้บริเวณเงาของวัตถุคือพื้นหลังเพียงอย่างเดียว แต่ถ้าวัตถุมีความหนาดังรูปที่ 3.9 ทำให้บริเวณเงาของวัตถุคือพื้นหลังและตัววัตถุซึ่งส่งผลผิดต่อสมมุติฐานที่ว่า บริเวณเงาที่เกิดจากวัตถุบดบังแสงอินฟราเรด คือ บริเวณส่วนของพื้นหลังซึ่งสามารถแก้ไขข้อมูลระยะลึกที่ขาดหายไปจากระยะลึกพื้นหลังข้างเคียง และส่งผลต่อการใช้ขอบวัตถุที่สร้างจากข้อมูลภาพระยะลึกมากำหนดขนาดบริเวณไม่ชัดเจน เพราะขนาดบริเวณไม่ชัดเจนถูกกำหนดจากความคลาดเคลื่อนของตัวอุปกรณ์ ทำ

ให้ขอบวัตถุที่ผิดพลาดมีผลต่อการกำหนดความถูกต้องของบริเวณไม่ชัดเจนเนื่องด้วยบริเวณส่วนของวัตถุอาจถูกกำหนดเป็นพื้นหลังได้



รูปที่ 3.8 วัตถุที่มีความหนาน้อยโดยบริเวณเงาที่เกิดจากบั้งแสงของวัตถุเป็นส่วนหนึ่งของพื้นหลัง

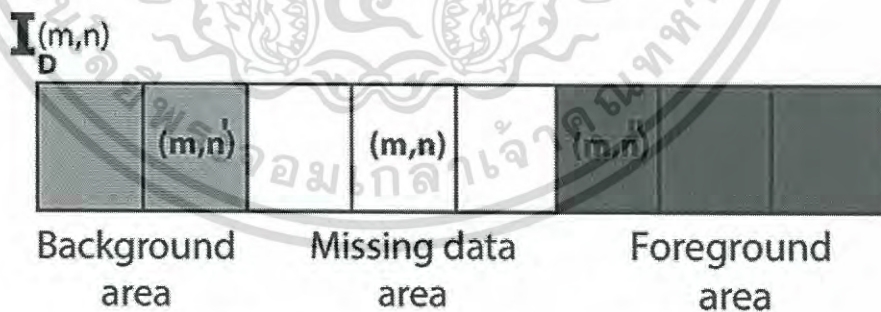


รูปที่ 3.9 วัตถุที่มีความหนามากโดยบริเวณบั้งแสงของวัตถุ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

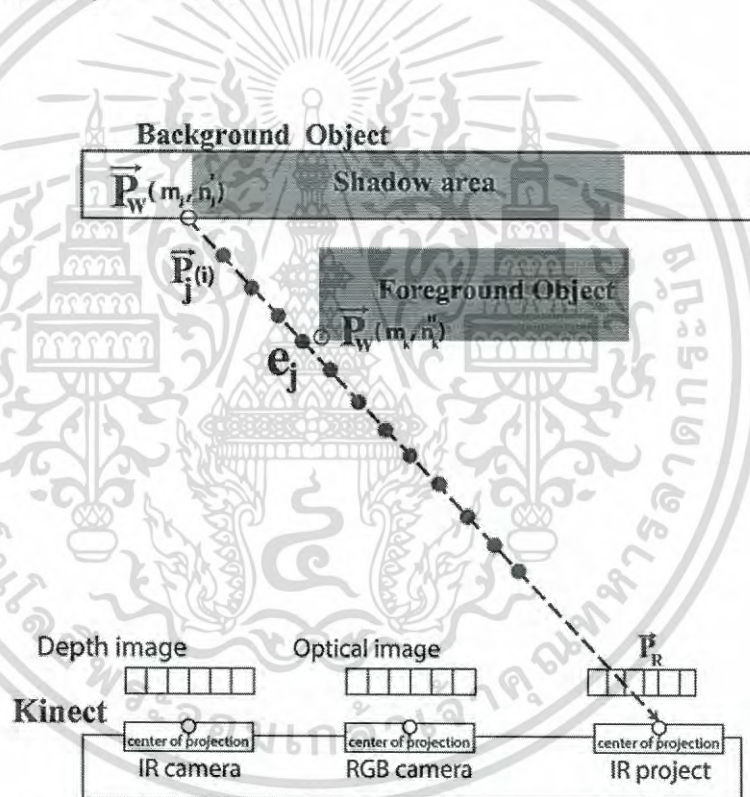
รายละเอียดของวิธีการตามแนวทางที่เสนอข้างต้น อธิบายเป็นขั้นตอนได้ดังต่อไปนี้

1. จากสมมุติฐานที่ว่าตามแนวแกน  $Z$  ของวัตถุ FG มีระยะห่างจากวัตถุอื่น ๆ ที่เป็นส่วนประกอบของพื้นหลัง ซึ่งวัตถุ FG เป็นกลุ่มที่มีค่าระยะลึกที่น้อยที่สุด ทำให้สามารถทำการแบ่งบริเวณของภาพออกเป็น 3 กลุ่มตามข้อมูลระยะลึกของวัตถุที่ปรากฏในภาพ โดยส่วนแรกคือ ส่วนของวัตถุ FG ที่เป็นกลุ่มของข้อมูลระยะลึกที่น้อยที่สุด , ส่วนของวัตถุอื่น ๆ หรือส่วนของพื้นหลังซึ่งเป็นกลุ่มข้อมูลระยะลึกกลุ่มอื่นนอกเหนือจากกลุ่มของวัตถุ FG และส่วนของกลุ่มที่ไม่มีข้อมูลระยะลึก โดยการแบ่งกลุ่มใช้อัลกอริทึม K-means แบ่งกลุ่มข้อมูลระยะลึกออกเป็น 3 กลุ่มตามที่กล่าวมาโดยกำหนด  $o$  ให้  $F$  คือ เซตของตำแหน่งพิกเซลของวัตถุเบื้องหน้า จากข้อมูลภาพระยะลึก ,  $B$  คือ เซตของตำแหน่งพิกเซลพื้นหลังจากข้อมูลภาพระยะลึก และ  $I_d(m, n)$  คือข้อมูลภาพระยะลึก โดยที่  $m, n$  แถวและคอลัมน์ในแต่ละพิกเซล ของภาพ,  $m = 0, 1, \dots, M-1$  และ  $n = 0, 1, \dots, N-1$  โดยที่  $M$  คือ ความสูง และ  $N$  คือ ความกว้างของภาพ
2. ตรวจสอบข้อมูลภาพระยะลึก  $I_d(m, n)$  แต่ละพิกเซลที่ไม่มีข้อมูลระยะลึกตามเงื่อนไข  $n' < n < n''$  จากเหตุผลลักษณะปัญหาการบังแสงของวัตถุ ดังรูปที่ 3.10 เพราะข้อเท็จจริงทางโครงสร้างของ Kinect ที่ IR project อยู่ทางขวา ทำให้เงาที่เกิดจากการบัง มีทิศทางของเงาพาดจากขวาไปซ้าย (เนื่องจากตำแหน่งของแสง IR อยู่ทางขวาของกล้อง) โดยที่ตำแหน่งพิกเซล  $(m, n') \in B$  คือตำแหน่งข้อมูลระยะลึกของพื้นหลังและ  $(m, n'') \in F$  คือตำแหน่งข้อมูลระยะลึกของวัตถุ



รูปที่ 3.10 ภาพตัวอย่างลักษณะการค้นหาริเวณที่ไม่ทราบค่า

3. ทำขั้นตอนที่ 2 จนครบทุกแถวของภาพระยะลึก โดยกำหนดให้  $S$  เป็น เซตทิวเปิล (Tuple) ของ  $J$  ที่  $m_j < n_j < n'_j$  ตามความสัมพันธ์ของเงื่อนไขขั้นตอนที่ 2,  $n'_j$  และ  $n_j$  คือพิกเซลถัดไปที่อยู่ทางด้านซ้ายและขวาของ  $I_d(m_j, n_j)$  จำนวน  $j$  พิกเซล
4. ตรวจสอบแต่ละองค์ประกอบ  $m_j < n_j < n'_j$  ของ  $S$  ว่ามีความสอดคล้องเป็นบริเวณที่เกิด จากเงาของวัตถุหรือไม่จากการแบ่งเส้นระหว่าง  $\vec{P}_W(m_j, n'_j)$  และ  $\vec{P}_R$  ออกเป็นช่วงๆ โดย กำหนดให้  $\vec{P}_W(m_j, n'_j)$  คือตำแหน่งจริงใน World coordinate ของ  $I_d(m_j, n'_j)$ ,  $\vec{P}_W(m_k, n'_k)$  คือ ตำแหน่งจริงใน World coordinate ของ  $I_d(m'_k, n'_k)$ ,  $k = 0, 1, 2, \dots, J - 1$ ,  $\vec{P}_R$  คือตำแหน่งจริงของ IR projector ของ Kinect ที่อยู่ใน World coordinate ( $\vec{P}_R = [0.05, 0, 0]^T$ ) และ  $\vec{P}_j(i)$  ที่  $i = 0, 1, 2, \dots, K - 1$  ซึ่ง  $K$  คือจุด ระหว่าง  $\vec{P}_W(m_j, n'_j)$  และ  $\vec{P}_R$  (รูปที่ 3.11)

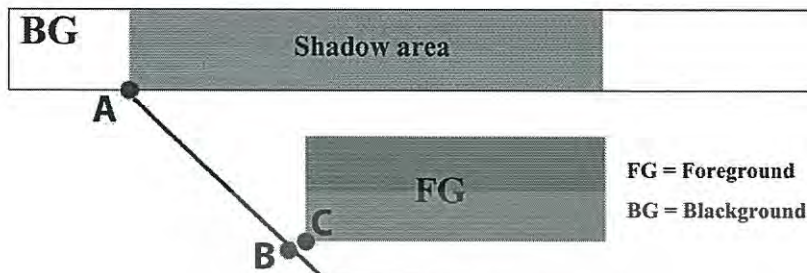


รูปที่ 3.11 ภาพตัวอย่างลักษณะการแบ่งเส้นระหว่าง  $\vec{P}_W(m_j, n'_j)$  และ  $\vec{P}_R$  ออกเป็นช่วงๆ

นำมาคำนวณหาระยะห่างระหว่างตำแหน่งที่เป็นไปได้ของ  $\vec{P}_j(i)$  และ  $\vec{P}_W(m_k, n'_k)$  เพื่อนำมาหาระยะทางที่น้อยที่สุดตามสมการต่อไปนี้ (3.1) จากเหตุผลบริเวณที่เป็นเงาของ วัตถุจากการบังแสงอินฟราเรด เมื่อตรวจสอบตามแนวแสงอินฟราเรด มีจุดบนแนวแสงที่ ระยะลึกใกล้เคียงกับระยะลึกของวัตถุและมีระยะห่างระหว่างจุดน้อย ดังรูปที่ 3.12 ที่จุดบน แนวแสงอินฟราเรดกับวัตถุมีระยะห่างกันระหว่างจุดน้อย

$$e_j(i) = \min_{k=0, \dots, J-1} \|\vec{P}_j(i) - \vec{P}_w(m_k, n_k)\| \quad (3.1)$$

$$e_j = \min_{i=0, \dots, K-1} e_j(i) \quad (3.2)$$



รูปที่ 3.12 ภาพการตรวจสอบแนวแสงอินฟราเรดที่จุด A พบว่าที่จุด B บนแนวแสงกับจุด C ที่เป็นขอบของวัตถุที่ทำให้เกิดเงามีระยะห่างระหว่างจุดน้อยและมีระยะลึกใกล้เคียงกัน (ในทางทฤษฎีเป็นจุดเดียวกัน แต่ในทางปฏิบัติไม่ได้เป็นเช่นนั้นเนื่องจากความคลาดเคลื่อนของอุปกรณ์)

ในทางทฤษฎีจุดที่เป็นขอบของวัตถุที่บังแสงอินฟราเรดจนเกิดเงาบนพื้นหลังจากจุดตามแนวเส้นที่ลากจากตำแหน่งแสงอินฟราเรดที่มายังจุดดังกล่าว ที่สัมผัสกับผิวของวัตถุสามารถตรวจสอบตำแหน่งของจุดที่มีค่าระยะลึก และระยะห่างระหว่างจุดบนแนวแสงอินฟราเรดที่ตรงกัน

แต่ในทางปฏิบัติความคลาดเคลื่อนของกล้อง Kinect ทำให้เกิดความคลาดเคลื่อนของตำแหน่งระหว่างจุด ดังนั้นจึงกำหนดเงื่อนไขการตรวจสอบบริเวณที่ไม่มีค่าระยะลึกเกิดจากการบังแสงอินฟราเรดหรือไม่ เมื่อตำแหน่งจุดที่ใกล้ที่สุดของผิววัตถุไปยังเส้นที่ลากผ่านตำแหน่งแสงอินฟราเรด กำหนดให้คือ  $t$  ค่าเกณฑ์ในการตรวจสอบระยะห่างที่ใกล้ที่สุด เมื่อ

- $e_j(i) < t$  สรุปได้ว่า  $I_d(m_j, n_j)$  คือบริเวณที่เกิดจากเงาของวัตถุเบื้องหน้า จึงสามารถแก้ไขข้อมูลระยะลึกที่หายไปของบริเวณดังกล่าวจากการให้ค่าน้ำหนักของพื้นหลังข้างเคียง  $I_d(m_j, n_j)$
- $e_j(i) > t$  สรุปได้ว่าบริเวณดังกล่าวไม่ได้เกิดจากเงาของวัตถุแต่เกิดจากการที่วัตถุไม่สะท้อนแสงอินฟราเรด ซึ่งสามารถแก้ไขข้อมูลระยะลึกบริเวณดังกล่าวได้จากการให้ค่าน้ำหนักของพื้นหลังและวัตถุในบริเวณดังกล่าว

โดยค่า  $t$  ที่เหมาะสมได้จากการทดลองหรือกำหนดค่าโดยผู้ใช้

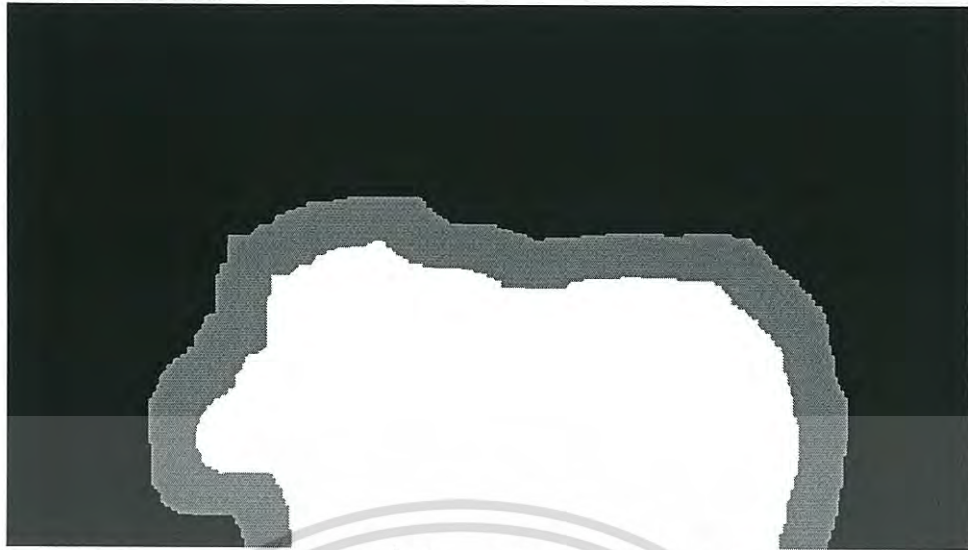
### 3.3 การตรวจสอบหาขอบวัตถุจากข้อมูลระยะลึก

หัวข้อนี้กล่าวถึงการตรวจสอบค้นหาเส้นขอบระยะลึกที่ผ่านการปรับปรุงแก้ไขในขั้นตอนที่ 3.2 โดยอาศัยหลักการของ K-means ในการแบ่งกลุ่มของข้อมูลภาพระยะลึกเพื่อค้นหาตำแหน่งวัตถุจากสมมุติฐานที่ว่าตำแหน่งของวัตถุมีระยะห่างจากพื้นหลังอย่างชัดเจน :จากการแบ่งกลุ่มทำให้ได้ข้อมูลบริเวณของวัตถุและพื้นหลัง เพื่อนำข้อมูลดังกล่าวที่ส่วนของวัตถุมาค้นหาเส้นขอบระยะลึก ด้วยวิธีการ Sobel edge detection กำหนดให้เส้นขอบระยะลึกคือ  $E_d$

### 3.4 การตรวจสอบหาขอบวัตถุจากเส้นขอบสี

หัวข้อนี้กล่าวถึงการตรวจสอบค้นหาเส้นขอบสีของภาพวัตถุ เพื่อนำข้อมูลเส้นขอบสีมาใช้เป็นส่วนประกอบในการกำหนดขนาดบริเวณ ไม่ชัดเจนกับเส้นขอบระยะลึก จากสมมุติฐานที่ว่าคุณภาพของภาพไตรแมบมีผลต่อคุณภาพและประสิทธิภาพในการซ้อนภาพ ซึ่งการสร้างภาพไตรแมบบนอัตโนมัติสามารถทำได้โดยอาศัยข้อมูลเส้นขอบระยะลึกในการอ้างอิงตำแหน่งของวัตถุ แต่ข้อมูลภาพระยะลึกมีความคลาดเคลื่อนจากความเป็นจริงซึ่งเกิดจากตัวอุปกรณ์ ทำให้การใช้เส้นขอบระยะลึกมาอ้างอิงตำแหน่งของภาพวัตถุจำเป็นต้องชดเชยความคลาดเคลื่อนที่เกิดขึ้น ซึ่งวิธีการหนึ่งที่สามารถนำมาใช้ปรับปรุงความถูกต้องในการสร้างภาพไตรแมบบนอัตโนมัติคือ การนำข้อมูลเส้นขอบสีมาช่วยในการอ้างอิงตำแหน่งของภาพวัตถุจากเหตุผลที่ว่า เส้นขอบสีคือบริเวณที่ค่าสีมีความแตกต่างผ่านเกณฑ์ที่กำหนดทำให้เส้นขอบสีมีความถูกต้องในการอ้างอิงตำแหน่งสูง เพราะไม่มีความคลาดเคลื่อนของตัวอุปกรณ์มาเกี่ยวข้อง สำหรับขั้นตอนการตรวจสอบค้นหาเส้นขอบสีของภาพวัตถุสามารถทำได้ดังต่อไปนี้

1. การตรวจสอบเส้นขอบสีของภาพวัตถุใช้ข้อมูลภาพไตรแมบเริ่มต้น (Initial Trimap) ดังรูปที่ 3.13 ซึ่งสร้างจากข้อมูลค่าความเคลื่อนไหวของตัวอุปกรณ์ Kinect [หัวข้อที่ 2.2.2] กับข้อมูลเส้นขอบระยะลึก โดยการใช้เส้นขอบระยะลึกเป็นตำแหน่งอ้างอิงบริเวณไม่ชัดเจนด้วยการขยาย (Dilation) ตามขนาดค่าความเคลื่อนไหว ทำให้สามารถกำหนดบริเวณของภาพพื้นหลังคือ ส่วนด้านนอกบริเวณไม่ชัดเจนและบริเวณของภาพวัตถุคือส่วนที่อยู่ด้านใน



รูปที่ 3.13 ตัวอย่างภาพไทรแมบเริ่มต้นที่สร้างจากค่าความเคลื่อนไหวของตัวอุปกรณ์ Kinect กับข้อมูลเส้นขอบระยะลึก

ซึ่งในการค้นหาข้อมูลเส้นขอบที่ใช้ภาพสีเฉพาะบริเวณไม่ชัดเจนเท่านั้นเพื่อลดการคำนวณส่วนที่ไม่เกี่ยวข้องออกไป โดยกำหนดให้  $F_a(m, n)$  คือส่วนข้อมูลสีของวัตถุ,  $B_a(m, n)$  คือส่วนของข้อมูลสีพื้นหลัง และ  $U_a(m, n)$  คือส่วนข้อมูลสีบริเวณไม่ชัดเจนดังรูปที่ 3.14



รูปที่ 3.14 (ก) ส่วนข้อมูลสีพื้นหลังจากภาพไทรแมบเริ่มต้น (ข) ส่วนข้อมูลสีวัตถุจากภาพไทรแมบเริ่มต้น (ค) ส่วนข้อมูลสีบริเวณไม่ชัดเจนจากภาพไทรแมบเริ่มต้น

- ทำการตรวจสอบค้นหาเส้นขอบของภาพสี ด้วยอัลกอริทึมในการค้นหาขอบของภาพ (Image edge detection method) โดยงานวิทยานิพนธ์ฉบับนี้ใช้ Sobel edge detection ในการค้นหาเส้นขอบสีเนื่องจาก เป็นทฤษฎีการหาค่าความแตกต่างระหว่างจุดในแนวนอนและแนวตั้ง ซึ่งมีความคมชัดของเส้นขอบที่ได้ สำหรับการค้นหาเส้นขอบจากข้อมูลรูปภาพสีหาจากค่าความแตกต่างรวมของค่าสีแต่ละชานแนลสี RGB เพื่อหาผลรวมเฉลี่ยของความแตกต่างค่าสีจากข้อมูลภาพ  $U_a$  โดยสามารถเขียนในลักษณะของฟังก์ชันที่ค่าแต่ละจุดถูกกำหนดโดยเวกเตอร์ (Vector) 3 มิติที่เป็นค่าสี RGB ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$E_{rgb} = f(I_r(m, n), I_g(m, n), I_b(m, n)) \quad (3.3)$$

โดยที่  $E_{rgb}$  คือรูปภาพที่จากการหาเส้นขอบ,  $(m, n)$  คือ ตำแหน่งอ้างอิงของภาพ,  $I_r$ ,  $I_g$  และ  $I_b$  คือ ข้อมูลรูปภาพสีในแต่ละชาแนล RGB ของภาพที่ใช้และ  $f$  คือ ฟังก์ชัน Sobel edge detection

3. สำหรับการค้นหาเส้นขอบสี จำเป็นต้องมีการกำหนดค่าเกณฑ์ (Threshold) ที่เหมาะสม เพื่อให้ได้บริเวณที่เป็นขอบของรูปภาพที่ชัดเจน หากค่าเกณฑ์ที่กำหนดมากไปทำให้สูญเสียขอบของรูปภาพหรือหากกำหนดน้อยไปเป็นการเพิ่มเส้นขอบที่มีไม่เกี่ยวข้องเพิ่มมากขึ้น สำหรับวิทยานิพนธ์ฉบับนี้นำเสนอแนวทางการกำหนดค่าเกณฑ์ที่เหมาะสมดังต่อไปนี้
  - หลักแนวคิดในการกำหนดค่าเกณฑ์ คือ อาศัยหลักเมื่อลดค่าเกณฑ์ตั้งต้นทำให้จำนวนเส้นขอบเพิ่มขึ้น ซึ่งเมื่อนำจำนวนกลุ่มของเส้นขอบและความยาวของเส้นที่มากที่สุดมาเปรียบเทียบกับความแตกต่างกับจำนวนกลุ่มของเส้นขอบตั้งต้น หากไม่เกินค่าที่กำหนด แสดงว่าสามารถลดค่าเกณฑ์ลงได้จนกว่าไม่ผ่านเกณฑ์ที่กำหนด โดยกำหนดให้  $t_a$  คือ ค่าเกณฑ์ที่ได้จากการคำนวณของฟังก์ชัน  $f$  เป็นค่าเกณฑ์เริ่มต้น
  - ใช้  $E_{rgb}$  และ  $t_a$  เป็นข้อมูลภาพตั้งต้นในการค้นหาค่าเกณฑ์ โดยคำนวณหาค่าความยาวเส้น ขอบสี  $E_{rgb}$  ที่มีค่าความยาวต่อเนื่องมากที่สุดและจำนวนกลุ่มของเส้นขอบสี  $E_{rgb}$  โดยกำหนดให้เป็น  $mt$  และ  $mg$  ตามลำดับ
  - หากภาพเส้นขอบสีใหม่จากการปรับลดค่าเกณฑ์ สำหรับการปรับลดค่าเกณฑ์ ลงนั้น ขึ้นอยู่กับผู้ใช้เป็นคนกำหนด หากปรับค่าเกณฑ์ ลงครั้งละมาก ๆ ทำให้คำนวณได้รวดเร็วขึ้นแต่ลดความละเอียดลง ในวิทยานิพนธ์ฉบับนี้ปรับลดค่าเกณฑ์ลงทีละ 0.02 จากค่า  $t_a$  เนื่องจากให้ผลที่ความละเอียดและไม่ช้าจนเกินไป ซึ่งค่าดังกล่าวได้จากการทดลอง และกำหนดให้  $E_c$  คือ เส้นขอบสีที่ได้จากการปรับค่าเกณฑ์และ  $t_c$  คือค่าเกณฑ์ที่ถูกปรับลง
  - คำนวณหาค่าความยาวเส้นขอบสี  $E_c$  มีค่าความยาวต่อเนื่องมากที่สุดและจำนวนกลุ่มของเส้นขอบสี โดยกำหนดให้คือ  $mt'$  และ  $mg'$  ตามลำดับ
  - คำนวณหาค่าความแตกต่างระหว่างเส้นขอบสี  $E_{rgb}$  กับเส้นขอบสี  $E_c$  จากความยาวต่อเนื่องมากที่สุดและจำนวนกลุ่มของเส้นได้จากสมการต่อไปนี้

$$e_{mt} = \frac{|mt - mt'|}{mt} \quad (3.4)$$

$$e_{mg} = \frac{|mg - mg'|}{mg} \quad (3.5)$$

- ตรวจสอบค่า  $e_{mt}$  และ  $e_{mg}$  ว่าผ่านเกณฑ์ที่กำหนดหรือไม่ โดยตั้งเกณฑ์ที่ความแตกต่างไม่เกิน 30 เปอร์เซ็นต์
- หากผ่านเกณฑ์ที่กำหนดควนทำซ้ำจนกว่าไม่ผ่านเกณฑ์ หากไม่ผ่านเกณฑ์ที่กำหนดเลือกใช้ภาพเส้นขอบสี  $E_t$  ที่ได้จากค่า  $t_t$  ที่ผ่านเกณฑ์ก่อนหน้ามาใช้
- กำหนดให้เส้นขอบสีจากการกำหนดค่าเกณฑ์ใหม่คือ  $C_{rgb}$

### 3.5 การคัดกรองเส้นขอบสีที่เกิดจากลวดลาย

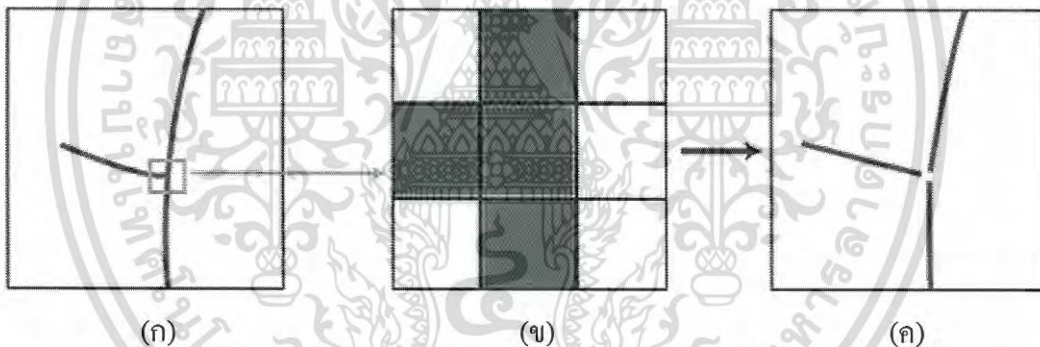
เป้าหมายของขั้นตอนนี้ คือ การคัดกรองเส้นขอบสีที่เกิดจากลวดลายวัตถุหรือพื้นหลัง ด้วยเหตุผลที่ว่า ภาพวัตถุและภาพพื้นหลังนั้นสามารถมีลวดลายและมีสีที่หลากหลาย ทำให้เกิดเส้นขอบสีที่ไม่ได้เป็นส่วนขอบวัตถุ สำหรับการคัดกรองเส้นขอบสีที่เกิดจากลวดลายสามารถแบ่งออกเป็นขั้นตอนเบื้องต้นได้ดังต่อไปนี้

1. การแยกกลุ่มของเส้นขอบที่รวมกัน เพื่อใช้ในการตรวจสอบเส้นขอบสี เนื่องด้วยเส้นขอบสีสามารถเกิดจากการรวมตัวของเส้นขอบได้มากกว่า 1 เส้น
2. การแบ่งเส้นขอบสีออกเป็นช่วง เพื่อลดความผิดพลาดการคัดแยกเส้นขอบสี ด้วยเหตุผลที่ว่าเส้นขอบสีเกิดจากความแตกต่างระหว่างค่าสีโดยรอบเส้นขอบ ซึ่งในแต่ละเส้นขอบสีสามารถเกิดจากความแตกต่างกันระหว่างค่าสีได้มากกว่า 2 สี
3. การตรวจสอบค่าสีรอบเส้นขอบ คือ การค้นหาค่าสีที่ทำให้เกิดเส้นขอบ เพื่อนำมาใช้ในการตรวจสอบหาเส้นขอบวัตถุที่เกิดจากความแตกต่างระหว่างสีของวัตถุและพื้นหลัง
4. การแบ่งกลุ่มสีของเส้นขอบ เพราะภาพวัตถุและภาพพื้นหลังสามารถมีลวดลายและมีสีที่หลากหลาย จึงทำให้เกิดเส้นขอบที่มีค่าสีที่แตกต่างกัน ดังนั้นการแบ่งกลุ่มสีของเส้นขอบเพื่อเตรียมข้อมูลในการนำไปตรวจสอบคัดกรองหากกลุ่มสีที่ใกล้เคียงกับค่าสีของภาพวัตถุและภาพพื้นหลังมากที่สุด
5. การตรวจสอบคัดกรองเส้นขอบสี เป็นขั้นตอนการคัดแยกเส้นขอบสีที่ไม่ได้เป็นขอบของวัตถุแต่เกิดจากสีของลวดลายที่มีความแตกต่างกัน โดยตรวจสอบความใกล้เคียงของค่าสีในแต่ละเส้นขอบกับค่าสีของภาพวัตถุและค่าสีของภาพพื้นหลัง

6. การคัดกรองความยาวของเส้นขอบสี เพื่อลดปัญหาเส้นขอบที่เป็นสัญญาณรบกวนเช่น เส้นขน เนื่องด้วยเส้นขอบสีควรมีความยาวต่อเนื่องในระดับหนึ่ง เพราะหากมีความยาว น้อยเกินกำหนด ทำให้ไม่สามารถระบุได้แน่ชัดว่าคือส่วนของเส้นขอบวัตถุ โดยที่ค่า เกณฑ์คือ ค่าความยาวของเส้นขอบสีขั้นต่ำที่ต่อเนื่อง

### 3.5.1 การแยกเส้นขอบสี

ในการคัดกรองเส้นขอบสีต้องทำการแยกเส้นขอบสีที่มีการรวมกลุ่มมากกว่าหนึ่งเส้น เนื่องด้วยในแต่ละเส้นขอบสีสามารถเกิดการรวมตัวของเส้นขอบได้มากกว่า 1 เส้น เพราะ ความแตกต่างของค่าสีที่ไม่เหมือนกันของวัตถุและพื้นหลัง จึงจำเป็นต้องแยกเส้นขอบสีจาก สำหรับการแยกกลุ่มเส้นขอบสี  $C_{rgb}$  ทำได้โดยใช้กรอบ (Window) ขนาด  $3 \times 3$  ครอบไป ในแต่ละพิกเซลของ  $C_{rgb}$  เพื่อหาค่าผลรวมของเส้นในแต่ละกรอบ ถ้าค่าผลรวมของกรอบมีค่า มากกว่า 3 แสดงว่าบริเวณดังกล่าวเกิดจากการรวมตัวของเส้นขอบสีมากกว่า 1 กลุ่ม ให้ทำการ ตัดส่วนบริเวณดังกล่าวออก เพื่อทำการแยกกลุ่มเส้นขอบสีดังรูปที่ 3.15 โดยกำหนดให้เป็น  $C_s$  คือ เส้นขอบสีที่ผ่านการแบ่งออกเป็นกลุ่มๆ



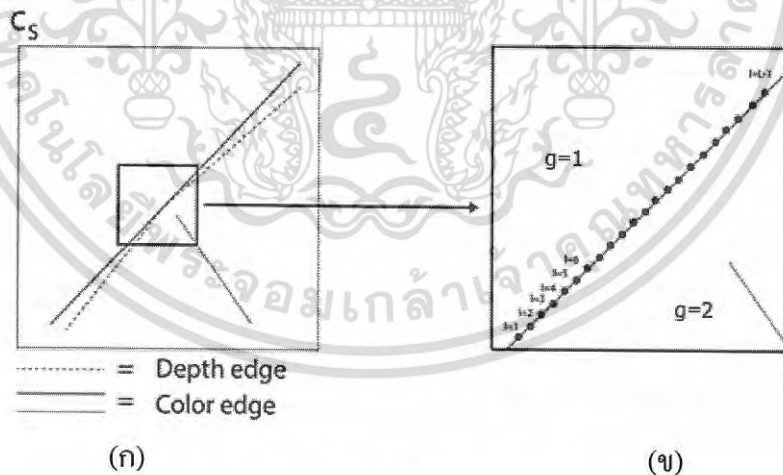
รูปที่ 3.15 (ก) จุดบริเวณที่เป็นรอยต่อเกิดจากการรวมตัวของกลุ่มเส้นขอบสีมากกว่า 1 กลุ่ม  
(ข) ตัวอย่างภาพพิกเซลบริเวณดังกล่าวซึ่งเกิดจากการรวมตัวของกลุ่มเส้นขอบสีมากกว่า 1 กลุ่ม  
(ค) ภาพหลังการแยกกลุ่มของเส้นขอบสี

### 3.5.2 การแบ่งเส้นขอบสีออกเป็นช่วง

การแบ่งเส้นขอบสีออกเป็นช่วง เพื่อลดความผิดพลาดการคัดแยกเส้นขอบสีเนื่องด้วย ในแต่ละเส้นขอบสีสามารถเกิดจากความแตกต่างของค่าสีได้มากกว่า 2 สี โดยในการแบ่งเส้น ขอบสีออกเป็นช่วงสามารถทำได้โดยการใช้กรอบ (Window) ตามตำแหน่งอ้างอิงของเส้นขอบ ระยะถึก ซึ่งขนาดของกรอบสามารถกำหนดได้จากค่าความคลาดเคลื่อนของตัวอุปกรณ์ Kinect เพราะ เส้นขอบวัตถุไม่เกินค่าความคลาดเคลื่อนที่เกิดขึ้น โดยการแบ่งเส้นขอบสีออกเป็นช่วง สามารถทำได้ตามขั้นตอนต่อไปนี้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. จัดเรียงตำแหน่งเส้นขอบระยะลึก  $E_d$  ตามรูปร่างของวัตถุ โดยกำหนดให้  $D_j = (m_j, n_j)$  โดยที่  $j = 0, 1, \dots, H-1$ , โดยที่  $(m_j, n_j) = \{(m_j, n_j) | E_d(m_j, n_j) = 1\}$  เป็นตำแหน่งพิกเซลที่อยู่บนเส้นวัตถุ ระยะลึกจำนวน  $H-1$  และ  $H$  คือ จำนวนพิกเซลของเส้นขอบระยะลึก  $E_d$
2. กำหนดให้  $W$  คือกรอบขนาด  $Q \times Q$  เพื่อใช้ตรวจสอบค้นหาเส้นขอบสี  $C_s$  ตามตำแหน่ง  $D_j$  โดยที่  $Q$  คือ ค่าความคลาดเคลื่อนอุปกรณ์ Kinect สามารถคำนวณได้ตามหัวข้อที่ 2.2.2 เรื่อง การคำนวณค่าความคลาดเคลื่อนของตัวอุปกรณ์ โดยเหตุผลที่กำหนดให้  $W$  มีกรอบขนาด  $Q \times Q$  เพราะ เส้นขอบวัตถุไม่เกินค่าความคลาดเคลื่อนของอุปกรณ์ Kinect หากเกินกรอบของภาพแสดงว่าเส้นขอบดังกล่าวคือ ส่วนของลวดลายวัตถุหรือพื้นหลัง
3. นำเส้นขอบสี  $C_s$  ที่พบใน  $W$  ตามตำแหน่ง  $D_j$  มาแบ่งส่วนออกเป็นกลุ่มตามความยาวที่ต่อเนื่องกัน (รูปที่ 3.16) กำหนดให้  $C_{l,g} = (m_{l,g}, n_{l,g})$  โดยที่  $g = 1, 2, \dots, G-1$  คือหมายเลขอ้างอิงกลุ่มที่พบใน  $W$ ,  $l = 1, 2, \dots, L-1$  โดยที่  $l$  คือตำแหน่งเส้นขอบสีที่พบใน  $W$  ของแต่ละกลุ่ม,  $G$  คือจำนวนกลุ่มเส้นขอบสีที่พบใน  $W$  และ  $L$  คือจำนวนพิกเซล ของแต่ละกลุ่ม สำหรับงานวิจัยฉบับนี้แบ่งส่วนความยาวของกลุ่มที่  $\frac{Q}{2}$  พิกเซลหากมีความยาวไม่ถึงที่กำหนดใช้ตามความยาวที่พบ



รูปที่ 3.16 (ก) บริเวณเส้นขอบสีที่ทำการกรอบ  $W$  ที่ตำแหน่ง  $D_j$ ,

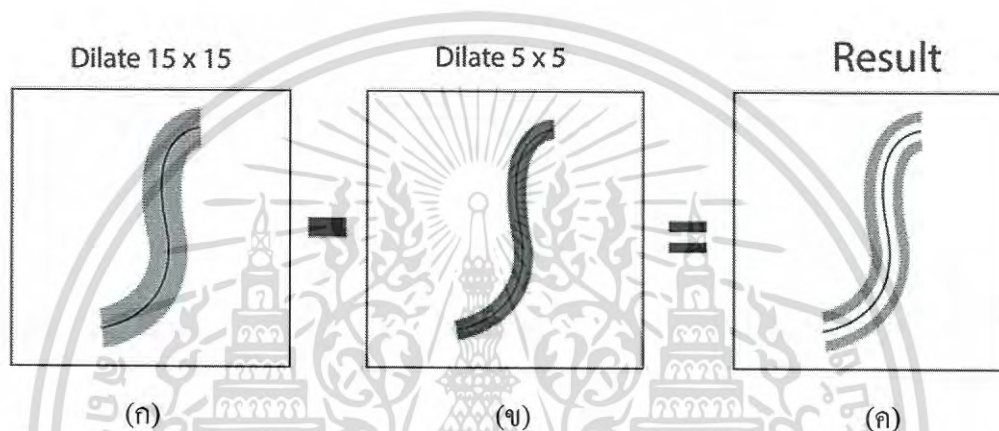
(ข) ตัวอย่างการอ้างอิงตำแหน่งเส้นขอบสีที่พบในกรอบ

### 3.5.3 การตรวจสอบค่าสีรอบเส้นขอบ

การตรวจสอบค่าสีรอบเส้นขอบ  $C_{l,g}$  เพื่อนำไปใช้ในการตรวจสอบหาเส้นขอบวัตถุที่เกิดจากสีของวัตถุและสีของพื้นหลัง ซึ่งในการตรวจสอบค่าสีรอบเส้นขอบนั้นไม่ได้ใช้ค่าสีที่เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ติดกับเส้นขอบโดยตรง เพราะค่าสีที่ติดกับเส้นขอบโดยตรงมีการผสมกันระหว่างค่าสีโดยรอบที่ทำให้เกิดเส้น ทำให้เมื่อนำมาตรวจสอบความแตกต่างระหว่างสีของวัตถุและพื้นหลังจึงเกิดความผิดพลาด

ด้วยเหตุผลดังกล่าวจึงใช้ข้อมูลค่าสีใกล้เคียงที่อยู่ถัดจากค่าสีรอบเส้นขอบ  $C_{1,g}$  โดยตรงสามารถทำได้จากการขยายเส้น  $C_{1,g}$  ที่ระยะแตกต่างกัน 2 ช่วง โดยงานวิจัยฉบับนี้ใช้การขยายที่ขนาด  $5 \times 5$  และ  $15 \times 15$  เพื่อหาพื้นที่ส่วนเกิน ดังรูปที่ 3.17 ซึ่งทำให้ได้ค่าสี 2 กลุ่มจากข้อมูลค่าสีใกล้เคียงรอบเส้นขอบ  $C_{1,g}$  โดยกำหนดให้คือ C1 และ C2 ตามลำดับ



รูปที่ 3.17 (ก) ภาพส่วนขยายของกลุ่ม  $C_{1,g}$  ที่ขนาด  $15 \times 15$  (ข) ภาพส่วนขยายของกลุ่ม  $C_{1,g}$  ที่ขนาด  $5 \times 5$  (ค) บริเวณส่วนต่างที่นำไปใช้ในการคำนวณ

### 3.5.4 การตรวจสอบคัดแยกค่าสีโดยรอบเส้นขอบ

การตรวจสอบคัดแยกค่าสีโดยรอบเส้นขอบออกเป็นกลุ่ม เพื่อนำไปใช้ในการตรวจสอบหาเส้นขอบวัตถุจากความใกล้เคียงของค่าสี เพื่อคำนวณหาค่ากลุ่มสีของวัตถุและพื้นหลังที่ใกล้กับเส้นขอบสี เพราะภาพวัตถุและพื้นหลังนั้นสามารถมีสีได้หลากหลาย จึงต้องทำการแบ่งค่าสีออกเป็นกลุ่ม ซึ่งการคำนวณค่าสีในแต่ละกลุ่มของวัตถุและพื้นหลังอ้างอิงตามตำแหน่งเช่นเดียวกับขั้นตอนที่ 3.5.2 โดยการแบ่งกลุ่มค่าสีของ  $F_a$  และ  $B_a$  ตามตำแหน่งอ้างอิงที่พบใน  $W$  ด้วย K-means สำหรับการทำ K-means ได้จุดศูนย์กลางของแต่ละกลุ่มซึ่งเป็นค่าเฉลี่ยของกลุ่มสีที่อยู่ใน  $W$  กำหนดได้ว่า  $M_{fg}(p)$  และ  $M_{bg}(q)$  คือค่าเฉลี่ยของค่าสีในแต่ละกลุ่มของ FG และ BG ที่พบใน  $W$  ตามลำดับโดยที่  $p = 0, 1, 2, \dots, P - 1$ ,  $q = 0, 1, 2, \dots, Q - 1$  และ  $P, Q$  คือจำนวนกลุ่มที่พบ

### 3.5.5 การคัดกรองเส้นขอบสีที่เกิดจากลวดลาย

ขั้นตอนการคัดกรองเส้นขอบสีโดยการเปรียบเทียบความแตกต่างระหว่างค่าสีในแต่ละกลุ่มกับกลุ่มสีของวัตถุและพื้นหลัง เพื่อหากกลุ่มสีที่มีความใกล้เคียงกับสีของภาพวัตถุและพื้นหลังซึ่งผ่านเกณฑ์ที่กำหนด โดยสามารถทำการตรวจสอบความใกล้เคียงของค่าสีของกลุ่ม C1 กับสีของ FG และ BG คิดเป็นทีเปอร์เซ็นต์ จากการหาค่าความแตกต่างที่น้อยที่สุดระหว่าง  $M_{fg}(p)$  และ  $M_{bg}(q)$  ตามสมการต่อไปนี้

$$t_1 = \operatorname{argmin}_{p=0, \dots, P-1} |C1 - M_{fg}(p)| \quad (3.6)$$

$$t_2 = \min_{q=0, \dots, Q-1} |C1 - M_{bg}(q)| \quad (3.7)$$

$$t_{f1} = \frac{t_1}{t_1 + t_2} \quad (3.8)$$

$$t_{b1} = \frac{t_2}{t_1 + t_2} \quad (3.9)$$

ซึ่งในการตรวจสอบทำให้ทราบว่า C1 มีค่าสีที่ใกล้เคียงกับ FG หรือ BG หรือมีความใกล้เคียงกันทั้งคู่โดยกำหนดให้  $t$  คือค่าเกณฑ์ที่ใช้ในการตรวจสอบมีค่าเท่ากับ 0.30 (เป็นค่าที่ได้จากการทดลอง ซึ่งหากมีค่าน้อยเกินไปทำให้เส้นขอบสีที่เกิดจากลวดลายไม่ถูกคัดกรองหรือมีค่ามากเกินไปทำให้เส้นขอบสีของวัตถุมีโอกาสถูกคัดกรองไปด้วย) เพื่อใช้ในการตรวจสอบได้ดังนี้

- ถ้า  $t_{f1}$  ค่าน้อยกว่า  $t$  แสดงว่า C1 มีค่าสีที่ใกล้เคียงกับ FG
- ถ้า  $t_{b1}$  ค่าน้อยกว่า  $t$  แสดงว่า C1 มีค่าสีที่ใกล้เคียงกับ BG
- ถ้า  $t_{f1}$  และ  $t_{b1}$  มีค่ามากกว่า  $t$  บริเวณดังกล่าวมีค่าสีที่ใกล้เคียงทั้ง FG และ BG

ทำการตรวจสอบ ค่าสีของกลุ่ม C2 ในลักษณะเดียวกันกับ C1 ซึ่งทำให้ทราบว่าสีที่ทำให้เกิดเส้นขอบนั้นมีความใกล้เคียงกับสีของวัตถุหรือพื้นหลัง โดยที่หาก C1 และ C2 มีลักษณะเหมือนกันคือ C1 และ C2 มีความใกล้เคียงกับสีของวัตถุทั้งคู่หรือมีความใกล้เคียงกับสีของพื้นหลังทั้งคู่ ทำให้สรุปได้ว่าเส้นขอบสีนั้นไม่ได้สีของภาพ ทำซ้ำขั้นตอนที่ 3.5.1 ถึง 3.5.4 ทุกตำแหน่งของ  $D_j$  กำหนดให้  $E_{ct}$  คือ เส้นขอบสีที่ผ่านการคัดกรองเส้นขอบที่ไม่เป็นขอบของวัตถุ

### 3.5.6 การคัดกรองความยาวของเส้นขอบสี

ขั้นตอนการคัดกรองความยาวของเส้นขอบสี  $E_{ct}$  เพื่อลดปัญหาเส้นขอบที่เป็นสัญญาณรบกวนเช่น เส้นขน เนื่องด้วยเส้นขอบสีควรมีความยาวต่อเนื่องในระดับหนึ่ง เพราะหากมีความยาวน้อย ทำให้ไม่สามารถใช้ในการอ้างอิงส่วนของเส้นขอบวัตถุได้แน่ชัด ดังนั้นค่าเกณฑ์ที่ใช้ในการคัดกรองความยาวของเส้นขอบต้องมีค่าไม่น้อยเกินไป และ ต้องมีค่าไม่มากเกินไปด้วยเหตุผลที่ว่า ถ้าค่าเกณฑ์ที่ใช้ในการคัดกรองความยาวของเส้นขอบมีความยาวมาก ส่งผลทำให้มีโอกาสที่เส้นขอบสีของวัตถุถูกตัดออกไปด้วย โดยงานวิจัยฉบับนี้กำหนดค่าเกณฑ์ความยาวอยู่ที่ 15 พิกเซล ซึ่งได้จากการทดลองการปรับความยาวในการคัดกรองเส้นขอบสีที่ระดับต่างๆ และกำหนดให้  $E_c$  คือ เส้นขอบสีที่ผ่านเกณฑ์

### 3.6 การสร้างเส้นขอบวัตถุ

ขั้นตอนการสร้างเส้นขอบวัตถุเพื่อนำไปสร้างภาพไทรแมบแบบอัตโนมัติจากข้อมูลเส้นขอบสีกับข้อมูลเส้นขอบระยะลึกจากเหตุผลที่กล่าวมาแล้วในหัวข้อที่ 3.1.5 ที่ว่าปัญหาในการสร้างภาพไทรแมบจากข้อมูลเส้นขอบสีมีข้อจำกัดเรื่องความสมบูรณ์ของเส้น หรือการใช้เส้นขอบระยะลึกที่ต้องกำหนดขนาดพื้นที่บริเวณไม่ชัดเจนให้ครอบคลุมความคลาดเคลื่อนของอุปกรณ์ ด้วยเหตุนี้จึงอาศัยข้อดีของเส้นขอบสีที่ให้ความถูกต้องสูง เมื่อวัดกับพื้นหลังมีสีแตกต่างกันและเส้นขอบระยะลึกที่มีความสมบูรณ์ของเส้น ในการสร้างเส้นขอบวัตถุ ด้วยวิธีการเปรียบเทียบระหว่างเส้นขอบสีกับเส้นขอบระยะลึกในแต่ละช่วงจากความเหมือนเชิงรูปร่าง และระยะทางที่มากที่สุดระหว่างเส้นขอบสีและเส้นขอบระยะลึกเพื่อนำเส้นขอบที่ผ่านการคัดเลือกมาใช้สร้างเป็นเส้นขอบวัตถุที่เหมาะสมโดยมีหลักการดังต่อไปนี้คือ

1. เปรียบเทียบเส้นขอบสีกับเส้นขอบระยะลึกในแต่ละช่วงกับค่าส่วนเบี่ยงเบนมาตรฐาน (SD) ของความเหมือนเชิงรูปร่างและระยะห่างมากที่สุดระหว่างเส้น ด้วยเหตุผลที่ว่าเส้นขอบระยะลึกให้ข้อมูลเชิงรูปร่างของวัตถุหากเส้นขอบสีมีความแตกต่างเชิงรูปร่างแสดงว่าเส้นขอบสีดังกล่าว ไม่ได้เป็นเส้นขอบวัตถุ และเส้นขอบสีที่เป็นขอบวัตถุนั้นมีระยะห่างมากที่สุดกับเส้นขอบระยะลึกไม่เกินค่าความคลาดเคลื่อนหากเกินแสดงว่าไม่ได้เป็นเส้นขอบวัตถุ ดังนั้นหากเส้นขอบสีมีความเหมือนเชิงรูปร่างและมีระยะห่างสูงสุดผ่านเกณฑ์ที่กำหนดเลือกใช้เส้นขอบสีในบริเวณนั้น
2. บริเวณที่ไม่มีเส้นขอบสีใช้เส้นขอบระยะลึกในบริเวณดังกล่าวเป็นตัวกำหนดขอบเขตบริเวณไม่ชัดเจนจากค่าความคลาดเคลื่อน
3. หากบริเวณที่เส้นขอบสีไม่ผ่านเกณฑ์ตามที่กำหนดใช้เส้นขอบระยะลึกในบริเวณดังกล่าวสำหรับการกำหนดขอบเขตบริเวณไม่ชัดเจน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับขั้นตอนการสร้างเส้นขอบวัตถุสามารถทำได้ดังต่อไปนี้

1. กำหนดให้  $E'_d$  คือ ตัวแปรตำแหน่งของเส้นขอบระยะลึกที่ใช้เปรียบเทียบค่าระยะห่างระหว่างจุด,  $E_d^m$  และ  $E_d^n$  คือตัวแปรตำแหน่งเส้นขอบระยะลึกเพื่อใช้ในการเก็บค่าตำแหน่งเส้นขอบสีที่ถูกเลือกและ  $E_d^s$  ตัวแปรตำแหน่งรูปภาพ เส้นขอบระยะลึกเพื่อใช้ในการเก็บค่าส่วนเบี่ยงเบนมาตรฐานเชิงรูปร่างของเส้นขอบสี
2. แบ่งเส้นขอบสี  $E_c$  ออกเป็นช่วง ๆ ตามลำดับความต่อเนื่อง โดยที่แต่ละช่วงมีความยาวไม่เกิน  $\frac{E}{2}$  Pixel ของเส้นขอบสีที่พบ กำหนดให้คือ  $C_{k,s} = (m_{k,s}, n_{k,s})$ ,  $k = 0, 1, 2, \dots, K-1$  และ  $s = 0, 1, 2, \dots, S-1$  โดยที่  $K$  คือ จำนวนกลุ่มของเส้นขอบสี  $E_c$  ที่แบ่งออกเป็นช่วงทั้งหมด และ  $S$  คือ จำนวนสมาชิกในแต่ละกลุ่ม
3. หาระยะห่างระหว่างจุด  $C_{k,s}$  กับตำแหน่ง  $D_j$  ใกล้ที่สุดซึ่งไม่ซ้ำตำแหน่งเดิมในกลุ่มของตัวเองมาหาค่าระยะห่างมากที่สุดของกลุ่มและค่า SD จากสมการต่อไปนี้

$$C'_{dis} = \underset{j,k,s}{\operatorname{argmin}}(|C_{k,s} - D_j|) \quad (3.10)$$

$$C'_{max} = \max(C'_{dis}) \quad (3.11)$$

$$C'_{sd} = \sqrt{\frac{1}{N_{k,s}} \sum_{k,s}^{K,S} (C'_{dis} - \mu_k)^2} \quad (3.12)$$

$$\mu_k = \frac{1}{N_{k,s}} \sum_{k,s}^{K,S} C'_{dis} \quad (3.13)$$

กำหนดให้  $C'_{dis}$  คือระยะห่างระหว่างจุด  $C_{k,s}$  กับตำแหน่ง  $D_j$  ใกล้ที่สุดซึ่งไม่ซ้ำตำแหน่งเดิมในกลุ่มของตัวเอง,  $C'_{max}$  คือค่าระยะห่างมากที่สุดของ  $C'_{dis}$  และ  $C'_{sd}$  คือ SD ของระยะห่างระหว่างจุดของ  $C_{k,s}$  กับตำแหน่ง  $D_j$

4. ในขั้นตอนที่ 3 ทราบถึงตำแหน่งจับคู่ระหว่าง  $C_{k,s}$  กับ  $D_j$  กำหนดให้คือ  $D_t = (m_t, n_t)$  โดยเก็บค่าข้อมูลที่ผ่านมาเงื่อนไขที่  $E'_d(D_t) > C'_{max}$  ได้ต่อไปนี้
  - $E'_d(D_t) = C'_{max}$  คือตัวแปรเก็บค่าระยะห่างมากที่สุดในการจับคู่ระหว่าง  $C_{k,s}$  กับ  $D_j$  บนตำแหน่ง  $E'_d$
  - $E_d^m(D_t) = m_{k,s}$  คือตัวแปรเก็บค่าตำแหน่ง  $m_{k,s}$  ของ  $C_{k,s}$  ตรงตำแหน่ง  $D_t$  ของรูปภาพ  $E_d^m$

- $E_d^n(D_t) = n_{k,s}$  คือตัวแปรเก็บค่าตำแหน่ง  $n_{k,s}$  ของ  $C_{k,s}$  ตรงตำแหน่ง  $D_t$  ของรูปภาพ  $E_d^n$
- $E_d^{sd}(D_t) = C_{sd}'$  คือตัวแปรเก็บ SD จากการจับคู่ระหว่าง  $C_{k,s}$  กับ  $D_j$  บนตำแหน่ง  $E_d^{sd}$

สำหรับการจับคู่ระหว่าง  $C_{k,s}$  กับ  $D_j$  ที่มีระยะทางสั้นที่สุดไม่ซ้ำตำแหน่งเดิมนั้นมีโอกาสเกิดช่วงที่ไม่มีคู่ระหว่าง  $C_{k,s}$  กับ  $D_j$  ดังรูปที่ 3.18 สามารถแก้ไขได้โดยการกำหนดให้มีตำแหน่งเดียวกับ  $C_{k,s}$  ที่จับคู่กับ  $D_j$  ที่มีระยะห่างมากที่สุดในกลุ่มเดียวกัน



รูปที่ 3.18 ตัวอย่างช่วงการจับคู่ของ  $C_{k,s}$  กับ  $D_j$  ที่เกิดช่วงที่ไม่มีคู่

5. ทำขั้นตอนที่ 2 ถึง 4 ซ้ำจนครบทุกกลุ่มของ  $C_{k,s}$
6. เลือกเส้นขอบวัตถุจากเส้นขอบสีหรือเส้นขอบระยะลึกตามเงื่อนไขที่  $E_d^{sd}(D_j) \leq 0.15$  และ  $E_d'(D_j) \leq E$  ถ้าเงื่อนไขเป็นจริงเลือกใช้ Color edge แทนพื้นที่บริเวณดังกล่าวได้จากตำแหน่ง  $E_d^m(d)$  และ  $E_d^n(d)$  โดยที่

$$E_d^m(d) = \{d | E_d^{sd}(D_j) \leq 0.15 \text{ และ } E_d'(D_j) \leq E\} \quad (3.14)$$

$$E_d^n(d) = \{d | E_d^{sd}(D_j) \leq 0.15 \text{ และ } E_d'(D_j) \leq E\} \quad (3.15)$$

หากเงื่อนไขเป็นเท็จเลือกใช้เส้นขอบระยะลึกแทน กำหนดให้  $E_c'$  คือ เส้นขอบวัตถุที่ได้จากการเลือกเส้นขอบสีและเส้นขอบระยะลึก

7. ทำการปรับปรุงเส้นขอบวัตถุ  $E_c'$  ด้วยเหตุผลที่ว่าเส้นขอบวัตถุเกิดจากเส้นขอบสีกับเส้นขอบระยะลึกทำให้เกิดรอยต่อระหว่างเส้นทั้งสอง โดยทำการเชื่อมจุดระหว่างเส้นทั้งสองด้วยวิธีการ Binary circular blob ของ [14] กำหนดให้  $E_{dt}$  คือเส้นขอบระยะลึกที่ถูกเลือก และ  $E_c$  คือเส้นขอบวัตถุที่ผ่านการคัดกรองและปรับปรุง

สำหรับการสร้างเส้นขอบวัตถุ สามารถเขียนเป็น pseudo code ได้ดังนี้

```

E'_d = E_d × 1000;
E_d^m = E_d;
E_d^n = E_d;
E_d^sd = E_d
for k=1:K-1
    for s=1:S-1
        for j=1:J-1
            C'_dis = abs(C_{k,s} - D_j);
            if E'_d(D_j) > C'_dis // เงื่อนไขหาจุดที่มีระยะทางน้อยที่สุด
                E'_d(D_j) = C'_dis; // เก็บข้อมูลระยะทาง
                D_t = D_j; // เก็บข้อมูลตำแหน่งเส้นขอบระยะลึก
                E_d^m(D_j) = m_{k,s} // เก็บข้อมูลตำแหน่ง m_{k,s} ที่จับคู่
                E_d^n(D_j) = n_{k,s} // เก็บข้อมูลตำแหน่ง n_{k,s} ที่จับคู่
            end
        end
    end
    C'_max = max(C'_dis) // ค่าระยะทางที่มากที่สุด
    μ_k = sum(C'_dis) / size(C'_dis) // ค่า mean ของระยะทางทั้งหมด
    C'_sd = sqrt(sum(pow(C'_dis - μ_k, 2)) / size(C'_dis)) // ค่า SD ของระยะทางทั้งหมด
    E'_d(D_t) = C'_max // เก็บค่าระยะทางที่มากที่สุดตามตำแหน่งของภาพ Depth edge
    E_d^sd(D_t) = C'_sd // เก็บค่า SD ตามตำแหน่งของภาพ Depth edge
    for j=1:J-1 // Loop การตรวจสอบการเลือก Edge
        if E_d^sd(D_j) ≤ 0.15 && E'_d(D_j) ≤ Q
            mt = E_d^m(D_j);
            nt = E_d^n(D_j);
            E_c(mt, nt) = 1;
        else
            E_dt(D_j) = 1;
        end
    end
end E_c = filledgegraps(E_c) // การเชื่อมเส้นที่เกิดช่องว่างด้วย Binary circular blob

```

รูปที่ 3.19 ตัวอย่าง Pseudo code การสร้างเส้นขอบวัตถุ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.7 การสร้างภาพไครแมบจากเส้นขอบวัตถุ

ขั้นตอนการสร้างภาพไครแมบจากข้อมูลเส้นขอบวัตถุ เนื่องจากข้อมูลเส้นขอบวัตถุเกิดจากเส้นขอบสี่และเส้นขอบระยะลึก ทำให้ในการกำหนดขอบเขตบริเวณไม่ชัดเจนสามารถแบ่งออกได้เป็น 2 กรณี คือ 1. ส่วนเส้นขอบวัตถุที่เกิดจากเส้นขอบระยะลึก และ 2 ส่วนเส้นขอบวัตถุที่เกิดจากเส้นขอบสี่หรือเส้นขอบที่เกิดจากการเชื่อมโยงเส้นระหว่างรอยต่อ ซึ่งการกำหนดขนาดบริเวณไม่ชัดเจนในส่วนของเส้นขอบระยะลึกกำหนดขนาดตามค่าความคลาดเคลื่อนของตัวอุปกรณ์ แต่สำหรับเส้นขอบวัตถุที่เกิดจากเส้นขอบสี่หรือรอยต่อพิจารณาการกำหนดขนาดบริเวณไม่ชัดเจนจากข้อมูลเส้นขนที่อยู่โดยรอบ โดยมีขั้นตอนการทำดังต่อไปนี้

1. ข้อมูลส่วนเส้นขนได้จากเส้นขอบที่ไม่ผ่านเกณฑ์ความยาวต่อเนื่องที่กำหนดในขั้นตอนที่ 3.6.5 ซึ่งข้อมูลเส้นขนที่นำมาใช้ปรับข้อมูลค่าเกณฑ์ค้นหาเส้นขอบสี่ลงครึ่งหนึ่งเพื่อเพิ่มข้อมูลส่วนที่มีโอกาสเป็นเส้นขน จึงต้องทำขั้นตอนที่ 3.4 ถึง 3.6 ใหม่อีกครั้ง กำหนดให้คือ  $E_r$
2. การกำหนดขนาดบริเวณไม่ชัดเจนอ้างอิงตามตำแหน่งของ  $E_c$  โดยการเรียงลำดับตำแหน่งตามความต่อเนื่องกำหนดให้คือ  $T_u = (m_u, n_u)$  โดยที่  $(m_u, n_u)$  คือตำแหน่งเส้นบนรูปภาพ  $E_c$  และ  $u = 1, 2, \dots, U - 1$  คือ จำนวนจุดที่พบทั้งหมดบนภาพ  $E_c$
3. ใช้กรอบ  $W$  ในการตรวจสอบบริเวณโดยรอบของตำแหน่งอ้างอิงเพื่อหาส่วนที่เกิดจากเส้นขน  $E_r$  ซึ่งในการคำนวณจำเป็นต้องแบ่ง  $T_u$  ออกเป็นช่วง ๆ โดยอ้างอิงจากลำดับของ  $T_u$  ที่หารด้วย  $Q$  ลงตัว
4. ทำการขยายเส้นขอบวัตถุจากตำแหน่ง  $T_u$  ที่อยู่ในกรอบ  $W$  โดยขยายเส้นขอบเพิ่มทีละ 1 จนถึงค่าความคลาดเคลื่อน ซึ่งในแต่ละรอบการขยายเส้นตรวจสอบค้นหาเส้นขน  $E_r$  ที่อยู่ตำแหน่งเดียวกันในกรอบ  $W$  โดยมีหลักเงื่อนไขดังนี้
  - หากตรวจสอบไม่พบเส้นขนหลังจากการขยายเส้นขอบวัตถุจนถึงค่าความคลาดเคลื่อนกำหนดขนาดบริเวณไม่ชัดเจนของเส้นขอบวัตถุที่อยู่ในกรอบ ได้จากการขยายเส้นขอบวัตถุในกรอบเท่ากับครึ่งหนึ่งของค่าความคลาดเคลื่อน
  - หากตรวจสอบพบเส้นขนจากการขยายให้ขนาดที่ตรวจสอบพบครั้งสุดท้ายมาใช้ในการกำหนดขนาดบริเวณไม่ชัดเจนของเส้นขอบวัตถุที่อยู่ในกรอบ
5. ตรวจสอบจนครบทุกตำแหน่ง  $T_u$
6. ขยายเส้นขอบวัตถุในส่วนที่ใช้เส้นขอบระยะลึก  $E_{dt}$  โดยขยายขนาดเท่ากับค่าความคลาดเคลื่อน

7. กำหนดบริเวณส่วนของพื้นหลังคือด้านนอกบริเวณไม่ชัดเจนและส่วนของวัตถุคือส่วนด้านใน

สำหรับขั้นตอนการสร้างภาพไทรแมบจากเส้นขอบวัตถุสามารถเขียนเป็น pseudo code ได้ดังนี้

```

tempEc = Ec
for u = 1:U - 1
    if u == 1 || mod(u, E) == 0
        imin = max(Tu - E, 1); // การตีกรอบ w ที่ตำแหน่ง Tu
        imax = min(Tu + Q, M);
        jmin = max(Tu - Q, 1);
        jmax = min(Tu + Q, N);
        wc = Ec(iMin:iMax, jMin:jmax);
        wr = Er(iMin:iMax, jMin:jmax);
        diTemp = checkDilate(wc, wr); // function ตรวจสอบขนาดที่ใช้ในการขยาย
    wd = imdilate(wc, diTemp); // ขนาดที่ขยาย
    tempEc(iMin:iMax, jMin:jmax) = wd; // ตำแหน่งที่ขยายเส้น
    end
end
tempEr = imdilate(Er, Q); // ส่วนการขยายเฉพาะเส้นขอบระยะลึก
unknown = tempEr + tempEc;
trimap = createdTri(unknown);

```

รูปที่ 3.20 ตัวอย่าง Pseudo code ขั้นตอนการสร้างภาพไทรแมบ

## บทที่ 4

### ผลการทดลอง

#### 4.1 วัตถุประสงค์ของการทดลอง

วัตถุประสงค์ของการทดลองเพื่อทดสอบเทคนิคการสร้างภาพไตรแมมแบบอัตโนมัติที่พัฒนาขึ้น โดยสามารถแบ่งออกเป็น 2 ส่วนดังต่อไปนี้

1. เพื่อทดสอบความถูกต้องของการตรวจจับเส้นขอบวัตถุจากภาพระยะใกล้ด้วยวิธีการที่เสนอ โดยใช้การเปรียบเทียบกับเส้นขอบวัตถุที่กำหนดด้วยมือโดยผู้ใช้งาน และวิธีการที่อาศัยข้อมูลภาพระยะใกล้ที่มีอยู่ก่อนแล้ว (ภาพที่ 4.1 แสดงตัวอย่างเส้นขอบวัตถุที่กำหนดโดยผู้ใช้งาน)



รูปที่ 4.1 ตัวอย่างเส้นขอบวัตถุที่กำหนดโดยผู้ใช้งาน

2. เพื่อทดสอบประสิทธิภาพของภาพไตรแมมที่สร้างขึ้นด้วยวิธีการที่เสนอ โดยเปรียบเทียบความถูกต้องกับภาพไตรแมมที่สร้างด้วยผู้ใช้งาน (Trimap-U) ในการทดสอบ มีวัตถุประสงค์เพื่อเปรียบเทียบผลกระทบของพื้นหลังที่แตกต่างกันว่าส่งผลอย่างไรต่อประสิทธิภาพในการสร้างภาพไตรแมมด้วยวิธีที่เสนอ

## 4.2 อุปกรณ์ที่ใช้ในการทดลอง

- เครื่องคอมพิวเตอร์ PC Intel core i7 3.4 gb
- กล้องถ่ายภาพ Kinect ที่ความละเอียด  $640 \times 480$  pixel
- กล้องถ่ายภาพ DSLR Cannon รุ่น E550D ที่ความละเอียด  $2592 \times 1728$  pixel
- อุปกรณ์ประกอบฉากในการถ่ายภาพ
- โปรแกรมจับภาพและปรับค่า [8]
- โปรแกรม Matlab

## 4.3 รายละเอียดและผลการทดลอง

### 4.3.1 การทดลองเปรียบเทียบความถูกต้องของการตรวจจับเส้นขอบวัตถุ

ในการทดสอบความถูกต้องของการตรวจจับเส้นขอบวัตถุด้วยวิธีการที่เสนอ ได้ทำการเปรียบเทียบความถูกต้องกับเส้นขอบที่กำหนดด้วยมือ และเส้นขอบที่ตรวจจับด้วยวิธีการอื่น ๆ ที่มีอยู่ในปัจจุบัน ดังต่อไปนี้

1. วิธีตรวจจับเส้นขอบวัตถุจากข้อมูลระยะลึกที่ปรับปรุงด้วยวิธีการเฉลี่ยแบบถ่วงน้ำหนักจากข้อมูลระยะลึกข้างเคียง (Object Edge Detection with Averaging Filter เรียกวิธีการนี้ย่อ ๆ ว่า OED-AF)
2. วิธีตรวจจับเส้นขอบจากข้อมูลระยะลึกที่ผ่านการปรับปรุงด้วยวิธีการ Bilateral Filter (Object Edge Detection with Bilateral Filter เรียกวิธีการนี้ย่อ ๆ ว่า OED-BF)
3. วิธีตรวจจับเส้นขอบจากข้อมูลระยะลึก (Object Edge Detection with Depth เรียกวิธีการนี้ย่อ ๆ ว่า OED-D)
4. วิธีตรวจจับเส้นขอบจากข้อมูลภาพสีและข้อมูลระยะลึกที่เสนอในงานวิจัยนี้ (Object Edge Detection with Color and Depth image เรียกวิธีการนี้ย่อ ๆ ว่า OED-CD)

สำหรับข้อมูลภาพในการทดลองเพื่อใช้ในการเปรียบเทียบความถูกต้องของเส้นขอบวัตถุกับภาพลักษณะดังต่อไปนี้

1. ภาพที่สีของวัตถุแตกต่างอย่างชัดเจนกับสีของพื้นหลังที่ไม่มีลวดลาย (มีค่าความแตกต่างของสีเฉลี่ยมากกว่า 85% หาได้จากค่าผลต่างเฉลี่ยหารด้วยค่าสีเฉลี่ยของวัตถุ) ดังตัวอย่างในรูปที่ 4.2 จากเหตุผลที่ว่าสีของวัตถุและพื้นหลังมีสีแตกต่างกันนั้นสามารถค้นหาขอบของภาพวัตถุได้ชัดเจนซึ่งมีความถูกต้องมากกว่าเส้นขอบระยะลึก ทำสามารถสร้างภาพภาพไตรแมบที่มีขนาดบริเวณไม่ชัดเจนได้แคบ



รูปที่ 4.2 ตัวอย่างภาพวัตถุที่มีสีแตกต่างกับภาพพื้นหลังชัดเจน

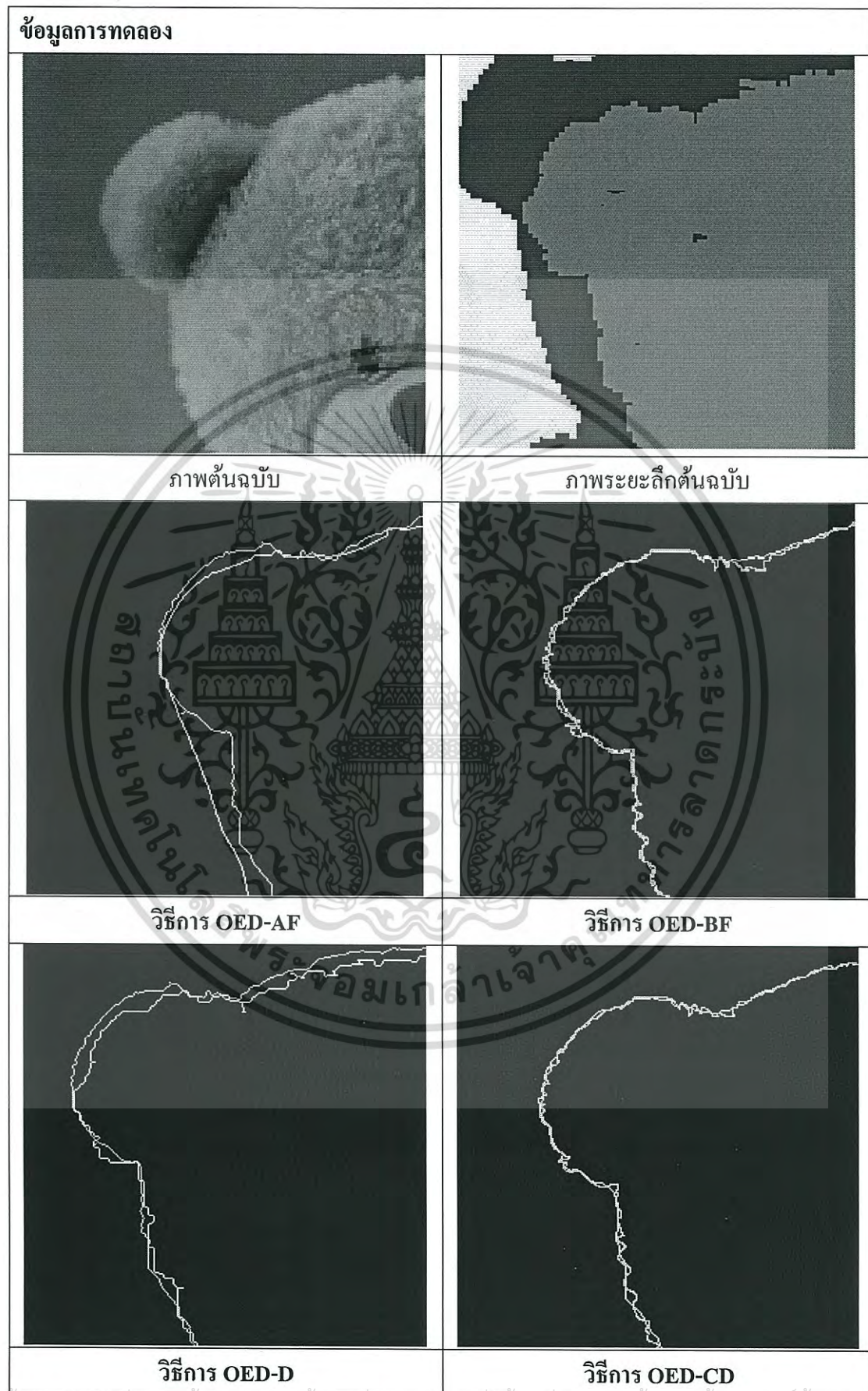
2. ภาพที่สีของวัตถุใกล้เคียงกับสีของพื้นหลังที่ไม่มีลวดลาย (มีค่าความแตกต่างของสีเฉลี่ยน้อยกว่า 15% หาได้จากค่าผลต่างเฉลี่ยหารด้วยค่าสีเฉลี่ยของวัตถุ) ดังตัวอย่างในรูปที่ 4.3 ด้วยเหตุผลที่ว่าภาพลักษณะดังกล่าวนี้ยากต่อการใช้ข้อมูลภาพสีในการค้นหาเส้นขอบวัตถุ ซึ่งมีผลต่อการสร้างภาพไตรแมบที่อาศัยข้อมูลภาพสี เพราะไม่สามารถคัดแยกวัตถุกับพื้นหลัง จึงส่งผลต่อการสร้างภาพไตรแมบ



รูปที่ 4.3 ตัวอย่างภาพวัตถุที่มีสีแตกต่างกับภาพพื้นหลังชัดเจนและไม่มีลวดลาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 ตัวอย่างการทดลองเพื่อหาเส้นขอบวัตถุ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลในการทดลองขั้นตอนที่ 4.3.1 ประกอบไปด้วยภาพวัตถุจำนวน 6 รูปที่มีสีพื้นหลังแตกต่างกับตัววัตถุและภาพวัตถุกับพื้นหลังที่มีสีคล้ายคลึงกันจำนวน 6 รูป รวมทั้งหมด 12 ภาพ

#### 4.3.1.1 การทดลองเส้นขอบวัตถุเมื่อภาพวัตถุและพื้นหลังมีสีแตกต่างกัน

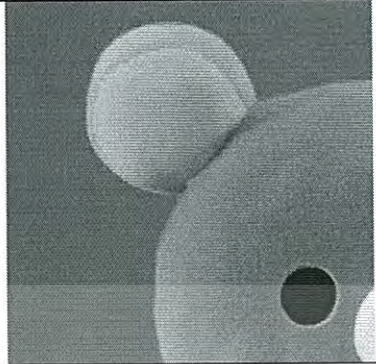
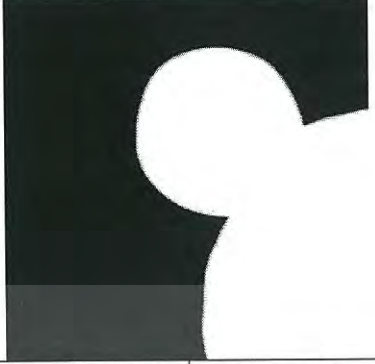




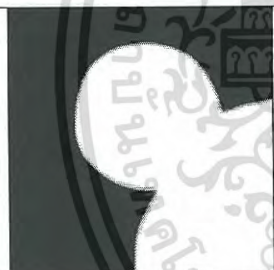



การทดลองนี้ เป็นการเปรียบเทียบผลการตรวจจับเส้นขอบวัตถุเมื่อภาพสีของวัตถุและพื้นหลังมีความแตกต่างกัน โดยอาศัยเส้นขอบวัตถุที่สร้างด้วยมือเป็นข้อมูลอ้างอิง (Ground truth) ค่าความคลาดเคลื่อนหาได้จากการวัดระยะห่าง (หน่วยพิกเซล) จากแต่ละจุดบนเส้นขอบวัตถุที่ตรวจจับได้ไปยังจุดบนเส้นขอบวัตถุที่สร้างด้วยมือที่ใกล้ที่สุด จากนั้นนำระยะห่างที่ได้ทั้งหมดมาหาค่าเฉลี่ย เป็นตัวเลขที่ใช้สะท้อนความถูกต้องของการตรวจจับเส้นขอบวัตถุ โดยต่อไปจะเรียกค่าเฉลี่ยดังกล่าวนี้ว่า  $\bar{E}_{oe}$

ตารางที่ 4.2 ผลการทดลองระยะระหว่างเส้นขอบเมื่อสีวัตถุแตกต่างกับสีพื้นหลัง

วิธีการตรวจจับเส้นขอบวัตถุ	รูปที่ 1	รูปที่ 2	รูปที่ 3	รูปที่ 4	รูปที่ 5	รูปที่ 6	$\bar{E}_{oe}$
OED-AF	5.39	9.60	8.43	8.58	7.82	7.68	7.916
OED-BF	1.21	3.66	1.02	4.97	2.89	0.81	2.446
OED-D	5.26	7.37	4.71	4.82	3.65	5.29	5.183
OED-CD	4.92	5.70	4.2	2.78	3.65	4.24	4.253

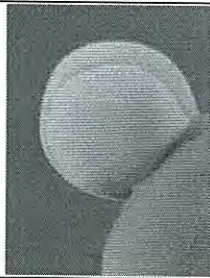








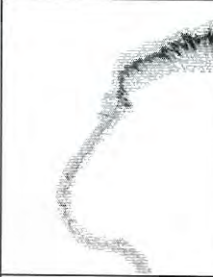





จากผลการทดลองการตรวจจับเส้นขอบ ที่สีของภาพวัตถุและพื้นหลังมีความแตกต่างกันข้างต้นพบว่า ในกรณีนี้วิธี OED-BF ให้ผลที่ดีที่สุด (มีค่า  $\bar{E}_{oe}$  ต่ำสุด) และเมื่อนำเส้นขอบทั้ง 4 วิธี ไปสร้างภาพไทรแมบและภาพแมทได้ตามตารางที่ 4.3

ตารางที่ 4.3 ตัวอย่างภาพไทรแมบและภาพแมทที่ภาพวัตถุมีสีแตกต่างกับภาพพื้นหลัง

ภาพสี		ภาพแมทจริง	
			
OED-AF	OED-BF	OED-D	OED-CD
			
ภาพ ไทรแมบ			
			
ภาพแมท			

จากตารางที่ 4.3 การทดลองตรวจจับเส้นขอบที่สีของภาพวัตถุและภาพพื้นหลัง มีความแตกต่างกัน พบว่าขอบเขตของภาพไทรแมบมีความแตกต่างกัน ซึ่งเมื่อนำ ข้อมูลภาพไทรแมบมาสร้างภาพแมทด้วยวิธีการซ้อนภาพแบบโคลสฟอร์ม (Closed form matting) พบว่าภาพแมทที่ได้มีความคล้ายคลึงกับภาพแมทจริง โดยตาราง ที่ 4.4 แสดงส่วนของภาพแมทแต่ละวิธีที่แตกต่างกับภาพแมทจริงที่สีวัตถุแตกต่างกับ พื้นหลัง

ตารางที่ 4.4 ส่วนของภาพเมทแต่ละวิธีที่แตกต่างกับภาพเมทจริงเมื่อสีวัตถุแตกต่างกับพื้นหลัง

Color image	OED-AF	OED-BF	OED-D	OED-CD
				
				
				

#### 4.3.1.2 การทดลองเส้นขอบเมื่อภาพวัตถุและพื้นหลังมีสีความคล้ายคลึงกัน

การทดลองนี้ เป็นการเปรียบเทียบผลการตรวจจับเส้นขอบวัตถุเมื่อภาพสีของวัตถุและพื้นหลังไม่แตกต่างกัน โดยอาศัยเส้นขอบวัตถุที่สร้างด้วยมือเป็นข้อมูลอ้างอิง (Ground truth) ค่าความคลาดเคลื่อนหาได้จากการวัดระยะห่าง (หน่วยพิกเซล) จากแต่ละจุดบนเส้นขอบวัตถุที่ตรวจจับได้ไปยังจุดบนเส้นขอบวัตถุที่สร้างด้วยมือที่ใกล้ที่สุด จากนั้นนำระยะห่างที่ได้ทั้งหมดมาหาค่าเฉลี่ย เป็นตัวเลขที่ใช้สะท้อนความถูกต้องของการตรวจจับเส้นขอบวัตถุ

ตารางที่ 4.5 ผลการทดลองระยะระหว่างเส้นขอบเมื่อสี่วัตถุกับสี่พื้นหลังมีความคล้ายคลึงกัน

วิธีการค้นหาเส้นขอบวัตถุ	รูปที่ 1	รูปที่ 2	รูปที่ 3	รูปที่ 4	รูปที่ 5	รูปที่ 6	$\bar{E}_{oe}$
OED-AF	5.39	9.60	8.43	8.58	7.82	7.68	7.916
OED-BF	5.13	8.08	6.60	6.9	6.05	9.82	7.101
OED-D	5.26	7.37	4.71	4.82	3.65	5.29	5.183
OED-CD	5.04	5.89	4.52	2.95	3.65	4.92	4.495

ผลการทดลองเปรียบเทียบระยะห่างระหว่างเส้นขอบเมื่อสี่ของภาพวัตถุกับสี่ของภาพพื้นหลังมีความคล้ายคลึงกันพบว่า วิธีการที่ใช้เฉพาะภาพระยะลึกให้ผลเหมือนเดิม เพราะไม่ได้มีการนำข้อมูลภาพสี่เข้ามาเกี่ยวข้องในการคิดคำนวณ สำหรับวิธี OED-BF นั้นพบว่ามีผลคลาดเคลื่อนมากกว่ากรณีแรก อันเป็นผลมาจากการที่สี่ของวัตถุและพื้นหลังใกล้เคียงกัน (ดูรูปที่ 4.4 ประกอบ) ขณะที่วิธีการ OED-CD ที่เสนอได้ผลที่ดีกว่า โดยผลที่ได้ใกล้เคียงกับกรณีของการทดลองในข้อ 4.3.1.1 แสดงให้เห็นว่าวิธีการที่เสนอ มีคุณสมบัติที่ดีกว่าในแง่ที่ว่าค่าสี่พื้นหลังที่แตกต่างกันส่งผลกระทบต่อประสิทธิภาพการทำงานน้อยกว่าวิธี OED-BF และเมื่อนำเส้นขอบทั้ง 4 วิธีที่สี่ของภาพวัตถุและพื้นหลังมีความคล้ายคลึงกันมาสร้างภาพไตรแมบและภาพแมทได้ตามตารางที่ 4.6

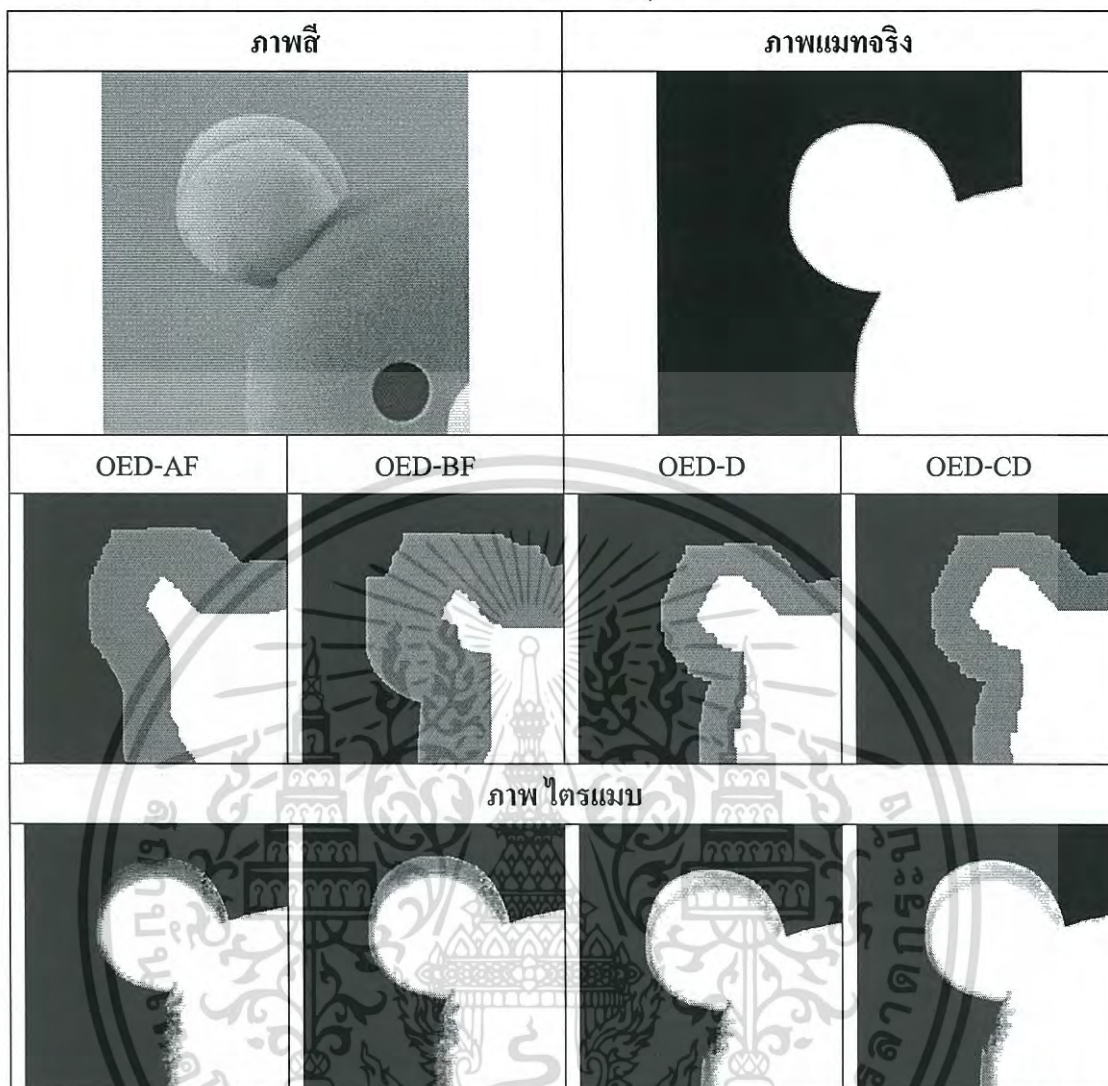


(ก)

(ข)

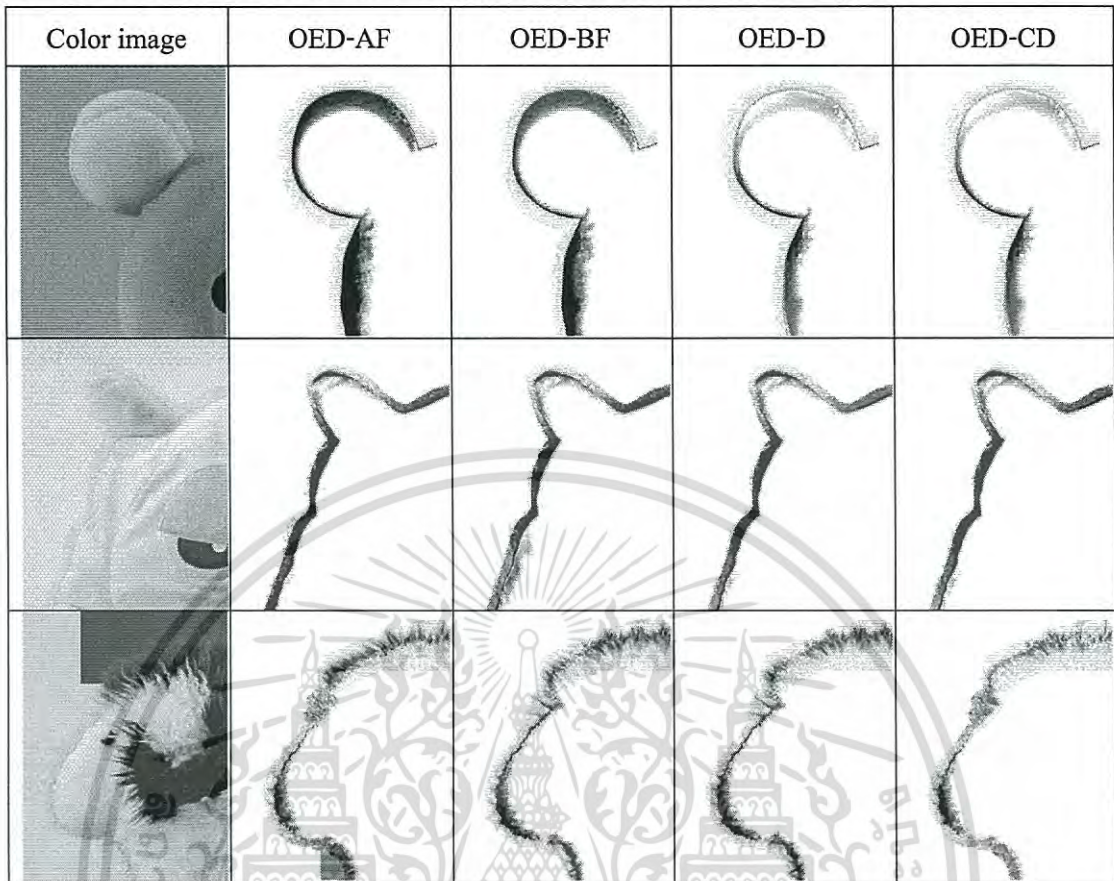
รูปที่ 4.4 ภาพตัวอย่างที่สี่ของวัตถุและพื้นหลังใกล้เคียงกันทำให้วิธี OED-BF

ตารางที่ 4.6 ตัวอย่างภาพไทรแมบและภาพแมทที่ภาพวัตถุมีสีคล้ายคลึงกับภาพพื้นหลัง



จากตารางที่ 4.6 การทดลองตรวจจับเส้นขอบที่สีของภาพวัตถุและภาพพื้นหลัง มีความคล้ายคลึงกัน พบว่าขอบเขตของภาพไทรแมบมีความแตกต่างกัน และเมื่อสร้างภาพแมทด้วยวิธีการซ้อนภาพแบบโคลสฟอร์ม (Close form matting) พบว่าภาพแมทมีความผิดพลาดเมื่อเทียบกับภาพแมทจริง โดยตารางที่ 4.7 แสดงส่วนของภาพแมทแต่ละวิธีที่แตกต่างกับภาพแมทจริงที่สีวัตถุคล้ายคลึงกับพื้นหลัง

ตารางที่ 4.7 ส่วนของภาพแมทแต่ละวิธีที่แตกต่างกับภาพแมทจริงเมื่อสวิตดูคล้ายคลึงกับพื้นหลัง



สรุปผลการทดลองเปรียบเทียบความถูกต้องของการตรวจจับเส้นขอบวัตถุ คือ เส้นขอบวัตถุจากวิธีการ OED-BF เหมาะกับภาพวัตถุกับพื้นหลังมีค่าสีที่ความแตกต่างกัน แต่เกิดปัญหาในการตรวจจับเส้นขอบวัตถุเมื่อภาพวัตถุกับพื้นหลังมีสีใกล้เคียงกัน ในขณะที่วิธีการที่เสนอ OED-CD ให้ผลเฉลี่ยจากทุกกรณีที่ดีที่สุด ( $E_{oe}$  เฉลี่ยจากทุกกรณีเท่ากับ 4.495) เมื่อเปรียบเทียบกับวิธีการอื่น ๆ ที่นำมาทดสอบ

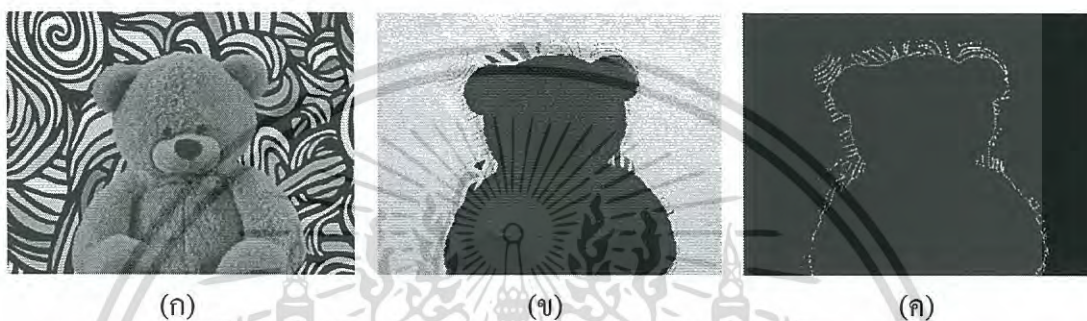
#### 4.3.2 การทดสอบประสิทธิภาพในการกำหนดขอบเขตของภาพไทรแมบ

การทดลองนี้เป็นการทดสอบประสิทธิภาพในการกำหนดขอบเขตของภาพไทรแมบที่สร้างขึ้นจากข้อมูลภาพสีและข้อมูลระยะลึก (Trimap-CD) โดยการเปรียบเทียบกับภาพไทรแมบที่สร้างด้วยมือ (Trimap-U) ซึ่งขั้นตอนนี้เปรียบเทียบการกำหนดขอบเขตของภาพไทรแมบกับภาพพื้นหลังลักษณะต่อไปนี้คือ

1. ภาพพื้นหลังที่มีลวดลาย (Texture background)
2. ภาพพื้นหลังที่ไม่มีลวดลาย (Plain background) โดยสีของภาพพื้นหลังมาจากค่าสีเฉลี่ยของภาพพื้นหลังที่มีลวดลาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับในการทดลองเพื่อวัดประสิทธิภาพและความถูกต้องการกำหนดขอบเขตของภาพไทรแมบ วิธีการ OED-BF ไม่ถูกนำมาทดลองด้วยเหตุผลที่ว่า วิธีการ OED-BF พบปัญหาการตรวจจับเส้นขอบวัตถุ เมื่อภาพพื้นหลังของวัตถุมีลวดลายหรือมีสีใกล้เคียงกัน เพราะวิธีการ OED-BF ใช้ข้อมูลภาพสีในการปรับปรุงข้อมูลภาพระยะลึก ทำให้เมื่อภาพพื้นหลังมีสีใกล้เคียงกับวัตถุหรือบริเวณที่ปรับปรุง ไม่มีสีที่ตรงกับวัตถุหรือพื้นหลัง ทำให้การปรับปรุงภาพระยะลึกเกิดความผิดพลาดเกิดเป็นเส้นขอบของลวดลาย ซึ่งส่งผลต่อการค้นหาเส้นขอบวัตถุ ดังรูปที่ 4.5



รูปที่ 4.5 ตัวอย่างปัญหาการใช้ Bilateral filter ตรวจจับเส้นขอบวัตถุ-(ก) ภาพสี (ข) ภาพระยะลึก ปรับปรุงด้วย OED-BF (ค) เส้นขอบวัตถุ








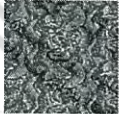

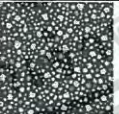

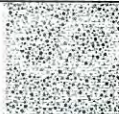














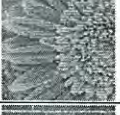









เพื่อตรวจสอบประสิทธิภาพความคงทนต่อการเปลี่ยนแปลงในแบบต่าง ๆ โดยภาพไทรแมบที่ได้จากภาพพื้นหลังทั้งสองแบบต้องมีขอบเขตบริเวณไม่ชัดเจนไม่แตกต่างกันมาก เพราะถูกสร้างมาจากภาพวัตถุชนิดเดียวกัน หากแตกต่างกันมากแสดงว่า Trimap-CD ไม่เหมาะกับการนำมาใช้สร้างภาพไทรแมบ สำหรับการคำนวณตรวจสอบความถูกต้องของบริเวณไม่ชัดเจนที่ตำแหน่งตรงกันกับภาพ Trimap-U คิคำนวณเป็นเปอร์เซ็นต์ตรงกันดังตารางที่ 4.8

ตารางที่ 4.8 แสดงลักษณะผลการคำนวณทดสอบประสิทธิภาพในการกำหนดขอบเขต

ภาพไครแมบโดยผู้ใช้			
รูปภาพสี	ภาพไครแมบ	บริเวณกำหนดถูกต้อง	บริเวณกำหนดผิด
			
			

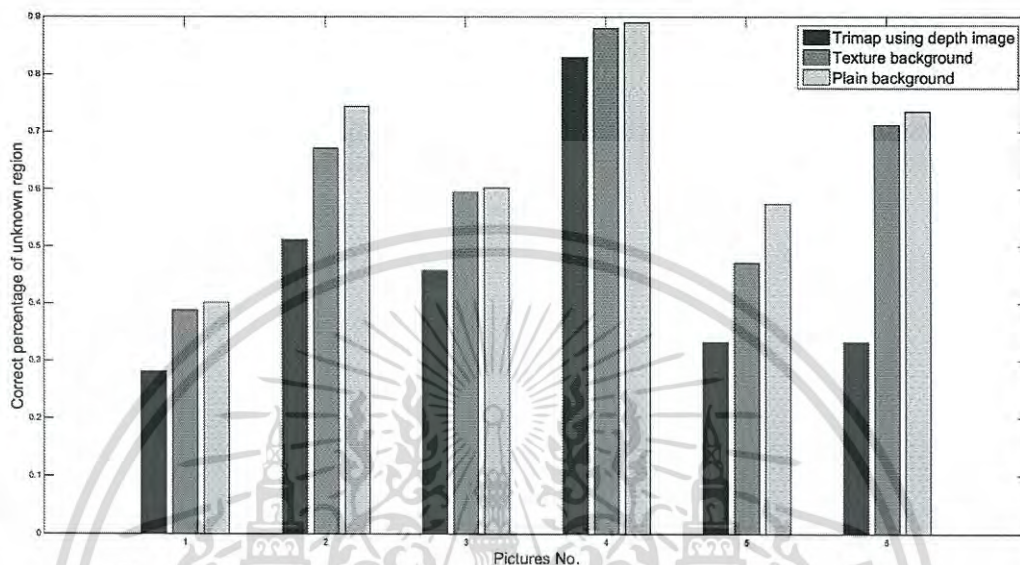
ชุดข้อมูลรูปภาพในการทดลองประกอบไปด้วยภาพวัตถุจำนวน 6 ชนิดที่มีพื้นหลังแตกต่างกัน 30 แบบรวมภาพทั้งหมด 180 ภาพตามตารางที่ 4.9

ตารางที่ 4.9 ข้อมูลรูปภาพวัตถุและภาพพื้นหลังที่ใช้ในการทดลอง

ภาพวัตถุ					
					
ภาพพื้นหลัง					
					
					
					
					
					

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลองวัดประสิทธิภาพการกำหนดขอบเขตของ Trimap-CD ระหว่างภาพพื้นหลังที่มีลวดลายและไม่มีลวดลาย ให้ผลที่ใกล้เคียงกันไม่แตกต่างกันมาก ดังรูปที่ 4.6 และมีความถูกต้องมากกว่า Trimap-D จึงสรุปได้ว่า Trimap-CD มีความคงทนต่อสภาพการเปลี่ยนแปลงของภาพพื้นหลัง



รูปที่ 4.6 กราฟผลการทดลองส่วนความถูกต้องในการกำหนดขอบเขตของภาพไตรแมบ

### 4.3.3 การวัดประสิทธิภาพความถูกต้องของภาพแมท

การทดลองวัดผลประสิทธิภาพความถูกต้องของภาพแมทที่คำนวณด้วยวิธีการ Closed form matting จากภาพไตรแมบที่สร้างขึ้นทั้ง 3 วิธีคือ

1. ภาพไตรแมบโดยผู้ใช้ ( User defined trimap หรือสามารถเรียกย่อ ๆ ว่า Trimap-U) คือ ภาพไตรแมบที่ผู้ใช้ระบุขอบเขตของวัตถุ , พื้นหลัง และบริเวณไม่ชัดเจน
2. ภาพไตรแมบโดยใช้ข้อมูลระยะลึก ( Trimap using depth image หรือสามารถเรียกย่อ ๆ ว่า Trimap-D) คือ ภาพไตรแมบที่สร้างจากข้อมูลระยะลึกซึ่งปรับปรุงตามลักษณะของปัญหาที่เกิดขึ้นมาระบุขอบเขตของวัตถุ , พื้นหลัง และบริเวณไม่ชัดเจน
3. ภาพไตรแมบโดยใช้ข้อมูลภาพสีและข้อมูลภาพระยะลึก ( Trimap using color and depth image หรือสามารถเรียกย่อ ๆ ว่า Trimap-CD) คือ ภาพไตรแมบที่สร้างจากข้อมูลระยะลึกซึ่งปรับปรุงตามลักษณะของปัญหาที่เกิดขึ้นและข้อมูลภาพสีมาระบุขอบเขตของวัตถุ , พื้นหลัง และบริเวณไม่ชัดเจน

โดยการวัดประสิทธิภาพความถูกต้องของภาพแมทจากค่า Root mean square error (RMSE) เมื่อเทียบกับภาพแมทที่แท้จริง ซึ่งค่า RMSE เป็นตัวบ่งบอกความผิดพลาดของภาพแมทที่เกิดขึ้น สามารถทำได้ตามสมการที่ 4.2

$$RMSE = \sqrt{\frac{\sum_{y=0}^Y (R_y - R'_y)^2}{Y}} \quad (4.2)$$

กำหนดให้  $R_y$  คือ ภาพแมทที่แท้จริงและ  $R'_y$  คือ ภาพแมทที่ได้จากการคำนวณ โดยที่  $y$  คือค่าในพิกเซลของรูปภาพซึ่งมาค่าอยู่ระหว่าง 0 ถึง 1 โดยการทดลองวัดประสิทธิภาพความถูกต้องของภาพแมทแบ่งออกเป็น 2 ส่วนดังต่อไปนี้

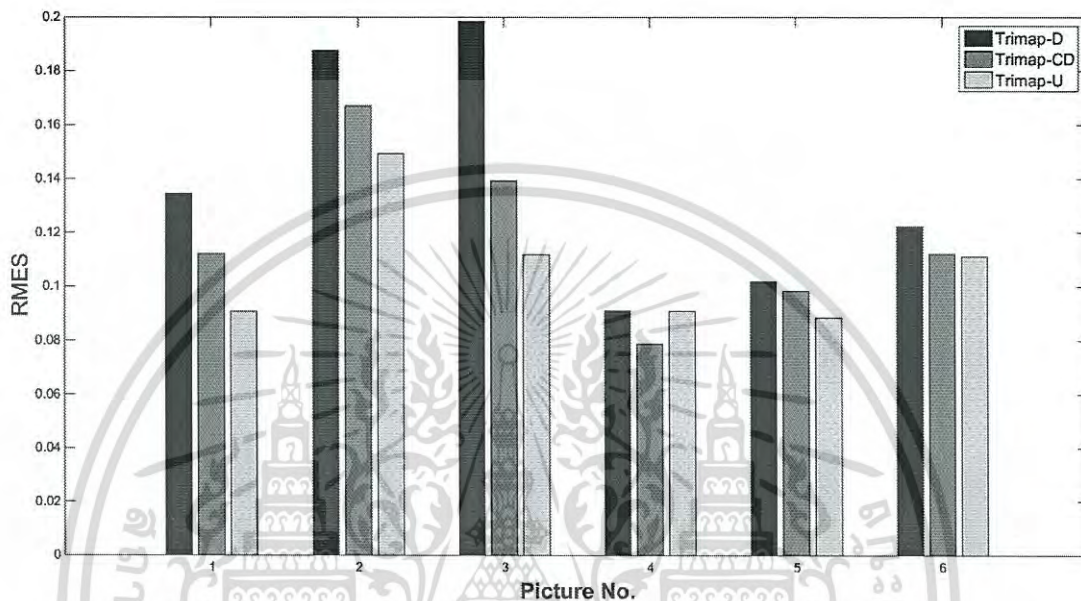
1. การทดลองเปรียบเทียบความถูกต้องของภาพแมทที่มาจากภาพไม่มีลวดลาย (Plain Background) ทั้ง 2 ลักษณะ คือ 1) ภาพพื้นหลังที่มีสีแตกต่างกับภาพวัตถุ และ 2) ภาพพื้นหลังที่มีสีคล้ายคลึงกับภาพวัตถุ เพื่อทดสอบปัจจัยของขนาดบริเวณไม่ชัดเจนส่งผลต่อความถูกต้องในการคำนวณหาภาพแมทกับภาพที่ไม่มีลวดลายหรือไม่ได้ตามตารางที่ 4.10

ตารางที่ 4.10 ผลการทดลองภาพพื้นหลังแบบไม่มีลวดลาย

ลักษณะภาพพื้นหลังไม่มีลวดลาย	RMSE จากภาพไครแมบ		
	Trimap-D	Trimap-CD	Trimap-U
1. ภาพพื้นหลังที่มีสีคล้ายคลึงกับภาพวัตถุ	0.231	0.212	0.149
2. ภาพพื้นหลังที่มีสีแตกต่างกับภาพวัตถุ	0.197	0.140	0.132

จากตารางที่ 4.10 ผลการทดลองความถูกต้องของภาพแมทสามารถสรุปได้ว่าขนาดบริเวณไม่ชัดเจนส่งผลต่อความถูกต้องในการคำนวณภาพแมท เมื่อภาพวัตถุและภาพพื้นหลังมีสีที่ใกล้เคียงกันแม้ว่าภาพพื้นหลังไม่มีลวดลาย แต่ส่งผลน้อยถ้าภาพวัตถุและภาพพื้นหลังมีสีที่แตกต่างกัน

2. การทดลองเปรียบเทียบความถูกต้องของภาพแมทจาก Trimap-D , Trimap-CD และ Trimap-U เมื่อภาพมีลักษณะเป็นลวดลายและมีสีที่หลากหลาย โดยอาศัยค่า RMSE ในการวัดผล เพื่อทดสอบความแตกต่างระหว่าง Trimap-D , Trimap-CD และ Trimap-U ดังรูปที่ 4.7



รูปที่ 4.7 ผลการทดลองเปรียบเทียบค่า RMSE ของภาพวัตถุกับพื้นหลังแบบลวดลายทั้ง 3 วิธี

ผลทดลองความถูกต้องของภาพแมทจาก Trimap-D , Trimap-CD และ Trimap-U พบว่าเมื่อเปรียบเทียบค่า RMSE ของ Trimap-U ให้ผลดีที่สุดเพราะการกำหนดขอบเขตโดยผู้ใช้ของ Trimap-U ทำให้สามารถสร้างภาพ Trimap ได้เหมาะสมที่สุดใกล้เคียงกับภาพวัตถุ แต่ในการสร้าง Trimap-U ผู้ใช้ต้องมีความเชี่ยวชาญในการกำหนดขอบเขต และการที่กำหนดขอบเขตโดยผู้ใช้จึงเกิดความไม่สะดวกล่าช้าในการสร้าง ด้วยเหตุนี้ทำให้ Trimap-CD ที่มีผลค่า RMSE ใกล้เคียงกับ Trimap-U โดยมีค่าความต่างเฉลี่ยกับ Trimap-U อยู่ที่ 0.0216 ที่ภาพพื้นหลังมีลวดลาย ทำให้ Trimap-CD เหมาะกับการนำมาใช้ในการสร้างภาพไตรแมบ

## บทที่ 5

### สรุปผลการทดลอง

จากบทที่ 4 ได้แสดงผลการทดสอบเทคนิคการตรวจจับเส้นขอบวัตถุและการสร้างภาพไตรแมมแบบอัตโนมัติโดยใช้ข้อมูลภาพสีกับข้อมูลระยะลึก สามารถสรุปผลการทดลองได้

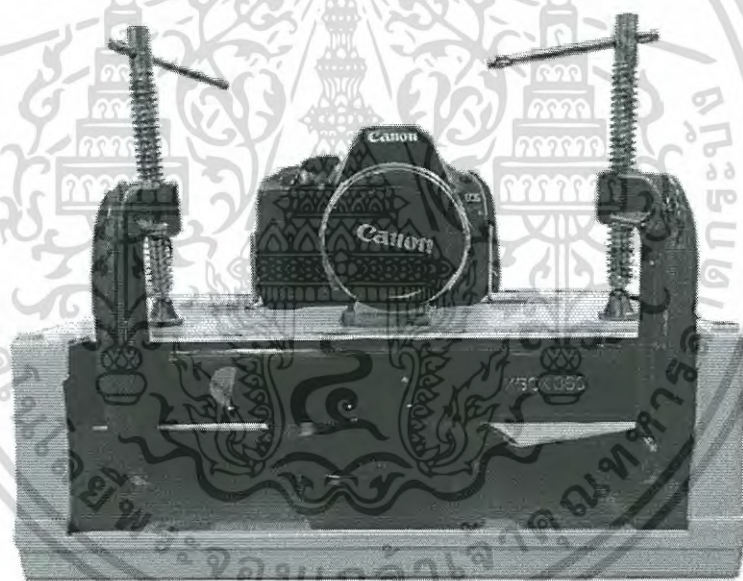
#### 5.1 สรุปผลการทดลอง

1. สำหรับผลทดลองเปรียบเทียบความถูกต้องของเส้นขอบวัตถุ เพื่อตรวจสอบความถูกต้องก่อนนำไปใช้ในการอ้างอิงตำแหน่งวัตถุพบว่า OED-BF ให้ผลดีที่สุด โดยมีค่า  $E_{oe} = 2.446$  เมื่อเทียบกับเส้นขอบภาพวัตถุ ในกรณีที่ภาพวัตถุกับภาพพื้นหลังมีสีที่แตกต่างกันและพื้นหลังไม่มีลวดลาย แต่ไม่เหมาะกับภาพวัตถุและภาพพื้นหลังที่มีสีคล้ายคลึงกัน ทำให้ OED-BF เหมาะกับกรณีที่มีการจัดเตรียมวัตถุกับพื้นหลัง แต่สำหรับ OED-CD มีความยืดหยุ่นในการใช้งานกับรูปภาพที่มีความหลากหลายของสีหรือภาพที่มีลวดลาย อีกทั้งให้ผลที่ใกล้เคียงกับ OED-BF เมื่อภาพวัตถุกับพื้นหลังมีสีใกล้เคียงกัน
2. ผลทดสอบประสิทธิภาพการกำหนดขอบเขตของภาพ Trimap-CD กับภาพพื้นหลังที่มีลวดลายและไม่มีลวดลาย มีความถูกต้องเมื่อนำไปตรวจสอบกับ Trimap-U มีความแตกต่างกันที่ 3.84% ในขณะที่ Trimap-D มีความแตกต่างอยู่ที่ 15.8 % แสดงว่าวิธีการ Trimap-CD มีความคงทนต่อความหลากหลายของลวดลายพื้นหลังมากกว่า
3. ผลทดสอบวัดผลประสิทธิภาพความถูกต้องของภาพแมท จาก Trimap-D, Trimap-CD และ Trimap-U พบว่า Trimap-U เป็นวิธีการที่ให้ผลที่ดีที่สุดในการสร้างภาพไตรแมมที่เหมาะสม ในขณะที่ Trimap-CD เป็นวิธีการสร้างภาพไตรแมมแบบอัตโนมัติที่ได้ผลดีที่สุด โดย Trimap-CD มีความแตกต่างจากกับวิธีการ Trimap-U อยู่ที่ 2.16%

## 5.2 การติดตั้งอุปกรณ์และการนำไปใช้งาน

การติดตั้งอุปกรณ์ใช้กล้อง DSLR สำหรับถ่ายภาพสี เพราะมีความละเอียดและความคมชัดสูง และกล้อง Kinect สำหรับถ่ายภาพระยะลึกระหว่างวัตถุกับพื้นหลัง โดยการติดตั้งจำเป็นต้องยึดตำแหน่งกล้องทั้ง 2 ตัวไม่ให้เคลื่อนที่ ดังรูปที่ 5.1 เพื่อลดปัญหาความคลาดเคลื่อนในการนำไปใช้หลังการปรับค่า (Calibration) โดยสามารถปรับค่ากล้องทั้ง 2 ตัวให้ตรงกันตามงานวิจัย [8]

สำหรับการนำไปใช้งาน เนื่องด้วยวิทยานิพนธ์ฉบับนี้ มีการปรับปรุงข้อมูลภาพถ่ายระยะลึกที่ได้จากกล้อง Kinect ที่อาศัยหลักการ Structured light ในการคำนวณข้อมูลระยะลึก ดังนั้นการใช้งานกับข้อมูลภาพระยะลึกด้วยวิธีการอื่น สามารถทำได้โดยการตัดขั้นตอน การปรับปรุงภาพระยะลึก (Depth image enhancement) เพราะขั้นตอนดังกล่าวเป็นการปรับปรุงข้อมูลภาพระยะลึกจากวิธีการ Structured light ดังนั้นจึงสามารถนำไปใช้งานกับข้อมูลภาพระยะลึกที่ได้จากวิธีการอื่นได้



รูปที่ 5.1 การจัดตำแหน่งของกล้อง DSLR และกล้อง Kinect

### 5.3 ข้อจำกัดและปัญหาที่พบ

1. ปัญหาการปรับค่า (Calibration) กล้อง Kinect กับกล้อง DSLR คือ การในปรับค่าของกล้องทั้ง 2 ตัวให้ตรงกัน จำเป็นต้องกำหนดตำแหน่งของกล้องที่แน่นอน ไม่มีการปรับเปลี่ยนไปจากเดิมหลังการปรับค่า เพราะตำแหน่งที่คลาดเคลื่อน ทำให้เกิดความผิดพลาดเมื่อซ้อนภาพระยะลึกกับภาพสี ดังนั้นแนวทางการแก้ไข คือ การกำหนดตำแหน่งที่มีความแน่นอนเพื่อป้องกันความคลาดเคลื่อนที่เกิดขึ้น ดังรูปที่ 5.1
2. ข้อจำกัดและปัญหาความโปร่งใสของวัตถุ เช่น ขวดน้ำพลาสติก เพราะไม่สามารถตรวจสอบหาข้อมูลระยะลึกของตัววัตถุ เนื่องจากไม่มีการสะท้อนของแสงอินฟราเรด ทำให้ไม่สามารถตรวจสอบหาตัววัตถุได้
3. ข้อจำกัดในการตัดแยกวัตถุกับพื้นหลังเมื่อมีระยะลึกหลายระดับ เนื่องด้วยงานวิทยานิพนธ์ฉบับนี้อาศัยสมมุติฐานที่ว่า กลุ่มข้อมูลระยะลึกที่ใกล้ที่สุดคือวัตถุเบื้องหน้าที่ต้องการ ดังนั้นถ้ามีวัตถุที่มีระยะลึกน้อยกว่าจะส่งผลทำให้เกิดความผิดพลาดในการตัดแยก เพราะวัตถุดังกล่าวถูกระบุเป็นวัตถุที่ต้องการแทน

### 5.4 แนวทางการทำวิจัยต่อ

สำหรับแนวทางการทำวิจัยต่อคือ การวิจัยเรื่อง การตัดแยกวัตถุกับพื้นหลังเมื่อมีระยะลึกหลายระดับ หรือการระบุวัตถุโปร่งแสง ด้วยเหตุผลที่ว่าตัววัตถุในภาพสามารถมีได้หลายชั้น และมีระยะลึกแตกต่างกันได้หลายระดับ หากสามารถตัดแยกระบุตำแหน่งของวัตถุที่ต้องการได้ จะทำให้สามารถใช้หลักการสร้างภาพไตรแมปด้วยข้อมูลภาพสีและข้อมูลระยะลึกร่วมกันของงานวิทยานิพนธ์ฉบับนี้ได้

## เอกสารอ้างอิง

- [1] A.R. Smith and J.F. Blinn, "Blue screens matting," in Proc. of Siggraph 96, pp. 259-268, 1996.
- [2] C. Yung-Yu, "Bayesian approach to digital matting," Computer vision and pattern recognition CVPR, pp. 264-271, 2001.
- [3] O. Wang et al, "Automatic natural video matting with depth", Computer graphics and applications, 2007. PG '07. 15th Pacific Conference on, pp. 469 – 472, 2007
- [4] A. Levin, D. Lischinski, and Y. Weiss "A closed-form solution to natural image matting" IEEE transactions on pattern analysis and machine intelligence, vol. 30, pp 1-15, 2008
- [5] F. Pitié and A. Kokaram , "Matting with a depth map" Image Processing (ICIP), 2010 17th IEEE International Conference on , pp. 21 – 24 ,2010
- [6] E. Mirante, M. Georgiev, and A. Gotchev, "A fast image segmentation algorithm using color and depth map", 3DTV Conference: The True Vision - Capture, pp. 1-4, 2011.
- [7] W. Sun et al., "Adaptive depth map assisted matting in 3D video" IEEE Multimedia and Expo (ICME), page 1-6, 2011
- [8] H.C Daniel., K. Juho, and H. Janne, "Joint depth and color camera calibration with distortion correction", TPAMI, pp 1-8, 2012.
- [9] E. Mirante et al., "A Fast image segmentation algorithm using color and depth map", Pattern Recognition (ICPR), IEEE, pp. 3787 – 3790, 2012.
- [10] J.H. Cho et al., "Enhancement of Depth Maps With Alpha Channel Estimation for 3-D Video", Selected Topics in Signal Processing, IEEE, pp. 483 – 494, 2012
- [11] J. Zhu et al., "Joint Depth and Alpha Matte Optimization via Fusion of Stereo and Time-of-Flight Sensor"
- [12] L. Chen, H. Lin and S. Li, "Depth image enhancement for Kinect using region growing and bilateral filter" Pattern Recognition (ICPR), pp. 3070-3073, 2012
- [13] L. Ting and L. Shutao, "Image Matting with Color and Depth Information" Pattern Recognition (ICPR), pp. 787 – 3790, 2012
- [14] A. S. Jalal and C. Bhatanagar, "Automatic Trimap and Alpha-Matte Generation For Digital Image Matting", Contemporary computing (IC3), pp. 202-208, 2013.
- [15] P. Kovesi , "Centre for Exploration Targeting" The University of Western Australia 2013



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ภาคผนวก

## บทความและผลงานวิจัยที่ได้รับการตีพิมพ์

M.Tamiyakul and N.Chotikakamthorn “Automatic Trimap Generation for Digital Image Matting Using a Consumer RGB-D Camera” The 29th International Technical Conference on Circuit/Systems Computers and Communications (ITC-CSCC), Phuket, Thailand, July 1-4, 2014



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2<sup>nd</sup> Call for Papers

# ITC-CSCC 2014

The 29<sup>th</sup> International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2014)

With the great success of the International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC) as the world leading conference devoted to the advancement of high technologies in Circuits/Systems, Computers and Communications, we would like to invite all the scholars and experts around the world to attend the 29<sup>th</sup> ITC-CSCC 2014 to be hosted in Phuket, Thailand.

**Topics**  
The conference is open to researchers from all regions of the world. Participation from Asia Pacific region is particularly encouraged. Proposals for special sessions are welcome. Papers with original work in all aspects of Circuits/Systems, Computers and Communications are invited. Topics include, but not limited to, the followings

**Circuits & Systems**

- Analog Circuits
- Computer Aided Design
- Semiconductor Devices & Technology
- VLSI Design
- Medical Electronics & Circuits
- RF Circuits
- Sensors & Related Circuits
- Power Electronics & Circuits
- Wireless Power Transmission
- Intelligent Transportation Systems & Technology
- Linear / Nonlinear Systems
- Robotics and Modern Control
- Neural Networks
- Verification & Testing

**Computers**

- Artificial Intelligence
- Bio Computing
- Computer Vision
- Computer Graphics and Virtual Reality (VR) Technology
- Face Detection & Recognition
- Image Coding & Analysis
- Image Processing
- Internet Technology & Applications
- Computer Systems & Applications
- Motion Analysis
- Multimedia Sensors & Technology
- Object Extraction & Technology
- Security
- Watermarking

**Communications**

- Antenna & Wave Propagation
- Audio / Speech Signal Processing
- Circuits & Components for Communications
- MIMO & Space-Time Codes
- IP Networks & QoS
- Mobile & Wireless Communications
- Network Management & Design
- Multimedia Communications
- Optical Communications & Components
- Radar / Remote Sensing
- Communication Signal Processing
- Underwater Communications
- Ultra-Wideband (UWB)
- Visual Communications
- Wireless Sensor Networks
- Ubiquitous Networks
- Radio Over Fiber Technology (ROF)
- Internet Protocol Television (IPTV)

**Submission of Papers**  
Prospective authors are invited to submit original paper(s) in either MS Word or PDF format written in English. Abstract is limited to two pages of text and figures. Abstract can be submitted on the official website. If you have any trouble in paper preparation and online submission, please contact the conference secretariat.

**Proceedings and Publications**  
All registered manuscripts will be published in conference proceedings. Full-length manuscripts are encouraged for submission to associated Journals: IEJE Journal (Korea), IEICE Transactions (Japan), ECTI Transactions (Thailand).

**Engineering Journal (Thailand) and Songklanakorn Journal of Science and Technology (SIST)**. The authors for their faculties are requested to pay the publication charge for the IEICE Transactions after manuscript acceptance.

**Important Deadlines**

Submission of Special Session Proposal:	March 1, 2014
Submission of Two-Page Extended Abstract:	April 1, 2014
Notification of Acceptance:	May 8, 2014
Submission of Camera Ready Paper:	June 1, 2014

**Contact:** secretary@itc-cscc2014.org  
**Official Website:** <http://www.itc-cscc2014.org>

**International Advisory Committee**

- Athikom Rucksabutr
- Booncharoen Sirinonvaval
- Jung Wha Chung
- Do Hyun Kim
- Hang Gu Bahk
- Hitashi Yamada
- Isao Shirakawa
- Jae Hong Lee
- Prabhat Chonratitwaitana
- Jin-Koo Rhee
- Jin Ok Park
- Kenji Onaga
- Kosin Channongchai
- Kukjin Chun
- Masao Nakagawa
- Makayuki Kawamura
- Monai Kraiakh
- Rokuya Ishii
- Satechi Goto
- Sawasdi Tantaratana
- Seong Dae Kim
- Seung Hong Hong

**International Coordination Committee**

- Byung Gook Park
- Chaodit Aswakul
- Cheon Won Choi
- Chiranut Sa-ngiamsak
- Daesik Hong
- Daichakorn Tancharoen
- Euisung Kang
- Hiroshi Tamami
- HyunWook Park
- Joosik Park
- Jun Heo
- Ken-ichi Nakano
- Lanchakorn Wuttisittikuljij
- Monkazu Nakamura

**Organizing Committee**

**General Chair**

- Chiranut Sa-ngiamsak
- General Co-Chair
- Nagsisa Ishiura
- Yong Seo Koo

**TPC Chair**

- Lanchakorn Wuttisittikuljij

**TPC Co-Chair**

- Kimhiko Hiratahi
- Jun Heo

**TPC Secretary**

- Paat Vanichchanant
- Daichakorn Tancharoen

**General Secretary**







- Piya Kovintavewat

**Special Session Chair**

- Keerasak Sripitumawat
- Local Arrangement Chair
- Chaiyachet Salivichit
- Chaodit Aswakul
- Jakapan Suaboot
- Phaphak Sirisak
- Wannarat Soriamorint
- Waradon Weerapan
- Publicity Chair
- Phongsak Keeratwintakorn
- Publication Chair
- Niithita Chirdeboi
- Industry Relation Chair
- Jitkaset Ngarmit
- Pivul Chamkeikong
- Finance Chair
- Rajjapin Sampanna
- Sirilock Suthachai
- Registration Chair
- Pichaya Tundayya
- Sasiphan Wongsubhavas
- Wannaree Wongtrairat

**July 1-4, 2014**

**Phuket Graceland Resort and Spa, Phuket**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# Automatic Trimap Generation for Digital Image Matting Using a Consumer RGB-D Camera

Man Tamiyakul

Faculty of Information Technology  
King Mongkut's Institute of Technology Ladkrabang  
Bangkok, Thailand  
E-mail tamiyakul@gmail.com

Nopporn Chotikakamthorn

Faculty of Information Technology  
King Mongkut's Institute of Technology Ladkrabang  
Bangkok, Thailand  
E-mail nopporn@it.kmitl.ac.th

**Abstract** This paper presents a method to automatically generate a trimap image from RGB-D data for digital matting with natural background. The method addresses the problem of missing depth data from a Kinect camera, due to IR light occlusion. The method automatically identifies the missing data area caused by light occlusion, and assigns appropriate depth value to it. In addition, the paper presents an approach to determine the width of the unknown region in a trimap image. The width of the region is set according to the depth accuracy of the RGB-D camera. Given the generated trimap image, the result based on the Bayesian-based matting technique is shown and compared with the case where a trimap image was manually created.

**Keywords:** Digital matting, natural background, RGB-D camera, Trimap, image segmentation

## I. INTRODUCTION

The objective of digital image matting for natural background is to separate foreground objects from real objects behind as appeared in still images or digital videos. Most methods achieve this by requiring rough segmentation of an image into three non-overlapping regions of foreground objects, background objects, and the foreground-background boundary. The area comprising the pixels that are at the foreground-background boundary is also known as an unknown region. The image representing such segmentation is commonly called a trimap (Fig. 1). The accuracy of the unknown region defines the quality of the trimap. In turn, the trimap quality is critical to digital matting. It can be created in several ways such as determining by hand, or making use of continuation of object borders, or depth information. [1-4].

This paper presents a method to automatically create the trimap by using depth information from a Kinect device. The device is a consumer RGB-D camera which uses structured light to calculate depths of objects in the scene. Kinect is an affordable device. However, it has two problems. The first is the incompleteness of depth data. The second problem is on accuracy of the measured depth value. In [5], a solution to the missing data problem is proposed. However, [5] requires data collection for an analysis (training) step.

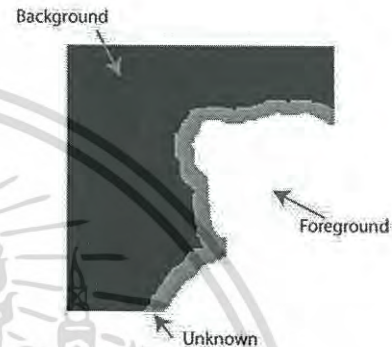


Fig. 1 Example of a trimap image.

This paper presents an approach to automatically identify the missing data area caused by Kinect's IR light occlusion, and assigns appropriate depth value to it. In addition, the paper presents an approach to determine the width of the unknown region in a trimap.

## II. DIGITAL MATTING WITH AN RGB-D CAMERA

Many digital matting techniques require a trimap, which separates the original image into three regions as described before. By using the RGB-D camera, the available depth information can be used to determine foreground-background object boundary image. This paper uses a calibrated Kinect camera [6] for trimap image determination and digital matting. An overview of a matting algorithm is described as follows [7] (see Fig. 2).

1. Take a picture data by an RGB-D camera. The data consists of an optical image and a depth image. Let  $I_C(m, n)$  be the optical image and  $I_D(m, n)$  be the depth image. Let  $m, n$  be the row and column of each pixel in the images,  $m = 0, 1, \dots, M-1$ , and  $n = 0, 1, \dots, N-1$ . Here,  $M$  is the image height and  $N$  is its width.
2. Generate a trimap from depth data by using the edge of an object caused by change in depth, to separate a foreground object from a background object.
3. Perform matting by using the generated trimap.

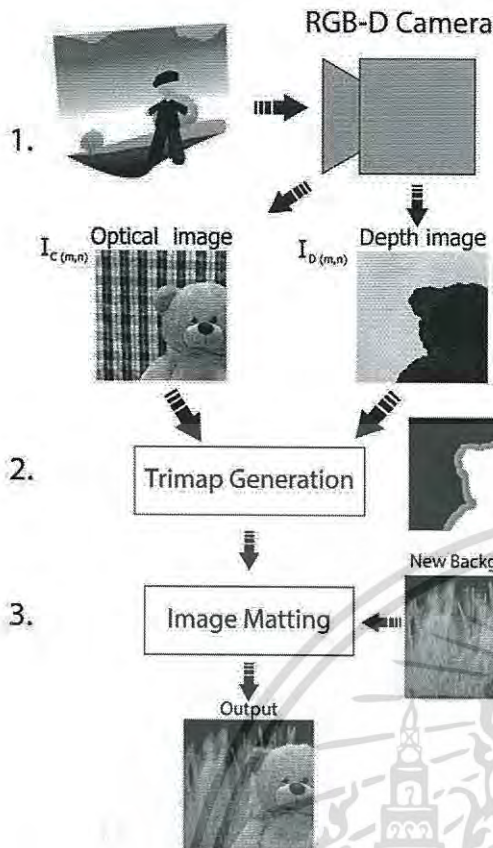


Fig. 2. Digital matting process

Generating a trimap using the depth data from a Kinect camera has two problems. These problems are

1. Depth data is usually incomplete. It is caused by two problems. The first problem is due to non-reflective object. For the second one, some infrared lights are obscured by a foreground object (see Fig. 3). Therefore, depth of the area inside the shadow caused by the occlusion cannot be measured.

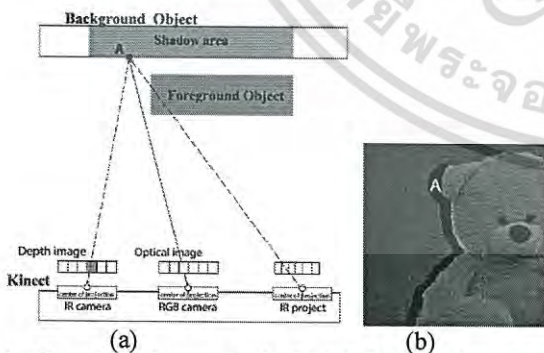


Fig 3 (a) The point A does not receive IR light because of occlusion by a foreground object. (b) An optical image overlay by a depth image showing the missing depth data area due to light occlusion.

2. Depth image taken by Kinect has discrepancy of depth value from the actual one. This discrepancy affects generation of a trimap from depth image data. The discrepancy causes the boundary between foreground and background regions that are determined by the depth image to differ from the boundary between foreground and background regions revealed in the optical image. From Fig. 4, the discrepancy of depth value causes some parts of the foreground-region boundary (that is determined by the depth image) to cover the background region in the optical image. From the same figure, the discrepancy of depth value causes some part of the background region (that is determined by the depth image) to overlap with the foreground region in the optical image. Therefore the width of the unknown area in the trimap must be appropriately determined. The width of unknown region must be determined to cover the discrepancy area.

Fig. 4 Incorrect foreground-background segmentation caused by the discrepancy of depth value.

The next section describes how to solve the two problems as mentioned above.

### III. PROPOSED TRIMAP GENERATION METHOD

#### 3.1) Determination of depth from missing data area

- 1) Divide the depth image into two parts. The first part is the area of the foreground object. The second part is the area of the background object. The division relies on the comparison of each pixel's depth value with a threshold value. The threshold value is determined using rough information on the depth of the foreground and background objects.
- 2) Define  $F$  as a set of pixel positions in the foreground area and  $B$  as a set of pixel positions in the background area.
- 3) For each row  $m$ , search for the segment  $I_d(m, n)$ ,  $n' < n < n''$ , which meets the following conditions (see Fig. 5)
  - a)  $I_d(m, n)$  are missing-data pixels (i.e., with the value NaN when read from a Kinect device).
  - b)  $(m, n') \in B$
  - c)  $(m, n'') \in F$

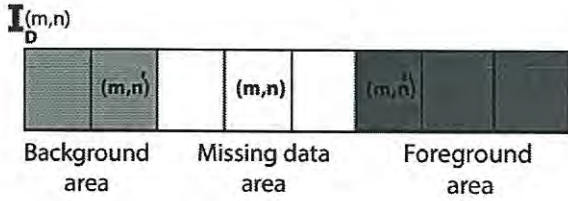


Fig. 5 Example of missing-depth pixels inside a shadow area.

4) Repeat Step 3 for all rows within the depth image. Let  $S$  be a set of  $J$  tuples  $(m_j, n'_j, n''_j)$  corresponding to groups of pixels that meet the three conditions in Step 2. Here,  $n'_j$  and  $n''_j$  are the next pixels on the left and right sides of  $I_d(m_j, n_j)$ .

5) Follow the steps below to check whether each element  $(m_j, n'_j, n''_j)$  of  $S$  corresponds to a cast shadow area.

5.1) Define  $\vec{P}_w(m_j, n'_j)$  as the position in the world coordinate of  $I_d(m_j, n'_j)$ . In addition, define  $\vec{P}_w(m_k, n''_k)$  as a point in the world coordinate of  $I_d(m_k, n''_k)$ ,  $k = 0, 1, 2, \dots, J-1$ . Define  $\vec{P}_R$  as the position of the IR projector in the world coordinate ( $\vec{P}_R = [0.05 \ 0 \ 0]^T [8]$ ).

5.2) Divide the line between  $\vec{P}_w(m_j, n'_j)$  and  $\vec{P}_R$ . Let  $\vec{P}_j(i)$  where  $i=0, 1, 2, \dots, K-1$ , be the  $K$  points between  $\vec{P}_w(m_j, n'_j)$  and  $\vec{P}_R$  (see Fig. 6).

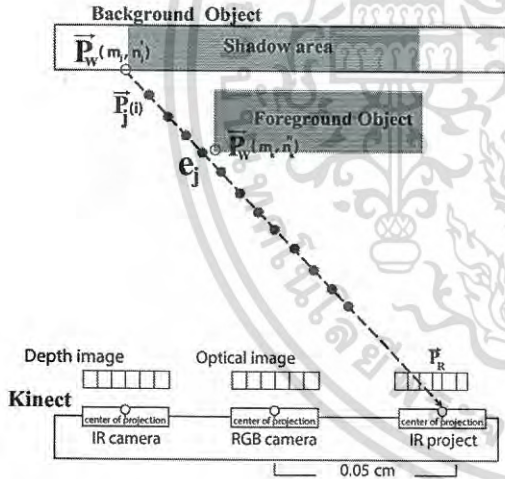


Fig. 6 Dividing the line between  $\vec{P}_w(m_j, n'_j)$  and  $\vec{P}_R$

5.3) Compute the distance between all possible pairs of  $\vec{P}_j(i)$  and  $\vec{P}_w(m_k, n''_k)$ . Then, find the pair with minimum distance using the following equations.

$$e_j(i) = \min_{k=0, \dots, J-1} \|\vec{P}_j(i) - \vec{P}_w(m_k, n''_k)\| \quad (1)$$

$$e_j = \min_{i=0, \dots, K-1} e_j(i) \quad (2)$$

5.4) Let  $t$  be the threshold used to detect the shadow area. If  $e_j$  is less than  $t$ , it is concluded that  $I_d(m_j, n_j)$  is in the shadow area. In this case,  $I_d(m_j, n_j)$  is assigned the value of  $I_d(m_j, n'_j)$ . But if  $e_j$  is greater than  $t$ , it is concluded that  $I_d(m_j, n_j)$  is not in the shadow area. The above steps are designed to use in case of only one object in foreground.

### 3.2) Trimap creation by using depth image information

The discrepancy of the depth value from a Kinect device affects determination of an appropriate width for the unknown region in a trimap as discussed above. It is proposed here that the width of the unknown region is obtained based on Kinect's depth error data. The method is briefly explained below.

- 1) Measure Kinect's depth error for each of the 3D points sampled over the working space. Store the measured depth error data in a lookup table.
- 2) From the taken depth picture, measure the averaged distance between the foreground objects and a camera. Use the obtained value to get a camera depth error from the lookup table, which is subsequently used to calculate a required width of the unknown region.

## IV. EXPERIMENTAL RESULTS

A Kinect camera was used to get the depth images at a resolution of  $640 \times 480$ . Color images were taken at a resolution of  $2592 \times 1728$  using a DSLR camera which was attached to the Kinect camera. The DSLR camera was used instead of a Kinect's color camera because of its better quality and resolution. The experiment was divided into two parts.

1). In the first part of the experiment, we calculate the accuracy of the proposed method for the detection of missing data area caused by IR light occlusion. The detection accuracy is calculated by comparing the detected missing data area with the actual missing data area. The actual missing data area is determined by human inspection. The calculation steps are explained below.

1. Define the actual missing data area by human inspection.
2. Detect the missing data pixels due to IR light occlusion by the method in Section 3.1.
3. Calculate the area of the difference between the missing data area, obtained from our method, with the actual missing data area.

By using five sample images, the average percentage of the area of the difference with respect to the missing data area is 6.74%. All images were taken at 1.5 m from the foreground object.

2). In the second part of the experiment, we examine the accuracy of trimap generation using our method. The detection accuracy is calculated by comparing the detected unknown region with the actual missing data area. The actual unknown region is determined by human inspection. The calculation steps are explained below.

1. Locate the unknown region by human inspection.
2. Detect the unknown region by the method in Section 3.2.
3. Calculate the area of the difference between the unknown region, obtained from our method, with the actual unknown region.

By using five sample images, the average percentage of the area of the difference with respect to the unknown region is 15.34%. All images were taken at 1.5 m from the foreground object.



Fig 8 A trimap based on camera depth accuracy data.

Fig. 9 compares the Bayesian-based matting result using the generated trimap and the result using the manually created trimap.

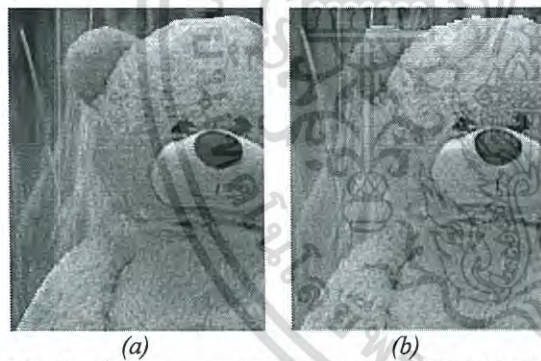


Fig. 9 Results of Bayesian matting (a) a manually determined trimap (b) automatic trimap determination

## V. CONCLUSION

This paper presents a method to automatically generate a trimap image from RGB-D data for digital matting with natural background. Our method can resolve the problem of incompleteness of depth data and inaccuracy of the measured depth value by approximating depth data in shadow area and automatically generating trimap. The method using the proposed trimap generation yields a comparable matting result to that of a manually generated trimap.

## REFERENCES

- [1] Mark A. Ruzon and Carlo Tomasi, "Alpha Estimation in Natural Images", Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on , 2000, pp. 18-25
- [2] Yung-Yu Chuang, Brian Curless and David H. Salesin "A Bayesian Approach to Digital Matting" Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on , 2001 , pp 264 -271
- [3] Olivier Juan and Renaud Keriven, "Trimap Segmentation for Fast and User-Friendly Alpha Matting", IEEE Workshop on Variational, Geometric & Level Set Methods, 2005, pp. 186-197
- [4] Oliver Wang, Jonathan Finger , et all. "Automatic Natural Video Matting with Depth", Computer Graphics and Applications, 2007. PG '07. 15th Pacific Conference on , 2007, pp. 469-472
- [5] Ju Shen and Sen-ching Cheung, "Layer Depth Denoising and Completion for Structured-Light RGB-D Cameras", 2013 IEEE Conference on Computer Vision and Pattern Recognition, , 2013 , pp. 1063-6919
- [6] Herrera C , Juho Kannala and Janne Heikkila "Joint depth and color camera calibration with distortion correction" Pattern Analysis and Machine Intelligence, IEEE Transactions on , 2012 , pp. 2058-2064
- [7] Ting Lu and Shutao Li "Image matting with color and depth information" Pattern Recognition (ICPR), International Conference on , 2012, pp 3787-3790
- [8] Jan Smisek, Michal Jancosek and Tomas Pajdla "3D with Kinect" Computer Vision Workshops (ICCV Workshops), IEEE International Conference on 2011, pp. 1154-1160