รายงานการวิจัยฉบับสมบูรณ์

การพัฒนาเทคนิคการแยกพื้นหลังธรรมชาติโดยใช้ข้อมูลจากภาพ RGBD

Development of a Natural Background Segmentation Technique from an RGBD Image

ผู้วิจัย

รศ.ดร.นพพร โชติกกำธร

นาย แมน เตมียกุล

ได้รับทุนสนับสนุนงานวิจัยจากเงินรายได้ ประจำปีงบประมาณ 2557

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายงานการวิจัยฉบับสมบูรณ์

การพัฒนาเทคนิคการแยกพื้นหลังธรรมชาติโดยใช้ข้อมูลจากภาพ RGBD

Development of a Natural Background Segmentation Technique from an RGBD Image

ผู้วิจัย

รศ.ดร.นพพร โชติกกำธร

นาย แมน เตมียกุล

ได้รับทุนสนับสนุนงานวิจัยจากเงินรายได้ ประจำปีงบประมาณ 2557

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

12785 74X

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดโครงงานวิจัย

ชื่อโครงการ (ภาษาไทย) การพัฒนาเทคนิคการแยกพื้นหลังธรรมชาติโดยใช้ข้อมูลจากภาพ RGBD

แหล่งเงิน เงินรายได้ คณะเทคโนโลยีสารสนเทศ สจล.

ประจำปีงบประมาณ 2557 จำนวนเงินที่ได้รับการสนับสนุน 50,000 บาท

ระยะเวลาทำการวิจัย 1 ปี ตั้งแต่ 1 ตุลาคม 2556 ถึง 30 กันยายน 2557

ชื่อ – สกุล หัวหน้าโครงการ และผู้ร่วมโครงการวิจัยพร้อมระบุ หน่วยงานต้นสังกัด และอีเมล์

1.รศ.ดร.นพพร โชติกกำธร คณะเทคโนโลยีสารสนเทศ สจล.

Email : nopporn@it.kmitl.ac.th

2.นาย แมน เตมียกุล

Email : tamiyakul@gmail.com

บทคัดย่อ

เอกสารฉบับนี้เป็นการรายงานผลการศึกษาและพัฒนาเทคนิคการแยกพื้นหลังธรรมชาติที่อาศัยข้อมูลระยะลึกที่ได้จากการถ่ายภาพด้วยกล้องถ่ายภาพประเภท RGB-D มาช่วยในการคำนวณภาพแมท (matt image) ซึ่งเป็นส่วนประกอบที่สำคัญในการซ้อนภาพ (matting) เพื่อแทนภาพพื้นหลังเดิมด้วยภาพพื้นหลังใหม่ โดยได้ทำการศึกษารูปแบบการใช้ประโยชน์จากข้อูลระยะลึกใน 2 ลักษณะ ได้แก่ การใช้ภาพระยะลึกเพื่อการปรับปรุงคุณภาพของเทคนิคการซ้อนภาพที่มีอยู่ในปัจจุบัน และการใช้ภาพระยะลึกในการสร้างแผนภาพไตรแมบ (trimap) อันเป็นข้อมูลที่จำเป็นต้องใช้ในเทคนิคที่สามารถใช้กับพื้นหลังธรรมชาติได้ ในส่วนที่หนึ่ง ได้มีการศึกษาเทคนิคต่าง ๆ ที่มีปรากฏในปัจจุบัน โดยแบ่งออกเป็น 2 กลุ่ม ตามวิธีการหลักที่ใช้ในการซ้อนภาพ ได้แก่ เทคนิคในกลุ่มที่พัฒนามาจากวิธีการซ้อนภาพด้วยหลักการของเบย์ (Bayesian-based matting) และกลุ่มที่พัฒนามาจากเทคนิคการซ้อนภาพด้วยสมการโคลสฟอร์ม (Closed-form matting) ในการศึกษา ได้ทำการเปรียบเทียบประสิทธิภาพของวิธีการทั้ง 2 กลุ่ม เปรียบเทียบกับวิธีการเดิมที่ไม่ได้มีการใช้ข้อมูลระยะลึกเข้ามาเป็นส่วนประกอบในการคำนวณ โดยอาศัยภาพจากฐานข้อมูลมาตรฐาน และภาพจากการบันทึกโดยใช้กล้อง Microsoft Kinect สำหรับในส่วนของการใช้ภาพระยะลึกในการสร้างแผนภาพไตรแมบ ในงานวิจัยนี้ได้ทำการศึกษาปัญหาความคลาดเคลื่อนและไม่สมบูรณ์ของภาพระยะลึกที่ได้จากกล้อง Kinect และได้ทำการพัฒนาวิธีการปรับปรุงคุณภาพของภาพระยะลึก เพื่อประโยชน์ในการสร้างแผนภาพไตรแมบที่มีความถูกต้องมากขึ้น อันจะส่งผลให้คุณภาพของการคำนวณภาพแมทดีขึ้นตามไปด้วย วิธีการดังกล่าวได้ถูกนำไปทดสอบกับภาพระยะลึกที่ได้จากการถ่ายภาพด้วยกล้อง Kinect

คำสำคัญ (Keywords) : Digital matting, Alpha channel, Depth sensor, Kinect, image processing

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Research Title: Development of a Natural Background Segmentation Technique from an RGBD Image

Researcher: 1. Nopporn Chotikakamthorn

2. Man Tamiyakul

Faculty: Information Technology **Department:** Information Technology

Abstract

This report describes a result from a study and development of digital matting techniques for natural background separation, using depth data obtained from an RGB-D camera. A study was focused on applying depth data to assist in computing the matt image, which plays an important role in image matting; a process to replace an original background with a new one. The study considered uses of depth data for two purposes. Firstly, use of depth data for improving the matting result of existing techniques was studied. Next, the study considered the use of depth data for automatic generation of a trimap image, which is required as part of the input to the matting process. The first part of the study considered existing techniques that make use of depth images. Those can be divided into two groups based on the matting computation approach employed. The first one consists of depth-assisted techniques which are based on the Bayesian-based matting method. The second group of depth-assisted techniques is based on the closed-form matting method. Performance of these methods were studied through experiments with standard data sets, as well as RGB-D images acquired by a Microsoft Kinect camera. For the second part on automatic trimap generation from depth data, problems due to inaccuracy and incompleteness of a depth image as obtained from a Microsoft Kinect camera has been addressed. A method has been developed to improve the quality of a Kinect's depth images, with the focus on its application to trimap generation. Improved trimap generation generally results in improved matting process. The proposed method was evaluated by using Kinect's depth images.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

โครงการวิจัยนี้ได้รับทุนสนับสนุนเงินวิจัยจากเงินรายได้ ประจำปีงบประมาณ 2557 คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
Abstract	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญภาพ	V
สารบัญตาราง	VI
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์	2
1.3 ขอบเขตงานวิจัย	2
1.4 ระเบียบวิจัย	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ	2
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง	3
2.1 ทฤษฎีการซ้อนภาพ (matting)	3
2.2 กล้องถ่ายภาพแบบ RGB-D	10
2.3 การปรับค่า (calibration) กล้อง Microsoft Kinect	11
บทที่ 3 การแยกพื้นหลังธรรมชาติ โดยอาศัยข้อมูลระยะลึก	12
3.1 การใช้ประโยชน์จากข้อมูลระยะลึก	12
3.2 เทคนิคการคำนวณภาพแมทโดยใช้ข้อมูลระยะลึก	13
3.3 การสร้างแผนภาพไตรแมบโดยอัตโนมัติจากข้อมูลระยะลึก	20
บทที่ 4 การทดลอง	26
4.1 การทดสอบเทคนิคในกลุ่ม Bayesian-based Matting Methods	27
4.2 การทดสอบเทคนิคในกลุ่ม Closed-form Matting Methods	31
4.3 การทดลองปรับปรุงคุณภาพของภาพระยะลึกเพื่อการสร้างแผนภาพไตรแมบอัตโนมัติ	42
บทที่ 5 สรุปผลการดำเนินงานและข้อเสนอแนะ	46
5.1 สรุปผลการดำเนินงาน	46
5.2 ข้อเสนอแนะ	47
บรรณานุกรม	48

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

ภาพที่	หน้า
2.1 ตัวอย่างการใช้การซ้อนภาพแบบบลูสกรีน เพื่อการซ้อนภาพฉากหลัง ในรายการโทรทัศน์ และอุตสาหกรรมภาพยนตร์	3
2.2 ภาพแสดงขั้นตอนการซ้อนภาพพื้นหลังโดยทั่วไป	5
2.3 การใช้ภาพแมทในการตัดส่วนของภาพที่ไม่ต้องการออก	6
2.4 แสดงการจัดเก็บค่าสีในแต่ละพิกเซลของภาพดิจิทัล	6
2.5 ตัวอย่างขั้นตอนโดยทั่วไปในการซ้อนภาพแบบดิจิทัล	7
2.6. ตัวอย่างภาพไตรแมบที่แสดงขอบเขตทั้งสามส่วนของภาพ	9
2.7 องค์ประกอบภายในของ Kinect	10
2.8 ตัวอย่างภาพ checkerboard	11
3.1 แสดงกราฟของ $H(\alpha)$ เมื่อเทียบกับค่า α ที่แตกต่างกัน	16
3.2 ตัวอย่างภาพระยะลึกจากอุปกรณ์ Kinect	21
3.3 ตัวอย่างภาพระยะลึกที่เติมค่าจุดที่ขาดหายโดยการประมาณค่าระยะลึกจากบริเวณใกล้เคียง	21
3.4 ภาพที่ได้จากการทำแบ่งกลุ่มพิกเซลด้วย K-mean	22
3.5 ตัวอย่างแผนภาพไตรแมบที่สร้างจากข้อมูลระยะลึกแบบอัตโนมัติ	22
3.6 ภาพตัวอย่างบริเวณที่เกิดจากการบังของวัตถุเบื้องหน้าของอุปกรณ์ Kinect	23
3.7 ภาพตัวอย่างบริเวณที่มีโอกาสเกิดจากการบังของวัตถุ	24
3.8 ภาพขั้นตอนการตรวจสอบบริเวณที่เกิดจากการบังของวัตถุหรือเป็นเงาของวัตถุ	25
4.1 ภาพต้นฉบับ ภาพไตรแมบ และภาพแมทที่สร้างขึ้นด้วยวิธี OBAY และค่า RMSE ที่ได้	27
4.2 อุปกรณ์ Kinect+DSLR ที่ใช้ในการทดลอง	28
4.3 ภาพสีจากกล้อง Kinect ภาพไตรแมบ ภาพแมทและภาพซ้อนที่สร้างขึ้นด้วยวิธีการ OBAY	28
4.4 ภาพต้นฉบับ ภาพไตรแมบ และภาพแมทที่สร้างขึ้นด้วยวิธี DBAY-W และค่า RMSE, SAD ที่ได้	29
4.5 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมบ ภาพแมทและภาพซ้อนด้วยวิธีการ DBAY-W	29
4.6 ภาพต้นฉบับ ภาพไตรแมบ และภาพแมทที่สร้างขึ้นด้วยวิธี DBAY-L และค่า RMSE, SAD ที่ได้	30
4.7 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมบ ภาพแมทและภาพซ้อนด้วยวิธีการ DBAY-L	31
4.8 ภาพต้นฉบับ ภาพไตรแมบ และภาพแมทที่สร้างขึ้นด้วยวิธี OCF และค่า RMSE, SAD ที่ได้	32
4.9 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมบ ภาพแมทและภาพซ้อนด้วยวิธีการ OCF	32
4.10 ภาพต้นฉบับ ภาพไตรแมบ และภาพแมทที่สร้างขึ้นด้วยวิธี DFC-P และค่า RMSE, SAD ที่ได้	34
4.11 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมบ ภาพแมทและภาพซ้อนด้วยวิธีการ DCF-P	35

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.12 ภาพต้นฉบับ ภาพไตรแมป และภาพแมทที่สร้างขึ้นด้วยวิธี DFC-S1 และค่า RMSE, SAD ที่ได้ 37
4.13 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมป ภาพแมทและภาพซ้อนด้วยวิธีการ DCF-S1 38
4.14 ภาพต้นฉบับ ภาพไตรแมป และภาพแมทที่สร้างขึ้นด้วยวิธี DFC-S2 และค่า RMSE, SAD ที่ได้ 39
4.15 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมป ภาพแมทและภาพซ้อนด้วยวิธีการ DCF-S2 40
4.16 ภาพเปรียบเทียบภาพแมท ที่คำนวณได้ของ OBAY, DBAY-W และ DBAY-L 41
4.17 เปรียบเทียบ ภาพสี ภาพระยะลึกที่ได้จากกล้อง และภาพระยะลึกที่ปรับปรุงแล้ว 43
4.18 ภาพสี ภาพไตรแมบที่สร้างจากข้อมูลภาพระยะลึก และภาพระยะลึกเมื่อซ้อนทับกับภาพสี 43
4.19 ตัวอย่างบริเวณ unknown region ในภาพไตรแมบที่กำหนดด้วยมือ ซึ่งไม่ครอบคลุมโดยภาพไตรแมบที่สร้างจากภาพระยะลึก 44
4.20 แสดงผลการสร้างภาพไตรแมบจากภาพระยะลึก เปรียบเทียบกับภาพที่กำหนดโดยใช้การวาดด้วยมือ 45

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
4.1 ค่าเฉลี่ยของ RMSE ที่ค่าพารามิเตอร์ต่าง ๆ โดยใช้วิธี DCF-P	33
4.2 ค่าเฉลี่ยของ RMSE ที่ค่าพารามิเตอร์ต่าง ๆ โดยใช้วิธี DCF-S1	36
4.3 ค่าเฉลี่ยของ RMSE ที่ค่าพารามิเตอร์ต่าง ๆ โดยใช้วิธี DCF-S2	38

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1. ความเป็นมาและความสำคัญของปัญหา

เทคนิคการซ้อนภาพ (matting techniques) เพื่อการแยกฉากหรือพื้นหลังออกจากวัตถุที่อยู่ด้านหน้า และซ้อนภาพฉากหลังใหม่เข้าไปแทนที่ เป็นกระบวนการที่มีความสำคัญต่อการปรับแต่งภาพถ่ายวีดีโอในงานด้านต่าง ๆ เช่น ในการถ่ายทำภาพยนตร์ โทรทัศน์ การตัดต่อวีดีโอเพื่อนำเสนอในลักษณะของสื่อประสม ตลอดจนการปรับแต่งรูปภาพนิ่ง เทคนิคดังกล่าว เมื่อถ่ายทำโดยใช้ฉากหลังเป็นพื้นสีน้ำเงิน (หรือสีเขียว) ในบางครั้งจะเรียกว่า blue-screen matting (หรือ green-screen matting) การใช้ฉากหลังเป็นสีพื้นดังกล่าวช่วยทำให้กระแยกพื้นหลังเดิมกระทำได้ง่ายกว่ากรณีที่ไม่สามารถควบคุมพื้นหลังได้ ในกรณีหลังจะเรียกรวม ๆ กันว่า เป็น พื้นหลังธรรมชาติ ด้วยเหตุที่ไมทราบสีของพื้นหลัง ทำให้โดยทั่วไปการแยกภาพพื้นหลังธรรมชาติในกระบวนการซ้อนภาพ จำเป็นต้องอาศัยข้อมูลขอบเขตของวัตถุซึ่งเรียกว่า ไตรแมบ (trimap) อันเป็นข้อมูลที่บ่งบอก (อย่างหยาบ ๆ) ว่าบริเวณใดในภาพเป็นส่วนของวัตถุด้านหน้า บริเวณใดเป็นส่วนของพื้นหลัง และบริเวณใดเป็นส่วนที่อยู่รอบ ๆ ขอบระหว่างวัตถุด้านหน้าและพื้นหลัง

โดยทั่วไปเทคนิคการซ้อนภาพดังกล่าว (นอกเหนือจากข้อมูลแผนภาพไตรแมบที่จำเป็นต้องใช้สำหรับฉากหลังธรรมชาติแล้ว) จะอาศัยเพียงข้อมูลภาพสีจากกล้องถ่ายภาพหรือกล้องวีดีโอ อย่างไรก็ดี ในปัจจุบันได้มีการพัฒนาและผลิตกล้องถ่ายภาพประเภท RGB-D ออกจำหน่ายอย่างแพร่หลาย ในหลายระดับราคา โดยกล้องดังกล่าว นอกเหนือจากข้อมูลภาพสี (RGB image) แล้วยังสามารถบันทึกข้อมูลระยะลึกของจุดต่าง ๆ ที่ปรากฏในภาพสีได้ด้วย ตัวอย่างของกล้องประเภทดังกล่าว เช่น กล้อง Kinect ของบริษัทไมโครซอฟต์ ด้วยเหตุที่กล้องประเภทดังกล่าว สามารถให้ข้อมูลเพิ่มเติมในรูปแบบของภาพระยะลึก ทำให้เกิดแนวคิดในการนำข้อมูลดังกล่าวมาใช้ให้เป็นประโยชน์ในการแยกวัตถุด้านหน้าออกจากพื้นหลัง ในกระบวนการซ้อนภาพ ทั้งนี้ เนื่องจากข้อมูลระยะลึกเป็นข้อมูลที่มีความสัมพันธ์โดยตรงต่อตำแหน่งของวัตถุในภาพที่สามารถใช้ช่วยจำแนกวัตถุด้านหน้าออกจากพื้นหลังได้ โครงการวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาการนำข้อมูลภาพระยะลึกมาใช้ในกระบวนการซ้อนภาพนอกเหนือจากการใช้ข้อมูลภาพสีตามปกติ โดยทำการศึกษาแนวทางการใช้ประโยชน์จากข้อมูลระยะลึกใน 2 ลักษณะ กล่าวคือ การใช้ภาพระยะลึกเพื่อการปรับปรุงคุณภาพของการคำนวณภาพแมบด้วยเทคนิคการซ้อนภาพที่มีอยู่ในปัจจุบัน และการใช้ภาพระยะลึกในการสร้างแผนภาพไตรแมบ โดยอัตโนมัติ เพื่อลดเวลาในการสร้างแผนภาพไตรแมบที่โดยปกติแล้วจำเป็นต้องให้ผู้ใช้เป็นผู้กำหนด (เช่นโดยการวาดเส้นแบ่งลงไปในภาพ) เป้าหมายของการศึกษา คือการทดลองประเมินประสิทธิภาพของวิธีการต่าง ๆ ที่มีอยู่ ตลอดจนการปรับปรุงกระบวนการสร้างแผนภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไตรแมบจากข้อมูลระยะลึกที่ถ่ายจากกล้อง RGB-D สำหรับการใช้งานตามบ้าน (consumer-graded camera) ซึ่งมีข้อจำกัดหลายประการทั้งในแง่ของความถูกต้องและความไม่สมบูรณ์ของภาพระยะลึกที่ได้

1.2 วัตถุประสงค์

1. เพื่อศึกษาเปรียบเทียบวิธีการต่าง ๆ ในการซ้อนภาพจากฉากหลังธรรมชาติ ที่ใช้ประโยชน์จากภาพระยะลึก

2. เพื่อศึกษาปัญหาของการใช้ข้อมูลภาพระยะลึกที่ได้จากกล้อง Microsoft Kinect เพื่อการสร้างแผนภาพไตรแมบ ตลอดจนการพัฒนาแนวทางการปรับปรุงภาพระยะลึกเพื่อสร้างแผนภาพไตรแมบแบบอัตโนมัติ

1.3 ขอบเขตของงานวิจัย

1. ศึกษาวิธีการซ้อนภาพโดยใช้ข้อมูลภาพระยะลึก กับเทคนิคการซ้อนภาพ สำหรับฉากหลังธรรมชาติ โดยพิจารณาการประยุกต์ใช้ข้อมูลระยะลึกกับเทคนิคที่มีการอ้างอิงและศึกษาอย่างแพร่หลาย 2 วิธี อันได้แก่ Bayesian-based matting technique และ Closed-form matting technique.

2. ศึกษาแนวทางการสร้างแผนภาพไตรแมบ จากข้อมูลภาพระยะลึก โดยพิจารณาปัญหาและแนวทางการปรับปรุงภาพระยะลึกที่ได้จากกล้อง Microsoft Kinect

3. ทำการทดสอบและทดลองโดยใช้ข้อมูลภาพจากฐานข้อมูลมาตรฐานที่เปิดให้ใช้งานโดยทั่วไป และข้อมูลภาพระยะลึกที่บันทึกโดยใช้กล้อง Microsoft Kinect โดยในส่วนของภาพสี ใช้กล้อง DSLR ในการบันทึกแทนเพื่อให้ได้คุณภาพที่ดีกว่ากล้องถ่ายภาพสีของอุปกรณ์ Kinect

1.4. ระเบียบวิจัย

1. ศึกษางานวิจัยที่เกี่ยวข้อง
2. จัดทำฐานข้อมูลภาพเพื่อการทดสอบ
3. เนินการทดสอบด้วยภาพจำลอง
4. ดำเนินการทดสอบจากภาพถ่ายจริง
5. สรุปผลการทดสอบเชิงเปรียบเทียบ

1.5 ประโยชน์ที่คาดว่าจะได้รับของโครงการวิจัย

1. ทราบขีดความสามารถของวิธีการที่มีอยู่ในปัจจุบัน
2. เป็นแนวทางในการปรับปรุงเทคนิคที่มีในปัจจุบันให้ดีขึ้น

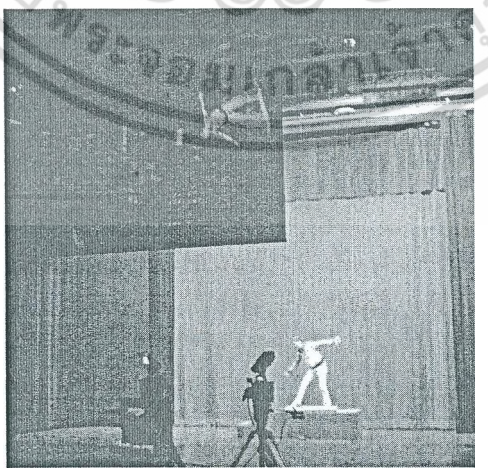
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีที่เกี่ยวข้อง

2.1 ทฤษฎี การซ้อนภาพ (matting)

เทคนิคการซ้อนภาพ (matting) เป็นวิธีการที่ใช้กันอย่างแพร่หลายในการถ่ายทำภาพยนตร์สื่อวีดีโอหรือโทรทัศน์ บางครั้งมีความจำเป็นต้องปรับเปลี่ยนพื้นหลัง (Background) ที่ปรากฏในภาพที่ถ่ายทำในครั้งแรกด้วยพื้นหรือฉากหลังใหม่โดยไม่ให้กระทบกับภาพของผู้แสดงหรือวัตถุที่อยู่ด้านหน้า (Foreground) ทั้งนี้อาจมีสาเหตุมาจากความจำเป็นในการลดค่าใช้จ่ายในการผลิต คิวนัดหมายนักแสดงในการถ่ายทำที่จำกัดหรือเพื่อประกอบการสร้างเทคนิคพิเศษด้านภาพ (visual/special effects) ที่ไม่สามารถหาสร้างหรือจำลองฉากหลังจริงขึ้นมาได้ เป็นต้น ซึ่งเทคนิคในการซ้อนภาพฉากหลังดังกล่าว เป็นที่รู้จักและใช้งานกันอย่างแพร่หลายมาหลายสิบปี แต่อาจเป็นที่รู้จักกันในชื่อที่แตกต่างกัน เช่น เทคนิคแบบบลูสกรีน (Blue screen matting) หรือเทคนิคโคร-มาคีย์ (chroma key)

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



ภาพที่ 2.1 ตัวอย่างการใช้การซ้อนภาพแบบบลูสกรีน เพื่อการซ้อนภาพฉากหลัง ในรายการโทรทัศน์ และอุตสาหกรรมภาพยนตร์ (ภาพโดย Mark Barker จาก http://commons.wikimedia.org)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

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

2.1.1 หลักการทั่วไปของการซ้อนภาพพื้นหลัง

การซ้อนภาพพื้นหลังมีหลักการและขั้นตอนโดยทั่วไป แสดงดังในภาพที่ 2.2 จากภาพสามารถอธิบายขั้นตอนต่าง ๆ ได้โดยสังเขป ดังนี้

1. ทำการถ่ายภาพนักแสดงหรือวัตถุที่จะปรากฏอยู่ด้านหน้ากล้องในภาพ โดยในขั้นตอนนี้ สำหรับเทคนิคการซ้อนภาพโดยทั่วไป จำเป็นต้องจัดฉาก (screen) ที่เป็นพื้นหลังสีเดียว เช่น สีน้ำเงิน หรือสีเขียว ในขณะที่เทคนิคการซ้อนภาพแบบดิจิทัลสมัยใหม่ อาจเป็นพื้นหลังใด ๆ เช่น พื้นหลังที่ปรากฏตามธรรมชาติ ณ สถานที่ที่ถ่ายทำนั้น ๆ ในภาพตัวอย่าง แสดงเฉพาะกรณีที่พื้นหลังเป็นฉากสีน้ำเงิน เพื่อง่ายต่อการนำเสนอ

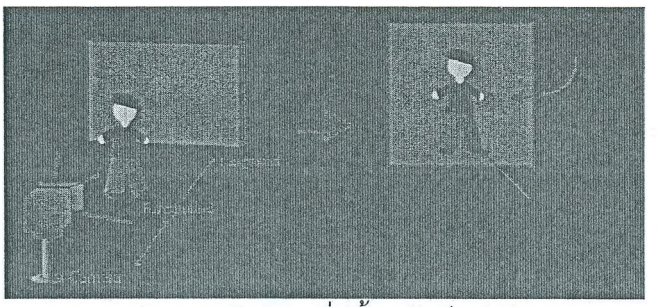
2. ทำการจัดเตรียมภาพฉากหลังใหม่ ซึ่งอาจจัดทำโดยการวาดหรือสร้างด้วยคอมพิวเตอร์ หรืออาจจะเป็นการถ่ายภาพวัตถุหรือสถานที่จริงก็ได้

3. ทำการตัดส่วนที่เป็นฉากหลังออกไปจากภาพที่ถ่ายได้จากขั้นตอนที่ 1 ซึ่งจะมีผลทำให้บริเวณดังกล่าวเป็นสีดำสนิท

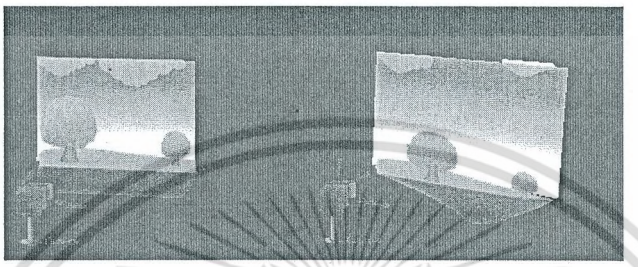
4. จากภาพที่ได้ในขั้นตอนที่ 2 ทำกระบวนการเดียวกับในขั้นตอนที่ 3 หากแต่ในขั้นตอนนี้ ตัดภาพในบริเวณที่ตรงกับตำแหน่งของนักแสดงในภาพที่ถ่ายจากขั้นตอนที่ 1 แทน

5 นำภาพที่ได้จากขั้นตอนที่ 3 และ 4 มาผสมเข้าด้วยกัน โดยส่วนที่เป็นสีดำในแต่ละภาพ จะไม่มีผลต่อสีหรือความสว่างของภาพบริเวณนั้นหลังการผสมเข้าด้วยกัน ทำให้ในภาพสุดท้ายบริเวณที่เป็นนักแสดงหรือวัตถุที่อยู่ด้านหน้าจะยังคงเหมือนภาพต้นฉบับที่ถ่ายได้จากในขั้นตอนที่ 1 จะมีการเปลี่ยนแปลงเฉพาะบริเวณฉากหลังเดิมที่ถูกแทนที่ด้วยฉากหลังใหม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



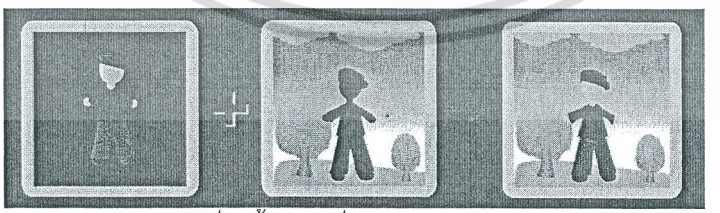
1. ถ่ายภาพนักแสดงที่มีพื้นหลังเป็นสีเดียว



2. ถ่ายหรือสร้างภาพฉากหลังใหม่ที่ต้องการ

3. ตัดฉากหลังออกจากภาพต้นฉบับ

4. ตัดบริเวณที่ซ้อนกับตำแหน่งนักแสดงออกจากภาพฉากหลังใหม่



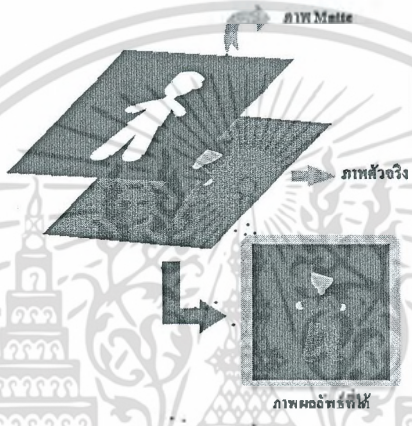
5. นำภาพที่ได้ขั้นตอนที่จาก 3 และ 4 มารวมกัน

ภาพที่ 2.2 แสดงขั้นตอนการซ้อนภาพพื้นหลังโดยทั่วไป

ในทางปฏิบัติ การตัดภาพส่วนที่ไม่ต้องการออกจากภาพต้นฉบับในขั้นตอนที่ 3 (หรือขั้นตอนที่ 4 ที่กระทำกับภาพฉากหลังใหม่) จำเป็นต้องทำการสร้างภาพแมท (matte) ซึ่งเป็นภาพที่มีขนาดเท่ากับภาพใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

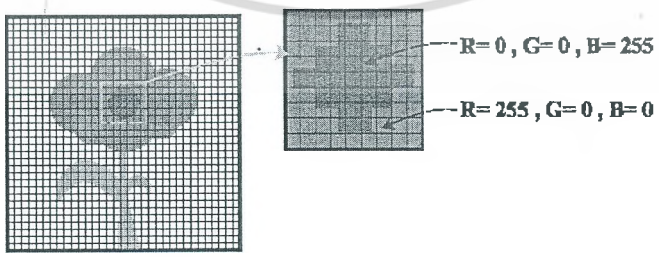
ขั้นตอนที่ 1 โดยภาพแมทส่วนที่ตรงกับบริเวณของภาพที่ไม่ต้องการให้ถูกตัดออกไป จะมีลักษณะโปร่งแสงเหมือนแผ่นใส ในขณะที่บริเวณที่ต้องการตัดออก จะมีลักษณะทึบแสง เมื่อนำภาพแมทมาวางไว้เหนือภาพต้นฉบับที่เตรียมไว้ (จากขั้นตอนที่ 1 หรือ 2) ก็จะได้ภาพสุดท้ายตามต้องการ (ดูภาพที่ 2.3 ประกอบ) กระบวนการในรายละเอียด สำหรับแต่ละขั้นตอน อาจแตกต่างกันออกไป ขึ้นอยู่กับเทคโนโลยีที่ใช้ เช่น ในกรณีของการใช้ฟิล์มภาพยนตร์ ภาพแมทจะถูกสร้างโดยกระบวนการถ่ายภาพลงบนแผ่นฟิล์มที่มีลักษณะเฉพาะ และในการรวมภาพเข้าด้วยกัน จะกระทำโดยการถ่ายภาพ [2-3] ที่ได้จากการนำฟิล์มของภาพแมทและภาพต้นฉบับมาซ้อนทับกัน



ภาพที่ 2.3 การใช้ภาพแมทในการตัดส่วนของภาพที่ไม่ต้องการออก

2.1.2 **เทคนิคการซ้อนภาพฉากหลังแบบดิจิทัล**

ภาพดิจิทัล ไม่ว่าจะเป็นภาพนิ่ง หรือภาพเคลื่อนไหว มีหลักการพื้นฐานเหมือนกัน กล่าวคือ ข้อมูลภาพจะถูกตัดแบ่งออกเป็นบริเวณหรือจุดย่อยๆเป็นจำนวนมาก แต่ละจุดเรียกว่าพิกเซล (pixel) โดยการจัดเก็บภาพในรูปแบบดิจิทัล จะทำการบันทึกค่าสีหรือความสว่างในแต่ละพิกเซลไว้ (ดูภาพที่ 2.4 ประกอบ)

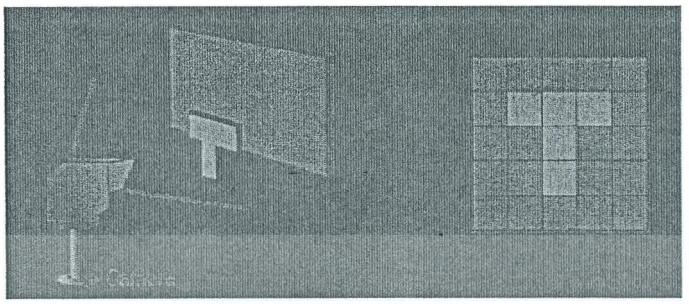


ภาพที่ 2.4 แสดงการจัดเก็บค่าสีในแต่ละพิกเซลของภาพดิจิทัล

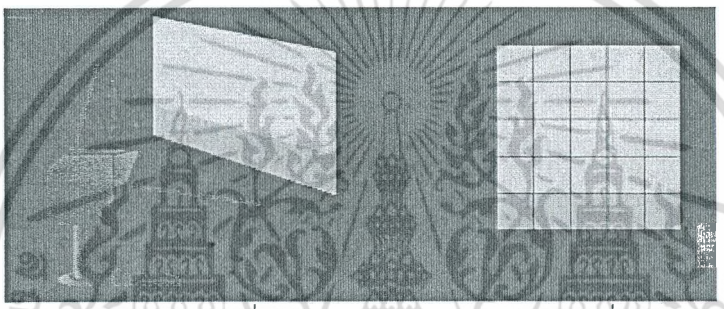
สำหรับภาพสี แต่ละพิกเซลจะประกอบไปด้วยค่าความสว่างของแม่สีทั้ง 3 อันได้แก่ สีแดง สีเขียว และสีน้ำเงิน (Red Green Blue หรือเรียกย่อๆว่า RGB) โดยแต่ละสีจะมีค่าความสว่างระหว่าง 0 ถึง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

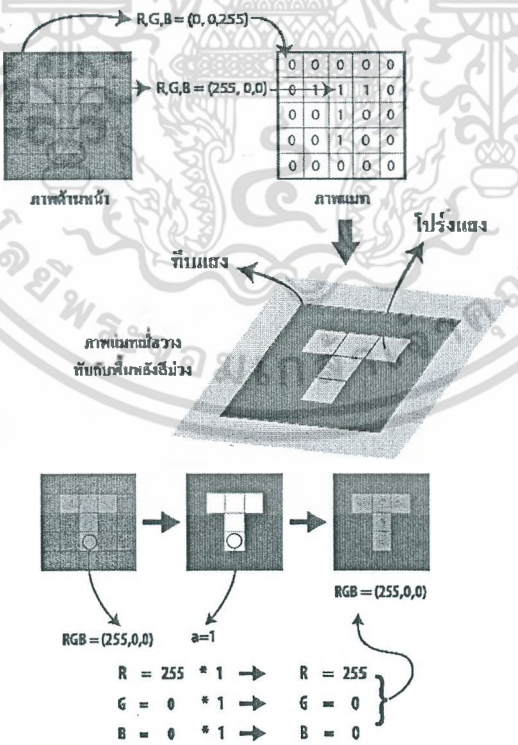
255 เป็นต้น เมื่อพิจารณากระบวนการซ้อนภาพดังที่ได้อธิบายไว้ในหัวข้อ 2.2 และ ภาพที่ 2 ในกรณีของภาพดิจิทัลจะ มีรายละเอียดดังแสดงในภาพที่ 2.5



1. ภาพดิจิทัล (ขวา) ที่ได้จากการถ่ายภาพวัตถุ โดยใช้พื้นหลังสีเดียว



2. ภาพดิจิทัล (ขวา) ที่ได้จากการถ่ายภาพฉากหลังใหม่ที่ต้องการ



3. การสร้างภาพแมท และกระบวนการในการรวมภาพเข้าด้วยกัน

ภาพที่ 2.5 ตัวอย่างขั้นตอนโดยทั่วไปในการซ้อนภาพแบบดิจิทัล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูป หลังผ่านขั้นตอนที่ 1 และ 2 จะได้ภาพดิจิทัลของนักแสดงหรือวัตถุที่อยู่ด้านหน้าและภาพฉากหลังใหม่ ตามลำดับ เมื่อเข้าสู่ขั้นตอนที่ 3 จะเริ่มจากการสร้างภาพแมทโดยภาพแมทเป็นภาพดิจิทัลที่มีขนาดเท่ากับภาพต้นฉบับ แต่มีข้อแตกต่างกันตรงที่ในแต่ละพิกเซลของภาพแมท จะทำการจัดเก็บค่าระดับความโปร่งแสงเป็นตัวเลขตัวเดียวที่มีค่าอยู่ระหว่าง 0 ถึง 1 แทนการเก็บค่าสี RGB จำนวน 3 ค่าเหมือนในกรณีของภาพต้นฉบับ โดยการสร้างภาพแมท จะขึ้นอยู่กับลักษณะของฉากหลังเดิม ในหัวข้อนี้จะขอยกตัวอย่างกรณีที่ฉากหลังเป็นพื้นหลังสีเดียว ที่ทราบค่าสีล่วงหน้า ในกรณีนี้ โปรแกรมที่ใช้ในการประมวลผลภาพเพื่อการซ้อนภาพฉากหลัง จะทำการอ่านค่าสีในแต่ละพิกเซลของภาพที่ได้จากขั้นตอนที่ 1 หาก ณ ตำแหน่งพิกเซลใดมีค่าสีตรงหรือใกล้เคียงกับค่าสีของฉากหลัง (ที่ทราบค่าอยู่ก่อนแล้ว) ก็จะทำการกำหนดค่าสีของภาพแมท ณ ตำแหน่งพิกเซลที่ตรงกันให้เป็นสีดำหรือทึบแสง (ค่าของพิกเซลเท่ากับ 0) หากสีไม่ตรงหรือไม่ใกล้เคียงกับสีฉากหลัง ก็จะกำหนดให้พิกเซลดังกล่าวของภาพแมทเป็นจุดโปร่งแสงแทน (ค่าของพิกเซลเท่ากับ 1)หลังจากได้ภาพแมทแล้ว การตัดภาพในส่วนที่ไม่ต้องการ จะกระทำโดยการคูณค่าความโปร่งแสงที่เก็บไว้ในแต่ละพิกเซลของภาพแมทเข้ากับค่าสีในพิกเซลที่ตำแหน่งตรงกันกับภาพต้น เป็นผลทำให้ภาพที่ได้ที่อยู่ในส่วนที่ทึบแสงของภาพแมท จะปรากฏเป็นสีดำเสมอ ในขณะที่ส่วนที่ตรงกับบริเวณโปร่งแสง จะไม่มีการเปลี่ยนแปลงของค่าสีจากภาพต้นฉบับแต่อย่างใด อนึ่งภาพแมทดังกล่าวซึ่งบ่งบอกระดับความโปร่งแสงของแต่ละพิกเซล บางครั้งจะทำการบันทึกรวมไปกับข้อมูลค่าสี RGB ของรูปภาพ โดยเรียกว่าค่าแอลฟา (alpha value)

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

หลังจากได้ภาพที่ผ่านกระบวนการตัดส่วนภาพออกแล้ว การผสมภาพ ทำได้โดยการนำค่าสีของแต่ละพิกเซลที่ตรงกันของภาพต้นฉบับและภาพฉากหลังใหม่ มาบวกเข้าด้วยกัน

2.1.3 วิธีการแยกภาพฉากหลังธรรมชาติ

ในกระบวนการซ้อนภาพ ขั้นตอนที่สำคัญ คือการแยกภาพวัตถุด้านหน้าออกจากภาพฉากหลัง กระบวนการดังกล่าวมีความยุ่งยากมากขึ้น เมื่อฉากหลังเป็นวัตถุใด ๆ ที่ไม่สามารถควบคุมสีและความสม่ำเสมอของสีหรือความสว่างได้ ตลอดจนภาพที่พิกเซลบริเวณขอบวัตถุด้านหน้ากับฉากหลังมีการผสมผสานกัน ทำให้กระบวนการแยกฉากหลังออกมีความซับซ้อนมากขึ้น

เพื่อให้สามารถทำการซ้อนภาพในกรณีดังกล่าวข้างต้น จำเป็นต้องกำหนดให้พิกเซลบริเวณรอยต่อระหว่างวัตถุด้านหน้ากับพื้นหลัง เป็นการผสมกันระหว่างสีของวัตถุด้านหน้ากับสีพื้นหลัง ตามสมการต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

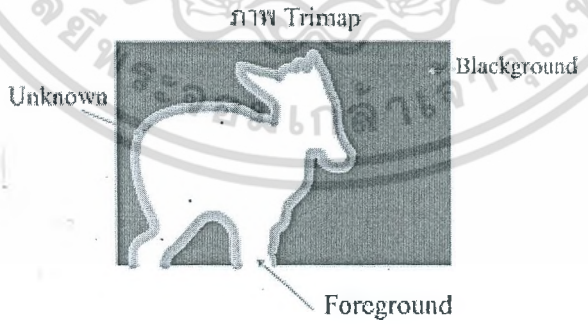
$$I_k = \alpha_k F_k + (1 - \alpha_k) B_k \quad (2.1)$$

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

จากสมการที่ 2.1 หากสามารถคำนวณหรือหาค่า α_k และ F_k จากภาพที่ถ่าย (I_k) ได้ ก็จะสามารถสร้างภาพที่ซ้อนฉากหลังใหม่ได้ด้วยสมการต่อไปนี้

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

โดยที่จากสมการข้างต้น B'_k คือค่าสี ณ พิกเซลที่ k ของภาพฉากหลังใหม่ และ J_k คือค่าสี ณ พิกเซลที่ k ของภาพใหม่ที่ได้ ปัญหาคือการคำนวณหาค่า α_k และ F_k จาก I_k ด้วยสมการที่ 2.1 โดยทั่วไป ไม่สามารถหาได้ เนื่องจากมีตัวแปรที่ไม่ทราบค่าถึง 3 ตัว อย่างไรก็ดี ในทางปฏิบัติ จะอาศัยการเลือกค่าสีของวัตถุที่อยู่ด้านหน้า และค่าสีของวัตถุในฉากหลัง ตรงบริเวณที่ใกล้เคียงกับขอบวัตถุที่ทราบค่าอยู่ก่อนแล้ว ในการประมาณค่า F_k และ I_k เพื่อให้สามารถแก้สมการหาค่าของ α_k ได้ ด้วยวิธีการดังกล่าว สิ่งสำคัญที่จะต้องกำหนด (โดยผู้ใช้) จากภาพต้นฉบับ คือบริเวณในภาพที่เป็นวัตถุที่อยู่ด้านหน้า และบริเวณในภาพที่เป็นวัตถุถัดไปที่อยู่ในฉากหลัง ข้อมูลดังกล่าว ซึ่งโดยทั่วไปจะกำหนดให้อยู่ในรูปของแผนภาพไตรแมบ (ดูภาพที่ 2.6 ประกอบ) จะถูกนำไปใช้ประกอบกับภาพสีต้นฉบับในการคำนวณหาภาพแมท เพื่อใช้แยกฉากหลังธรรมชาติ ซึ่งโดยทั่วไปมีขั้นตอนดังต่อไปนี้



ภาพที่ 2.6. ตัวอย่างภาพไตรแมบ ที่แสดงขอบเขตทั้งสามส่วนของภาพ

1. สำหรับแต่ละจุดพิกเซล k ใด ๆ ที่อยู่ใน unknown region ทำการเลือกจุดพิกเซลที่ใกล้เคียงจาก background region และใช้ค่าสีของพิกเซลดังกล่าวแทนค่า B_k ที่แท้จริง และในทำนองเดียวกัน ทำการเลือกจุดพิกเซลที่ใกล้เคียงจาก foreground region และใช้ค่าสีของพิกเซล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังกล่าวแทนค่า F_k ที่แท้จริง เช่น อาจเลือกใช้จุดใน foreground/background region ที่อยู่ใกล้กับจุดที่ต้องการคำนวณมากที่สุด เป็นต้น

2. ทำการคำนวณหาค่าความโปร่งแสง (α_k) จากสมการที่ (2.1)

3. สำหรับตำแหน่งของพิกเซลที่อยู่ใน background region กำหนดค่า α_k เป็น 0 และตำแหน่งที่อยู่ใน foreground region กำหนดค่า α_k เป็น 1

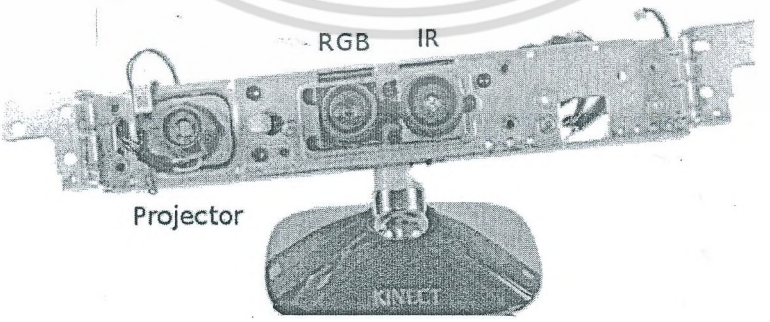
4. ทำการซ้อนภาพฉากหลังใหม่ โดยใช้สมการที่ (2.1) โดยอาศัยค่า α_k ที่คำนวณได้จากขั้นตอนที่แล้ว และค่าประมาณการของ F_k จากในขั้นตอนที่ 2

5. คุณภาพของการซ้อนภาพอาจปรับปรุงให้ดีขึ้น ด้วยการคำนวณค่าของ B_k และ F_k ให้ใกล้เคียงกับค่าสีที่แท้จริงมากขึ้น ส่งผลให้รายละเอียดภาพมีความถูกต้องมากขึ้น

2.2 กล้องถ่ายภาพแบบ RGB-D

กล้องถ่ายภาพแบบ RGB-D คือกล้องที่มีคุณสมบัติในการวัดและบันทึกระยะความลึกของภาพจากการถ่ายภาพจริง ซึ่งปัจจุบันสามารถแบ่งออกได้เป็น 3 ประเภท ตามลักษณะการคำนวณหาค่าระยะความลึกได้แก่ 1) แบบที่ใช้ time-of-flight sensor (TOF sensor) ใช้หลักการคำนวณจากระยะเวลาการเดินทางของแสง 2) . แบบที่ใช้ stereo camera ใช้หลักการภาพหลายมุมมองในการคำนวณหาระยะความลึก และ 3) แบบที่ใช้หลักการ structured light สำหรับรายงานวิจัยฉบับนี้จะใช้อุปกรณ์ Kinect ซึ่งทำงานด้วยหลักการแบบที่ 3 ในการหาค่าระยะลึก

Kinect คืออุปกรณ์ที่มีความสามารถในการถ่ายภาพระยะลึก ซึ่งภายในจะประกอบไปด้วย 3 ส่วนหลักคือ กล้องอินฟาเรด (Infrared-IR camera) , ส่วน IR โปรเจคเตอร์ (IR projector) และกล้องสี (RGB camera) ดังภาพที่ 2.7 สำหรับการทำงานของ Kinect เริ่มจากการฉายแสงอินฟาเรดออกจาก IR projector ซึ่งเป็นแสงที่ไม่สามารถมองเห็นด้วยตาเปล่า แสงที่ฉายจะมีลักษณะเป็นจุดๆตามแนวตั้ง 480 จุด และแนวนอน 640 จุด โดยในส่วนของกล้องอินฟราเรดจะรับแสงอินฟราเรดที่สะท้อนกลับมา เพื่อทำการคำนวณระยะความลึกของแสงอินฟาเรดที่ตกกระทบกับพื้นผิวหรือวัตถุ และกล้องสีจะเป็นตัวถ่ายภาพแบบปกติ



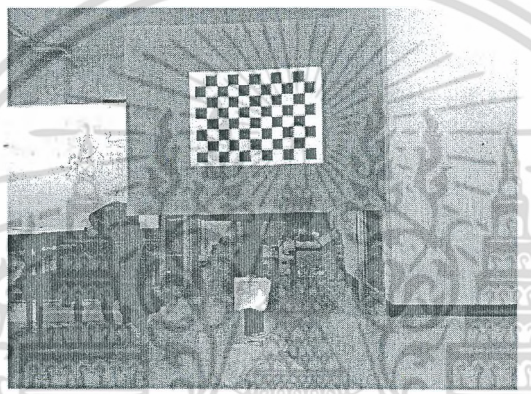
ภาพที่ 2.7 องค์ประกอบภายในของ Kinect

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 การปรับค่า (calibration) กล้อง Microsoft Kinect

กล้อง Kinect โดยทั่วไปจะมีค่า focal length และขนาด sensor ที่แตกต่าง ซึ่งเกิดจากสายการผลิตหรือความคลาดเคลื่อนในจัดทำ ทำให้ส่งผลต่อประสิทธิภาพการใช้งาน. จึงจำเป็นต้องทำกระบวนการปรับตั้งค่า (calibration) เพื่อทำการปรับค่าตัวอุปกรณ์ Kinect ให้มีความถูกต้องยิ่งขึ้น

การปรับตั้งค่าอุปกรณ์ Kinect กระทำได้โดยการใช้กล้อง Kinect ถ่ายภาพกระดานลายหมากรุก (checkerboard) จากมุมมองต่างๆ (จำนวนอย่างน้อย 30 ภาพ) เพื่อภาพที่ได้จากการถ่ายมาคำนวณหาค่าที่แท้จริงของตัวอุปกรณ์ Kinect เครื่องนั้นเอง ดังภาพที่ 2.8 นำภาพที่จากการถ่ายมาทำการปรับตั้งค่า ซึ่งในงานวิจัยนี้ใช้ซอฟต์แวร์ MATLAB Toolbox ที่เผยแพร่ใน [7] เพื่อการปรับค่าดังกล่าว



ภาพที่ 2.8 ตัวอย่างภาพ checkerboard

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การแยกพื้นหลังธรรมชาติ โดยอาศัยข้อมูลระยะลึก

3.1. การใช้ประโยชน์จากข้อมูลระยะลึก

ข้อมูลระยะลึกได้ถูกนำมาใช้ในกระบวนการซ้อนภาพในลักษณะที่แตกต่างกัน วิธีการที่แตกต่างกันดังกล่าว สามารถจำแนกลักษณะของการประยุกต์ใช้ตามวัตถุประสงค์ออกได้เป็น 2 ประเภท ดังต่อไปนี้

3.1.1 ใช้เพื่อเป็นข้อมูลเพิ่มความถูกต้องของการคำนวณภาพแมท

ปัญหาหาหลักที่ส่งผลต่อการแยกภาพวัตถุกับภาพพื้นหลังเพื่อนำมาซ้อนกับภาพพื้นหลังใหม่คือ การที่ภาพวัตถุและภาพพื้นหลังบริเวณดังกล่าวมีสีพื้นที่ใกล้เคียงกันทำให้ยากต่อการกำหนดบ่งบอกบริเวณดังกล่าวส่วนใดเป็นของภาพพื้นหลังหรือของวัตถุ ตัวอย่างเช่น ภาพสีพื้นหลังบริเวณดังกล่าวมีค่าสี R=200 , G = 180 , B = 190 และค่าสีของวัตถุคือ R= 195 ,G = 185 ,B =192 ซึ่งจะเห็นได้ว่าค่าของภาพพื้นหลังและภาพวัตถุบริเวณดังกล่าวมีความใกล้เคียงกันมาก ทำให้ส่งผลต่อความถูกต้องในการทำ matting และการคำนวณภาพแม็ทหรือค่า α จากปัญหาดังกล่าวข้อมูลระยะลึกถูกเพิ่มเข้ามาเพื่อช่วยเพิ่มความถูกต้องของการคำนวณภาพแมท โดยข้อมูลระยะลึกสามารถช่วยในการคัดกรองภาพพื้นหลังและวัตถุโดยอาศัยจากการที่วัตถุโดยทั่วไปจะมีระยะห่างจากภาพพื้นหลัง เช่นในตัวอย่างที่กล่าวมาในข้างต้นที่ ถ้าภาพพื้นหลังมีค่าระยะลึกคือ 3.5 เมตร และภาพวัตถุอยู่ที่ 1.2 เมตร จะเห็นได้ว่าแม้ว่าสีพื้นหลังกับวัตถุจะมีความใกล้เคียงกันแต่พื้นหลังกับภาพวัตถุนั้นมีค่าระยะลึกที่แตกต่างกันถึง 2.3 เมตร ทำให้สามารถแยกบริเวณดังกล่าวได้ ซึ่งจากกรณีที่กล่าวมาทำให้ข้อมูลระยะลึกจึงมีส่วนช่วยเพิ่มความถูกต้องของภาพแมท

3.1.2. ใช้ในการสร้างแผนภาพไตรแมบแบบอัตโนมัติ

แผนภาพไตรแมบเป็นแผนภาพที่มีความสำคัญต่อการคำนวณภาพแมท โดยประกอบไปด้วย 3 ส่วนคือส่วน foreground region, ส่วน background region และส่วน unknown region โดยแผนภาพไตรแมบ จะส่งผลต่อการประสิทธิภาพและความถูกต้องในการคำนวณภาพแมท สำหรับการสร้างแผนภาพไตรแมบ โดยทั่วไปที่มีพื้นหลังธรรมชาติจะต้องให้ผู้ใช้งานเป็นคนกำหนดกรอบขอบเขตของภาพทั้งสามส่วนดังกล่าว ซึ่งข้อมูลระยะลึกสามารถช่วยในการสร้างแผนภาพไตรแมบ ในแบบอัตโนมัติได้ เพราะอาศัยค่าระยะความลึกที่แตกต่างกันของวัตถุด้านหน้าและพื้นหลัง เป็นตัวกำหนด ทำให้ช่วยลดขั้นตอนและระยะเวลาในการกำหนดแผนภาพไตรแมบโดยผู้ใช้ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 เทคนิคการคำนวณภาพแมทโดยใช้ข้อมูลระยะลึก

ด้วยความแพร่หลายมากขึ้นของกล้องถ่ายภาพแบบ RGB-D ที่สามารถเก็บข้อมูลภาพระยะลึกได้ นอกเหนือจากภาพสี อีกทั้งความก้าวหน้าในด้านการคำนวณระยะลึกจากภาพหลายมุมมอง ทำให้สามารถคำนวณระยะลึกของจุดต่าง ๆ ที่ปรากฏในภาพ จากมุมมองที่แตกต่างกันของวัตถุได้ (เช่น โดยใช้เฟรมที่แตกต่างกันของภาพวีดีโอ หรือใช้กล้องถ่ายภาพแบบสามมิติ) ทำให้เกิดมีแนวคิดในการปรับปรุงคุณภาพของการคำนวณภาพแมท โดยอาศัยข้อมูลระยะลึกให้เป็นประโยชน์ เพิ่มเติมจากข้อมูลสีหรือความสว่างที่ได้จากกล้องถ่ายภาพปกติโดยทั่วไป ในหัวข้อนี้ จะได้กล่าวถึงงานวิจัยต่าง ๆ ที่ใช้ประโยชน์จากข้อมูลระยะลึกเพื่อจุดประสงค์ดังกล่าว ดังต่อไปนี้

3.2.1 การใช้ภาพระยะลึกกับเทคนิค Bayesian-based Matting

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

$$
\begin{aligned}
(F,B,\alpha) &= \arg\max_{F,B,\alpha} P(F,B,\alpha \mid C) \\
&= \arg\max_{F,B,\alpha} P(C|F,B,\alpha)P(F)P(B)P(\alpha)/P(C) \\
&= \arg\max_{F,B,\alpha} L(C|F,B,\alpha) + L(F) + L(B) + L(\alpha)
\end{aligned}
\tag{3.1}
$$

จากสมการข้างต้น C, F, และ B เป็นเวกเตอร์ขนาด 3×1 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 α ตามสมการต่อไปนี้ [2]

$$
\begin{bmatrix} \Sigma_F^{-1} + I\alpha^2/\sigma_c^2 & I\alpha(1-\alpha)/\sigma_c^2 \\ I\alpha(1-\alpha)/\sigma_c^2 & \Sigma_F^{-1} + I(1-\alpha)^2/\sigma_c^2 \end{bmatrix} \begin{bmatrix} F \\ B \end{bmatrix} = \begin{bmatrix} \Sigma_F^{-1}\overline{F} + C\alpha/\sigma_c^2 \\ \Sigma_B^{-1}\overline{B} + C(1-\alpha)/\sigma_c^2 \end{bmatrix} \tag{3.2}
$$

และ

$$
\alpha = \frac{(C-B)(F-B)}{\|F-B\|^2} \tag{3.3}
$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการแก้สมการแรกข้างต้น จะอาศัยค่าเริ่มต้นของ α ที่กำหนดโดยผู้ใช้หรืออาจเป็นค่าที่ได้จากการคำนวณในรอบที่แล้ว (previous iteration) มาใช้ โดยปกติการคำนวณกลับไปมาข้างต้นจะกระทำจนกระทั่งการเปลี่ยนแปลงของค่า $\max_{F,B,\alpha} P(F,B,\alpha \mid C)$ ที่ได้จากสมการ (3.1) น้อยกว่าระดับค่าที่ตั้งไว้

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

1. ข้อมูลระยะลึกกับข้อมูลค่าสีเป็นข้อมูลคนละประเภทที่มีสเกลที่แตกต่างกัน ทำให้ยากต่อการกำหนดความสำคัญหรือให้น้ำหนักที่เหมาะสมระหว่างข้อมูล 2 ประเภทดังกล่าว

2. ในความเป็นจริง ข้อมูลระยะลึกไม่ได้มีความสัมพันธ์โดยตรงกับค่าแมท (α) เนื่องจากค่า α มีความสัมพันธ์กันโดยตรงกับส่วนที่เป็นค่าสีในเวกเตอร์ C, F, และ B ไม่ใช่ค่าระยะลึก และในขณะที่ค่าสีใน C เกิดจากการผสมกันระหว่างค่าสีของ F และ B ในขณะที่ค่าระยะลึกที่จุดเดียวกันนั้น ไม่ได้มีความสัมพันธ์ในทางกายภาพกับค่าระยะลึกใน F และ B แต่อย่างใด (และหากมี ก็ไม่ได้มีความสัมพันธ์โดยตรงกับค่า α ที่ต้องการทราบ)

ด้วยเหตุผลดังกล่าวข้างต้น ทำให้จำเป็นต้องมีการกำหนดโมเดลที่เป็นสมมติฐานที่ใช้ในการระบุความสัมพันธ์ระหว่างค่าระยะลึก กับการคำนวณค่า α ซึ่งโมเดลหรือสมมติฐานที่แตกต่างกันนำไปสู่วิธีการที่แตกต่างกันด้วย ดังจะได้กล่าวถึงแต่ละวิธีการและสมมติฐานความสัมพันธ์อันเป็นที่มาของวิธีการต่าง ๆ ดังกล่าว ดังต่อไปนี้

ก) วิธีของ Wang [3]

โดยหลักการแล้ว หากวัตถุที่อยู่ด้านหน้าและพื้นหลังที่ต้องการแยกออกจากกัน อยู่ห่างกันมากพอตามแนวทิศทางการมองของกล้องที่ใช้บันทึกภาพ (แกน z) และข้อมูลระยะลึกที่ได้มีความถูกต้องเพียงพอ เราสามารถแยกวัตถุด้านหน้าออกจากพื้นหลังได้ โดยการทำการแบ่งกลุ่มของข้อมูลระยะลึกออกกลุ่ม ๆ โดยใช้เทคนิคการแบ่งกลุ่มข้อมูล (clustering techniques) เช่น k-mean เป็นต้น พิกเซลที่ค่าระยะลึกอยู่ในกลุ่มที่ใกล้กับกล้องมากกว่า จะเป็นพิกเซลที่ประกอบกันขึ้นเป็นภาพวัตถุด้านหน้า ส่วนพิกเซลอื่น ๆ ที่อยู่ในกลุ่มที่อยู่ห่างออกเป็น จะเป็นภาพที่เกิดจากวัตถุที่เป็นส่วนหนึ่งของพื้นหลัง วิธีการดังกล่าวซึ่งจัดเป็นลักษณะของการแบ่งแยกวัตถุในภาพแบบ hard segmentation (คือจุดใด ๆ ในภาพเป็นสมาชิกในกลุ่มวัตถุกลุ่มใดกลุ่มหนึ่งเท่านั้น) ไม่เหมาะกับการใช้งานในลักษณะของภาพที่จำเป็นต้องใช้เทคนิคการคำนวณภาพแมทแบบดิจิทัล ด้วยเหตุผลหลัก 2 ประการ คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. กรณีขอบหรือรอยต่อระหว่างวัตถุด้านหน้าและพื้นหลังที่ไม่ชัดเจน เช่น บริเวณเส้นผมหรือขน หรือวัตถุที่โปร่งแสงบางส่วน พิกเซลในบริเวณดังกล่าว ไม่สามารถระบุว่าเป็นของวัตถุด้านหน้าหรือพื้นหลัง หากต้องการแยกส่วนของสีวัตถุด้านหน้าออกจากสีของพื้นหลังที่ผสมในพิกเซลเหล่านั้น จำเป็นต้องคำนวณค่า α และสีที่แท้จริงของวัตถุด้านหน้าที่ตรงกับพิกเซลดังกล่าว แต่เนื่องจากค่าระยะลึกไม่ได้มีความสัมพันธ์โดยตรงกับค่า α เราจึงไม่สามารถคำนวณค่า α จากค่าระยะลึกได้

2. ค่าระยะลึกที่วัดได้ด้วยกล้อง RGB-D มีความคลาดเคลื่อนจากความเป็นจริง มากน้อยแตกต่างกันออกไป โดยเฉพาะสำหรับ consumer-grade camera เช่น กล้อง Kinect จะมีความคลาดเคลื่อนที่สามารถสังเกตเห็นได้ด้วยตา เมื่อเปรียบเทียบขอบวัตถุที่ได้จากค่าระยะลึก กับขอบวัตถุที่แท้จริง ดังนั้น แม้แต่การแบ่งวัตถุด้านหน้าออกจากฉากหลัง กรณีที่เป็นวัตถุทึบแสง ก็ไม่สามารถใช้ข้อมูลระยะลึกในการแบ่งได้อย่างถูกต้องโดยลำพัง

ด้วยเหตุผลและข้อจำกัดดังกล่าวข้างต้น ทำให้แนวทางในการใช้ข้อมูลระยะลึกสำหรับการคำนวณภาพแมทโดยทั่วไป จะเป็นการใช้ข้อมูลดังกล่าว (โดยตรงหรือโดยอ้อมก็ตาม) ในการเสริมความถูกต้องให้กับการแบ่งขอบเขตภาพอย่างหยาบ (coarse segmentation) เพื่อระบุส่วนของวัตถุด้านหน้า และส่วนของพื้นหลัง ที่ไม่อยู่ในบริเวณใกล้เคียงหรือตรงรอยต่อระหว่างวัตถุสองส่วนดังกล่าว (คือส่วน foreground region และ background region ในภาพไตรแมบ) โดยวิธีทางตรงคือการใช้ข้อมูลระยะลึกในการคำนวณหรือสร้างภาพไตรแมบ ซึ่งกล่าวถึงในหัวข้อ 3.3 ส่วนวิธีการโดยอ้อมคือการใช้ข้อมูลระยะลึกเป็นส่วนหนึ่งในกระบวนการคำนวณภาพแมท

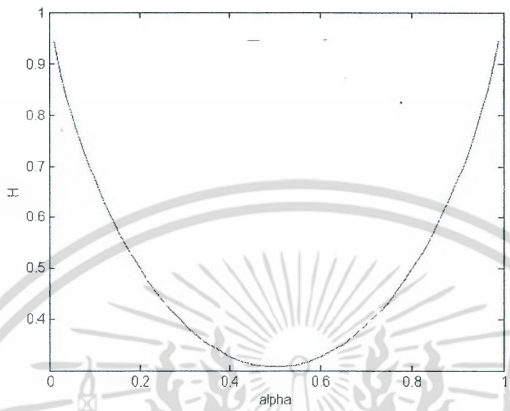
ใน [3] ผู้เขียนใช้ข้อมูลระยะลึกช่วยในการคำนวณภาพแมท โดยอาศัยสมมติฐานที่ว่ามีความเชื่อมโยงสัมพันธ์กันระหว่างค่า α กับความถูกต้องของกลุ่มข้อมูลระยะลึก กล่าวคือ สำหรับพิกเซลที่มีค่า α ที่สูงมาก หรือต่ำมาก (ใกล้ 1 หรือ 0) การแบ่งส่วนภาพโดยอาศัยข้อมูลระยะลึกจะมีความถูกต้องสูง เนื่องจากวัตถุในบริเวณดังกล่าวทึบแสงและไม่ได้อยู่บริเวณขอบหรือรอยต่อของวัตถุ ที่การวัดค่าระยะลึกมีโอกาสคลาดเคลื่อนสูง (ไม่ว่าจะใช้กล้องที่อาศัยการวัดค่าระยะลึกด้วยวิธีหรือเทคโนโลยีใดก็ตาม เช่น structured light stereoscopic sensor หรือ time-of-flight sensor) ในทางกลับกันค่า α ที่มีค่ากลาง ๆ (ใกล้เคียง 0.5) มักเกิดจากบริเวณผิววัตถุที่โปร่งแสงบางส่วน หรือบริเวณขอบวัตถุ ที่การวัดค่าระยะลึกมักมีความคลาดเคลื่อนสูงกว่ากรณีแรก จากสมมติฐานดังกล่าว ข้อมูลภาพในแต่ละพิกเซล ที่ประกอบไปด้วยค่าสีและค่าระยะลึก ก่อนถูกนำไปใช้ในการแก้สมการ (3.3) ส่วนของข้อมูลที่เป็นระยะลึกจะถูกนำมาคูณกับค่าน้ำหนักที่สะท้อนความน่าเชื่อถือของค่าระยะลึกนั้น ๆ ตามสมการต่อไปนี้

$$H(\alpha) = 1 + \alpha\log(\alpha) + (1-\alpha)\log(1-\alpha) \qquad (3.4)$$

โดยที่ α ในสมการข้างต้น คือค่า α ที่จุดหรือพิกเซลเดียวกับค่าระยะลึกที่ต้องการกำหนดค่าน้ำหนัก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 3.1 แสดงกราฟของค่าน้ำหนัก $H(.)$ ที่ค่า α ต่าง ๆ สังเกตว่าค่าน้ำหนักจะสูงเมื่อ α ใกล้ 0 หรือ 1 และต่ำสุดเมื่อ α เข้าหาค่า 0.5 โดยการใช้ค่าน้ำหนักดังกล่าว ทำให้ค่าระยะลึกในส่วนที่มีความถูกต้องต่ำ เมื่อนำไปคูณกับค่าน้ำหนักแล้ว จะมีผลต่อการคำนวณค่า α น้อยตามไปด้วย ทั้งนี้เพื่อไม่ให้ความผิดพลาดที่มีอยู่ในค่าระยะลึกส่งผลในทางลบต่อการคำนวณภาพแมท



ภาพที่ 3.1 แสดงกราฟของ $H(\alpha)$ เมื่อเทียบกับค่า α ที่แตกต่างกัน

ปัญหาที่สำคัญของวิธีการนี้ ในการนำไปใช้งาน คือ สมการที่ 3.4 ที่ใช้ในการคำนวณค่าน้ำหนัก ต้องอาศัยค่า α ซึ่งสำหรับพิกเซลบริเวณขอบวัตถุ (ซึ่งอยู่ใน unknown region ของภาพไตรแมบ) จะไม่ทราบมาก่อน ดังนั้นในทางปฏิบัติ การคำนวณในรอบแรก (1^{st} iteration) สำหรับพิกเซลเหล่านี้จำเป็นต้องอาศัยค่าเริ่มต้น เช่น $\alpha = 0.5$ ซึ่งจะมีผลทำให้ผลการคำนวณในรอบแรกนี้ ค่าระยะลึกในบริเวณดังกล่าวจะมีผลน้อยต่อผลลัพท์ที่ได้ หลังจากนั้นในรอบการประมวลผลถัด ๆ มา สามารถใช้ค่า α ที่ได้จากการคำนวณในรอบก่อนหน้านั้น สำหรับใช้แทนลงในสมการ 3.4 ได้

ข) วิธีการของ Lu [8]

เช่นเดียวกับวิธีการของ Wang วิธีการที่ 2 นี้ อาศัยค่าระยะลึกในการช่วยแบ่งขอบเขตระหว่างวัตถุด้านหน้ากับพื้นหลังร่วมกับข้อมูลค่าสี โดยอาศัยสมมติฐานต่อไปนี้

1. บริเวณพิกเซลในภาพที่ค่าสี foreground (F) มีค่าใกล้เคียงกับค่าสีที่วัดได้ (C) มีความเป็นไปได้สูงที่จะเป็นบริเวณของวัตถุด้านหน้า ในทำนองเดียวกัน บริเวณพิกเซลในภาพที่ค่าสี background (B) มีค่าใกล้เคียงกับค่าสีที่วัดได้ (C) มีความเป็นไปได้สูงที่จะเป็นบริเวณของวัตถุพื้นหลัง (หมายถึงทั้งสองกรณีมีความเป็นไปได้น้อยที่จะเป็นบริเวณรอยต่อระหว่างวัตถุด้านหน้าและพื้นหลัง) ดังนั้นค่าระยะลึกที่อ่านได้จากบริเวณดังกล่าว น่าจะมีความถูกต้องหรือน่าเชื่อถือสูง จึงควรให้น้ำหนักมากกับค่าระยะลึกของพิกเซลในบริเวณดังกล่าว

2. บริเวณพิกเซล ที่ค่าสี foreground (F) และค่าสี background (B) มีค่าใกล้เคียงกัน ควรจะให้น้ำหนักกับค่าระยะลึกของพิกเซลในบริเวณเดียวกันนั้นให้มาก เนื่องจากในกรณีนี้ ข้อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มูลค่าสีมีความน่าเชื่อถือหรือความแม่นยำต่ำในการแยกหรือจำแนกว่าบริเวณดังกล่าวเป็นวัตถุด้านหน้าหรือด้านหลัง และการคำนวณค่า α ก็มีแนวโน้มความคลาดเคลื่อนสูง การใช้ข้อมูลระยะลึกในบริเวณดังกล่าวจึงควรให้น้ำหนักมาก

หากพิจารณาสมมติฐานข้อแรก จะพบว่าเหมือนหรือสอดคล้องกับสมมติฐานที่ใช้ในวิธีการแรกของ Wang กล่าวคือบริเวณของภาพที่เป็นส่วนของวัตถุด้านหน้าหรือพื้นหลังที่ทึบแสง และไม่ใช่บริเวณรอยต่อ ค่าระยะลึกมีความถูกต้องสูงและควรให้ค่าน้ำหนักมาก หากแต่การระบุบริเวณดังกล่าว ใช้วิธีที่แตกต่างกัน (แต่สัมพันธ์กัน) กล่าวคือ ในขณะที่วิธีการของ Wang ใช้ค่า α มาเป็นตัววัด วิธีการของ Lu ใช้ผลต่างหรือความคล้ายคลึงกันระหว่าง F (B) กับ C แทน ซึ่งสัมพันธ์กับค่า α ในแง่ที่ว่า ค่า F (B) ที่ใกล้เคียงกับ C มีผลทำให้ค่า α เข้าใกล้ 1 (0 ในกรณีของ B) ซึ่งทั้งสองวิธี จะให้ค่าน้ำหนักกับค่าระยะลึกมากกว่าบริเวณอื่น ๆ ที่ ค่า F (B) แตกต่างจาก C ซึ่งในกรณีหลัง ค่าของ α จะเข้าใกล้ 0.5

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

$$w_d = exp(-\frac{w_F \cdot w_B}{\sigma_d^2}) \tag{3.5}$$

โดยที่ w_F (และ w_B) คำนวณจากสมการต่อไปนี้

$$w_F = 1 - exp(-\|F^* - C^*\|^2) \tag{3.6}$$

$$w_B = 1 - exp(-\|B^* - C^*\|^2) \tag{3.7}$$

จากสมการข้างต้น ในที่นี้ F^* และเวกเตอร์อื่น ๆ ในสมการ คือเวกเตอร์ที่มีเฉพาะข้อมูลสีเท่านั้น ไม่รวมข้อมูลระยะลึก ใน [8] ผู้เขียนได้เสนอการกำหนดค่า $\sigma_d = 0.1$

และเช่นเดียวกับวิธีการของ Wang สำหรับพิกเซลบริเวณ unknown region ค่า F (B) เป็นค่าที่ไม่ทราบและต้องการคำนวณหา ซึ่งวิธีการของ Lu หลีกเลี่ยงปัญหาดังกล่าว โดยในขั้นตอนของการคำนวณค่า F (B) ในสมการที่ (3.2) จะดำเนินการโดยไม่มีการคูณค่าน้ำหนักให้กับข้อมูลระยะลึกแต่อย่างใด (คือคำนวณเหมือนวิธีการดั้งเดิม ยกเว้นแต่เพิ่มค่าระยะลึกเข้ามานอกเหนือจากค่าสี) ต่อเมื่อจะทำการคำนวณค่า α ด้วยสมการ (3.3) จึงจะคำนวณค่าน้ำหนักด้วยสมการข้างต้น แล้วนำไปคูณเข้ากับข้อมูลระยะลึก (ค่าที่ 4 ใน F, B, และ C) ก่อนที่จะนำไปใช้คำนวณหาค่า α ตามสมการที่ (3.3) ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.2 การใช้ภาพระยะลึกกับเทคนิค Closed-form matting

Closed-form matting เป็นเทคนิคการคำนวณภาพแมทอีกวิธีการหนึ่งที่มีการอ้างอิงบ่อยครั้ง ส่วนหนึ่ง น่าจะเนื่องมาจากการที่เทคนิคดังกล่าว มีการคำนวณและการใช้งานที่ไม่ซับซ้อน และได้ผลดี โดยเฉพาะอย่างยิ่งสำหรับภาพที่ไม่มีลวดลาย ทั้งในส่วนวัตถุด้านหน้าและพื้นหลัง เนื่องจากเทคนิคดังกล่าวพัฒนาอยู่บนสมมติฐานที่ว่า ค่าสีไม่เปลี่ยนแปลงในบริเวณใกล้เคียงกันบนภาพ จากสมมติฐานดังกล่าว ค่า α สามารถคำนวณได้โดยการแก้สมการต่อไปนี้ [4]

$$\alpha \;=\; arg\,min\,\alpha^T L\alpha + \lambda(\alpha^T - b_s^T)D_s(\alpha - b_s) \tag{3.8}$$

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

$$L_{i,j} \;=\; \sum_{k|(i,j)\in w_k} \left\{ \delta_{i,k} - \frac{1}{|w_k|}\left(1 + (I_i - \mu_k)\left(R_k + \frac{\epsilon}{|w_k|}I_3\right)^{-1}(I_j - \mu_k)\right)\right\} \tag{3..9}$$

จากสมการข้างต้น R_k คือ covariance matrix สำหรับจุดหรือพิกเซลในภาพที่อยู่ภายในขอบเขตตามที่กำหนดในเซ็ต w_k ซึ่งกำหนดขอบเขตโดยใช้กรอบวินโดว์สี่เหลี่ยมเลื่อนไปตามจุดต่าง ๆ บนภาพต้นฉบับ (หากไม่นับจุดบริเวณขอบของภาพ จำนวนเซ็ต w_k จะเท่ากับจำนวนพิกเซล N ที่ต้องการคำนวณหาค่าแมท) ในสมการเดียวกัน ค่า ϵ คือค่าคงที่ที่เอาไว้ควบคุมหรือกำหนดอัตราการเปลี่ยนแปลงของค่า α ในบริเวณที่อยู่ข้างเคียงกัน (smoothing factor) ตลอดจนใช้เพื่อให้การประมวลผลมีเสถียรภาพ ส่วน I_3 ในสมการเดียวกัน คือเมทริกซ์ เอกลักษณ์ ขนาด 3×3 ในสมการ 3.8 ข้างต้น α คือเวกเตอร์ขนาด $N \times 1$ ที่ประกอบไปด้วยค่าแอลฟ่าของจุดต่าง ๆ ในภาพ

หลังจากคำนวณค่า α ได้แล้ว ค่าสีของวัตถุด้านหน้า และสีพื้นหลัง สามารถหาได้โดยการแก้สมการต่อไปนี้ [4]

$$\begin{aligned}(F^c, B^c) \;=\; \arg\min \sum_{i\in I} & (\alpha_i F_i^c + (1-\alpha_i)B_i^c - I_i^c)^2 \\ & + |\nabla_x \alpha_i|((\nabla_x F_i^c)^2 + (\nabla_x B_i^c)^2) + |\nabla_y \alpha_i|((\nabla_y F_i^c)^2 \\ & + (\nabla_y B_i^c)^2)\end{aligned} \tag{3.10}$$

จากสมการข้างต้น F^c คือเวกเตอร์ที่ประกอบไปด้วยค่าสี foreground ที่จุดที่ i ต่าง ๆ (F_i^c) โดย c แทน color channel ใด ๆ (red, green, blue) B^c ในสมการข้างต้นมีความหมายในทำนองเดียวกัน นอกจากนี้ สัญลักษณ์ $\nabla_x(\cdot)$ แทน derivative operator ตามแนวแกน x ส่วน $\nabla_y(\cdot)$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แทน derivative operator ตามแนวแกน y เพื่อให้ได้ค่าสีที่ครบถ้วน สมการข้างต้น ต้องกระทำ 3 ครั้ง กับ color channel ที่แตกต่างกัน (red, green, blue)

สำหรับแนวทางการใช้ค่าระยะลึกร่วมกับวิธีการดังกล่าว ในที่นี้ จะกล่าวถึงอยู่ 2 วิธี ดังต่อไปนี้

ก) วิธีการของ Pitie [5]

วิธีการนี้ใช้แนวทางเดียวกับการปรับใช้ค่าระยะลึกกับ Bayesian-based matting [2] กล่าวคือ เพิ่มข้อมูลระยะลึกเข้าไปในสมการเพื่อคำนวณค่า α ซึ่งในกรณีนี้ คือการเพิ่มค่าที่ 4 เข้าไปในเวกเตอร์ I_i ที่เดิมมีขนาด 3×1. ทำให้สมการ (3.9) เปลี่ยนเป็น

$$L_{i,j} = \sum_{k|(i,j)\in w_k} \left\{ \delta_{i,k} - \frac{1}{|w_k|}(1 + (I_i - \mu_k)(R_k + E_k)^{-1}(I_j - \mu_k)) \right\} \quad (3.11)$$

โดยที่

$$E_k = \frac{1}{|w_k|} \begin{bmatrix} \epsilon & 0 & 0 & 0 \\ 0 & \epsilon & 0 & 0 \\ 0 & 0 & \epsilon & 0 \\ 0 & 0 & 0 & \epsilon_D \end{bmatrix} \quad (3.12)$$

และเช่นเดียวกับวิธีการที่ผ่านมา คือจำเป็นต้องมีการกำหนดค่าน้ำหนักให้กับข้อมูลค่าระยะลึก โดย [5] เสนอการกำหนดค่าน้ำหนักด้วยการปรับเปลี่ยนค่า ϵ_D ในสมการข้างต้น โดยค่ามากจะมีผลทำให้น้ำหนักของข้อมูลระยะลึกมีผลต่อการคำนวณน้อย และในทางกลับกัน หาก ϵ_D มีค่าน้อย จะทำให้ค่าระยะลึกมีน้ำหนักต่อการคำนวณมาก อย่างไรก็ดี [5] ไม่ได้กำหนดหรือแนะนำวิธีในการคำนวณหรือหาค่า ϵ_D ที่เหมาะสมแต่อย่างใด

ข) วิธีการของ Sun

แทนที่จะเพิ่มข้อมูลระยะลึกเข้าไปในสมการเพื่อการคำนวณ เช่น ในกรณีของ [5] ในบทความ [6] อาศัยค่าข้อมูลระยะลึกในการปรับค่าพารามิเตอร์ ϵ แทน โดยกำหนดให้ค่า ϵ ไม่คงที่ตลอดทั้งภาพ สามารถแตกต่างกันได้ในแต่ละพิกเซล โดยมีค่าขึ้นอยู่กับระดับการผันแปรของค่าระยะลึกในแต่ละบริเวณของภาพที่ทำการคำนวณ ตามสมการต่อไปนี้

$$\epsilon_i = k_1 e^{-k_2 . var(d_i)} \quad (3.13)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ $var(d_i)$ คือค่าความผันแปร (variance) ของค่าระยะลึก ณ. พิกเซลที่อยู่รอบ ๆ จุด i ค่า ระยะลึกดังกล่าวถูกปรับให้มีค่าอยู่ระหว่าง 0 ถึง 1 โดยในบทความ [6] กำหนดให้ค่า k_1 = 0.1, และ $k_2 = 10^6$

จากสมการข้างต้น แนวคิดสมมติฐานอันเป็นที่มาของสมการคือ หากค่าความผันแปรของระยะลึกสูง มีแนวโน้มว่าบริเวณดังกล่าวจะเป็นบริเวณขอบวัตถุ ดังนั้น จึงกำหนดค่า ϵ_i ให้ต่ำ และหากค่าความผันแปรของระยะลุกต่ำ ค่า ϵ_i จะสูง ซึ่งค่า ϵ_i เป็นตัวควบคุมระดับความผันแปรของค่า α_i ที่คำนวณได้ หาก ϵ_i ต่ำทำให้ค่า α_i ที่คำนวณได้มีการเปลี่ยนแปลงได้มากในระยะใกล้ ๆ แต่หาก ϵ_i มีค่ามาก จะทำให้ α_i ที่คำนวณได้มีแนวโน้มค่อนข้างคงที่หรือเปลี่ยนแปลงน้อย หรืออีกนัยหนึ่ง ϵ_i เป็นพารามิเตอร์ที่มีลักษณะที่เรียกว่าเป็น smoothing factor ความสัมพันธ์ข้างต้น ถูกกำหนดจากสมมติฐานที่ว่า ความผันแปรของระยะลึกในบริเวณ foreground หรือ background จะค่อนข้างต่ำ เมื่อเปรียบเทียบกับความผันแปรในบริเวณขอบวัตถุ เช่นเดียวกันกับการผันแปรของค่า α_i ในบริเวณที่เป็น foreground หรือ background ที่ค่อนข้างต่ำหรือคงที่ ในขณะที่ค่า α_i ในบริเวณขอบวัตถุจะมีแนวโน้มที่สูง นำไปสู่ข้อสรุปว่าความผันแปรของค่าระยะลึกและค่า α_i มีแนวโน้มไปในทิศทางเดียวกัน อันเป็นพื้นฐานแนวความคิดของการกำหนดค่า ϵ_i ตามสมการข้างต้น อนึ่งพึงสังเกตว่าสมมติฐานดังกล่าวมีความคล้ายคลึงสอดคล้องกับที่ใช้ในวิธีการอื่น ๆ ดังที่ได้กล่าวมาแล้วข้างต้น

ในบทความ [6] ไม่ได้กล่าวถึงการใช้ข้อมูลระยะลึกในการแก้สมการที่ 3.9 โดยตรงเหมือนกรณีของ [5] อย่างไรก็ดี ในงานวิจัยนี้ จะได้ทำการทดสอบวิธีการนี้ทั้งโดยการใช้กับไม่ใช้ข้อมูลระยะลึกเป็นข้อมูลลำดับที่ 4 ในเวกเตอร์ I_j นอกจากเหนือจากค่าสี RGB

3.3 การสร้างแผนภาพไตรแมปโดยอัตโนมัติจากข้อมูลระยะลึก

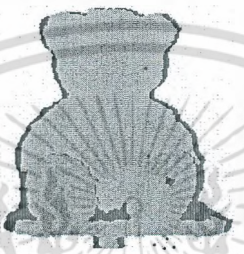
ดังที่ได้กล่าวมาแล้วในช่วงต้น ว่าการใช้ประโยชน์จากภาพระยะลึก ในการแยกวัตถุจากพื้นหลังธรรมชาติ หรือ digital matting นั้น สามารถแบ่งออกได้เป็น 2 แนวทางหลัก ๆ คือ การใช้ข้อมูลดังกล่าวเป็นส่วนหนึ่งในการคำนวณภาพแมท ดังเช่น วิธีการต่าง ๆ ที่ได้กล่าวไปแล้วในหัวข้อ 3.2 ที่ผ่านมา กับอีกแนวทางหนึ่ง คือการใช้ข้อมูลระยะลึกในการแบ่งส่วนภาพออกเป็น ส่วนวัตถุด้านหน้า ส่วนพื้นหลัง และส่วนรอยต่อของขอบวัตถุกับพื้นหลัง หรือการสร้างแผนภาพไตรแมปตามที่เรียกกันโดยทั่วไป ซึ่งประโยชน์ของการใช้ข้อมูลระยะลึกในการสร้างไตรแมป คือใช้แทนการสร้างแผนภาพดังกล่าวโดยการกำหนดหรือวาดด้วยมือ ในหัวข้อนี้ จะกล่าวถึงขั้นตอนโดยทั่วไปในการสร้างแผนภาพไตรแมปจากภาพระยะลึกโดยอัตโนมัติหรือกึ่งอัตโนมัติ จากนั้นจะได้กล่าวถึงปัญหาในการสร้างภาพไตรแมปจากข้อมูลระยะลึกในทางปฏิบัติ และเทคนิคที่พัฒนาขึ้นเพื่อลดปัญหาดังกล่าว ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.1 ขั้นตอนการสร้างภาพไตรแมบจากภาพระยะลึก

สำหรับขั้นตอนการสร้างภาพไตรแมบจากข้อมูลระยะลึกเบื้องต้นจะสามารถทำได้ตามดังต่อไปนี้

1. ทำการถ่ายภาพโดยใช้อุปกรณ์ที่มีความสามารถในการวัดระยะลึก สำหรับรายงานวิจัยฉบับนี้ใช้อุปกรณ์ Kinect ซึ่งใช้หลักการ structured light ในการคำนวณหาค่าระยะลึกของภาพวัตถุ



ภาพที่ 3.2 ตัวอย่างภาพระยะลึกจากอุปกรณ์ Kinect

2. ข้อมูลภาพระยะลึกของอุปกรณ์ Kinect โดยทั่วไปจะมีบริเวณที่ไม่ค่าข้อมูลระยะลึกสืบเนื่องมาจากการที่วัตถุบริเวณดังกล่าวไม่สะท้อนแสงหรือการที่บริเวณดังกล่าวเกิดจากการบังของวัตถุเบื้องหน้าทำให้เกิดเป็นเงาของวัตถุทำให้ไม่สามารถวัดข้อมูลระยะลึกบริเวณดังกล่าว ในการแก้ไขเบื้องต้นอย่างง่ายๆ จะอาศัยการเฉลี่ยข้อมูลระยะลึกบริเวณข้างเคียงแบบถ่วงน้ำหนักเชิงพื้นที่

ภาพที่ 3.3 ตัวอย่างภาพระยะลึกที่เติมค่าจุดที่ขาดหายโดยการประมาณค่าระยะลึกจากบริเวณใกล้เคียง

3. หากลุ่มของข้อมูลภาพระยะลึกที่มีค่าระยะลึกน้อยที่สุดจากสมมติฐานที่ว่ากลุ่มของภาพระยะลึกที่มีค่าน้อยที่สุดคือบริเวณที่เป็นกลุ่มของภาพวัตถุที่สนใจ สามารถทำได้โดยใช้ K-mean ในการหาค่าเทรดโชลด์ (threshold) เพื่อหากลุ่มของข้อมูลภาพระยะลึกที่น้อยที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 3.4 ภาพที่ได้จากการทำแบ่งกลุ่มพิกเซลด้วย K-mean

4. กำหนดขนาดขอบเขตของ unknown region ตามขอบภาพที่ได้โดยอาศัยค่าเทรดโชลด์ ดังภาพที่ 3.4 โดยทั่วไปแผนภาพระยะลึกจะมีความคลาดเคลื่อนหรือผิดพลาดไปจากวัตถุตามความเป็นจริง ซึ่งเกิดจากความถูกต้องของแต่ตัวอุปกรณ์ โดยที่ในแต่ละระยะความลึกจะมีค่าความคลาดเคลื่อนหรือค่าความผิดพลาดที่แตกต่างกันไป สำหรับการกำหนดขนาดของ unknown region ในภาพไตรแมบอย่างง่าย จะอาศัยค่าความผิดพลาดของตัวอุปกรณ์เป็นตัวกำหนดโดยสามารถหาได้จากการวัดขนาดของบริเวณที่คลาดเคลื่อนไปจากความเป็นจริงในแต่ละระยะ

ภาพที่ 3.5 ตัวอย่างแผนภาพไตรแมบที่สร้างจากข้อมูลระยะลึกแบบอัตโนมัติ

3.3.2 ปัญหาและข้อจำกัดในการสร้างภาพไตรแมบ

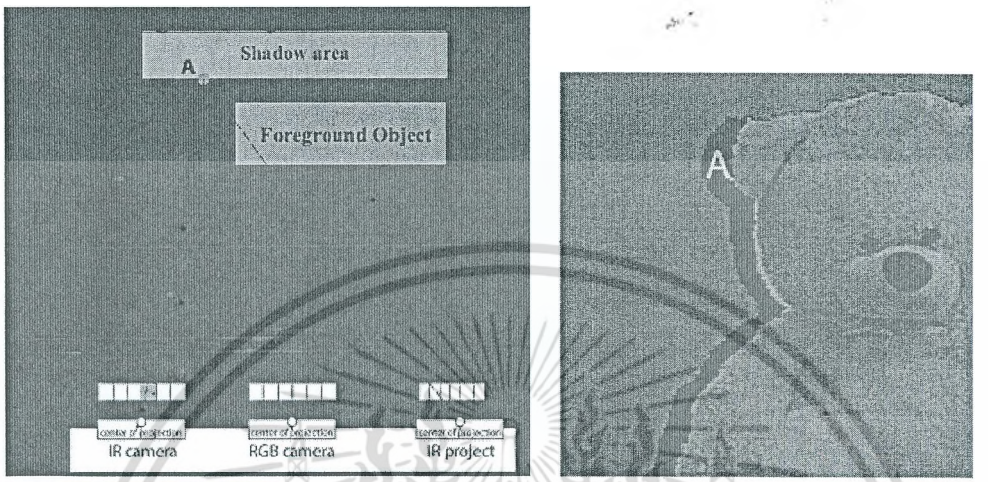
เนื่องด้วยการสร้างภาพไตรแมบที่อาศัยข้อมูลระยะลึกที่กล่าวมาในขั้นตอนนั้นเกิดจากการที่อาศัยค่าความผิดพลาดของตัวอุปกรณ์เป็นตัวกำหนด ที่ได้จากการวัดขนาดของบริเวณที่คลาดเคลื่อนไปจากความเป็นจริงในแต่ละระยะ แต่หากข้อมูลระยะลึกที่ได้มีความผิดพลาดไปจากความเป็นจริงโดยเฉพาะบริเวณที่ไม่สามารถวัดค่าระยะความลึกได้ซึ่งสามารถแบ่งได้ 2 กรณีคือ

1. การที่วัตถุไม่สะท้อนแสง
2. บริเวณที่เกิดจากการบังของวัตถุเบื้องหน้า

ในหัวข้อที่ 3.3.1 ใช้วิธีการแก้ไขอย่างง่ายๆ โดยการเฉลี่ยแบบถ่วงน้ำหนักจากค่าระยะลึกของบริเวณที่มีข้อมูลระยะลึกข้างเคียง ซึ่งข้อมูลระยะลึกที่ได้รับการแก้ไขบริเวณนั้นอาจเกิดความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เคลื่อนไปจากความเป็นจริงเพราะเกิดจากเฉลี่ย ทำให้การกำหนดขอบเขตของ unknown region ที่อาศัยค่าระยะลึกนั้นผิดพลาดไปด้วยเช่น ขอบเขตของ unknown region มีขนาดกว้างจนเกินไปซึ่งจะส่งผลต่อประสิทธิภาพและความถูกต้องในการสร้างภาพแมท



(ก) (ข)

ภาพที่ 3.6 ภาพตัวอย่างบริเวณที่เกิดจากการบังของวัตถุเบื้องหน้าของอุปกรณ์ Kinect (ก) จุด A คือบริเวณที่อยู่ในเงาของวัตถุ Foregroundซึ่งไม่ได้รับแสงอินฟราเรดของ Kinect (ข) จุด A คือบริเวณที่เป็นเงาของภาพที่เกิดจากภาพระยะลึก ซ้อนทับกับภาพสี

3.3.3 การปรับปรุงคุณภาพของภาพระยะลึกเพื่อการสร้างแผนภาพไตรแมบ

จากปัญหาดังที่กล่าวมาในหัวข้อที่ 3.3.2 ซึ่งเกิดจากการที่ข้อมูลระยะลึกผิดพลาดไปจากความเป็นจริงทำให้ขอบเขตของ unknown region ในภาพไตรแมบผิดพลาดหรือมีขนาดของ unknown region ที่ไม่เหมาะสม รายงานวิจัยนี้เสนอแนวทางการปรับปรุงคุณภาพของแผนภาพระยะลึกที่ได้จากอุปกรณ์ Kinect ในการค่าข้อมูลระยะลึก โดยอาศัยแนวทางหลักการทำงานของ structured light ในการตรวจสอบว่าค่าข้อมูลระยะลึกที่ขาดหายไปนั้นเกิดจากการบังของวัตถุหรือไม่ หากเกิดจากการบังของวัตถุจะสามารถสรุปได้ว่าบริเวณดังกล่าวมีค่าระยะลึกเท่ากับค่าระยะลึกของภาพพื้นหลังบริเวณใกล้เคียง แต่ถ้าไม่ได้เกิดจากการที่วัตถุบังจะสามารถสรุปได้ว่าบริเวณดังกล่าวเกิดจากการที่วัตถุไม่สะท้อนแสงซึ่งจะสามารถแก้ไขได้โดยการเฉลี่ยข้อมูลระยะลึกข้างเคียงแบบถ่วงน้ำหนักทั้งหมด สำหรับขั้นตอนการตรวจสอบสามารถทำได้ดังต่อไปนี้

1. กำหนดตัวแปรของข้อมูลรูปภาพสีเป็น $I_c(m,n)$ โดยที่ I_c คือรูปภาพสีที่ได้จากอุปกรณ์ Kinect หรือกล้อง DSLR และกำหนดตัวแปรของข้อมูลแผนภาพระยะลึกคือ โดยที่ I_d คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แผนภาพระยะลึก โดยที่ $m = 0,1 \ldots, M-1$ และ $n = 0,1 \ldots, N-1$ โดยที่ M คือความสูง และ N คือความกว้างของรูปภาพ

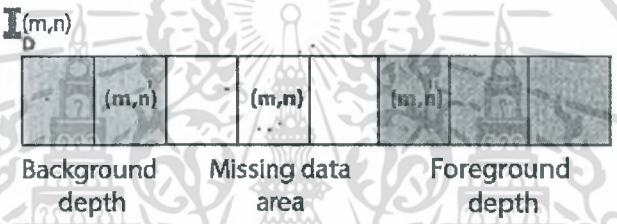
2. ใช้ข้อมูลของแผนภาพระยะลึกของวัตถุในการเป็นเกณฑ์แบ่งบริเวณแผนภาพระยะลึกของวัตถุและแผนภาพระยะลึกของพื้นหลัง

3. ตรวจสอบบริเวณที่ไม่มีข้อมูลระยะความลึกในแต่ละแถวของ $I_d(m,n)$ โดยกำหนดให้ $n' < n < n''$ ตามเงื่อนไขดังต่อไปนี้

1) $I_d(m;n)$ ณ. ตำแหน่งดังกล่าวไม่สามารถวัดค่าระยะลึกได้ โดยมีค่า NaN

2) $(m,n') \in B$

3) $(m,n'') \in F$



ภาพที่ 3.7 ภาพตัวอย่างบริเวณที่มีโอกาสเกิดจากการบังของวัตถุ

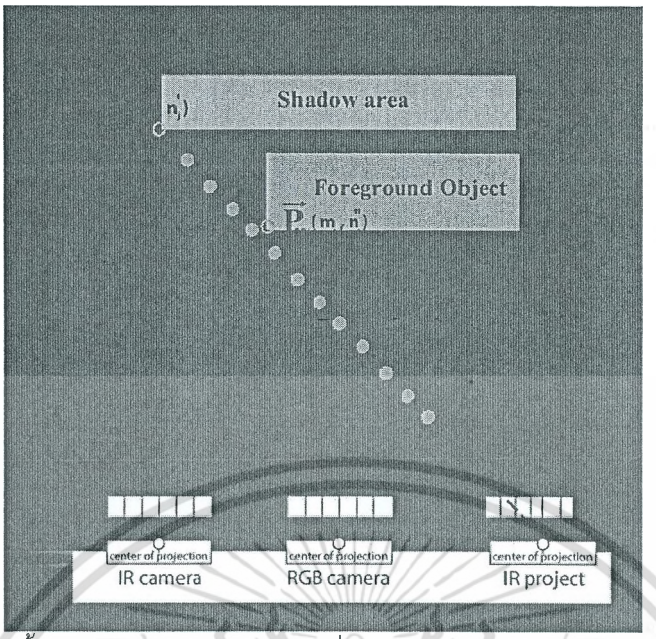
4. ทำขั้นตอนที่ 3 จนครบทุกแถวของแผนภาพระยะลึก กำหนดให้ S คือเซตของ tuples (m_j, n_j', n_j'') ที่ M_j และ $j = 0,1, \ldots, J-1$ คือแถวของตำแหน่งพิกเซล $I_d(m,n)$.จำนวน J ตัว และ n_j', n_j'' คือตำแหน่งพิกเซลที่อยู่ทางซ้ายและทางขวาที่ใกล้กับ $I_d(m,n)$.

5. ทำตามขั้นตอนต่อไปนี้เพื่อตรวจสอบในแต่ละสมาชิก (m_j, n_j', n_j'') ของ S ว่าเป็นเงาของวัตถุหรือไม่

5.1 กำหนดให้ $\vec{P}_w(m_j, n_j')$ คือตำแหน่งในความเป็นจริง (world coordinate) ของ $I_d(m_j, n_j')$ และให้ $\vec{P}_R$ คือตำแหน่งของ IR โปรเจคเตอร์ในความเป็นจริงโดยที่ $\vec{P}_R = [0.05, 0, 0]$

5.2 ทำการแบ่งจุดระหว่าง $\vec{P}_w(m_j, n_j')$ และ $\vec{P}_R$ โดยกำหนดให้เป็น $\vec{P}_j(i)$ ที่ i=0,1,2,...,K-1 และ K คือจำนวนของจุดระหว่าง $\vec{P}_w(m_j, n_j')$ และ $\vec{P}_R$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.8 ภาพขั้นตอนการตรวจสอบบริเวณที่เกิดจากการบังของวัตถุหรือเป็นเงาของวัตถุ

5.3 กำหนดให้ $\vec{P}_w(m_k, n_k'')$ คือจุดในความเป็นจริงของ $I_d(m_k, n_k'')$ ที่ k=0,1,2,..,J-1

5.4 กำหนดให้ e_j คือระยะห่างที่ใกล้ที่สุดระหว่าง $\vec{P}_j(i)$ และ $\vec{P}_w(m_j, n_j'')$ จากสมการต่อไปนี้

$e_j(i) =$	$\min_{k=0,..,J-1} \left\|\vec{P}_j(i) - \vec{P}_w(m_k, n_k'')\right\|$	(3.14)
$e_j =$	$\min_{i=0,..,K-1} e_j(i)$	(3.15)

5.5 กำหนดให้ t คือค่าที่ใช้เป็นเกณฑ์ในการตรวจสอบเงา ถ้า e_j มีค่าน้อยกว่า t จะสรุปได้ว่าบริเวณดังกล่าวนั้นเกิดจากเงาของวัตถุซึ่งสามารถแก้ไขข้อมูลระยะลึกบริเวณดังกล่าวได้จากการเฉลี่ยค่าระยะความลึกบริเวณข้างเคียงเฉพาะของภาพพื้นหลัง แต่หากค่า e_j มีค่ามากกว่า t แสดงว่าบริเวณดังกล่าวไม่ได้เกิดจากเงาของวัตถุแต่เป็นการที่วัตถุไม่สะท้อนแสง จะให้ข้อมูลระยะลึกได้จากการเฉลี่ยแบบถ่วงน้ำหนักเชิงพื้นที่ข้างเคียงจากข้อมูลระยะลึกของวัตถุและภาพพื้นหลัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
การทดลอง

สำหรับการแยกพื้นหลังธรรมชาติ โดยอาศัยข้อมูลระยะลึกจากบทที่ 3 สามารถแบ่งเป็นการทดลองเพื่อประเมินประสิทธิภาพได้ดังนี้

1) วิธีการในกลุ่มที่ใช้หลักการ Bayesian Matting ประกอบด้วยวิธีการต่อไปนี้

- Original Bayesian method (OBAY) คือวิธีการที่ไม่ได้อาศัยข้อมูลภาพระยะลึก
- Depth-assisted Bayesian method ของ Wang (DBAY-W) คือวิธีการของ Wang [3] ตามที่อธิบายในหัวข้อ 3.2.1
- Depth-assisted Bayesian method ของ Lu (DBAY-L) คือวิธีการใน [8] ตามที่อธิบายในหัวข้อ 3.2.1

2) วิธีการในกลุ่มที่ใช้หลักการ Closed-form Matting ประกอบด้วยวิธีการต่อไปนี้

- Original closed-form matting (OCF) คือวิธีการเดิมที่ไม่ได้อาศัยข้อมูลระยะลึก
- Depth-assisted closed-form matting method ของ Pitie (DCF-P) คือวิธีการใน [5] ตามที่ได้อธิบายไปในหัวข้อ 3.2.2
- Depth-assisted closed-form matting method ของ Sun ทั้ง 2 วิธี (DCF-S1, DCF-S2) ตามที่อธิบายในหัวข้อ 3.2.2 โดย DCF-S1 คือวิธีการที่อาศัยข้อมูลนระยะลึกเป็นองค์ประกอบที่ 4 ในเวกเตอร์ I_j ในขณะที่ DCF-S2 ใช้เพียงค่าสีเท่านั้น

ข้อมูลที่ใช้ในการทดลองสำหรับข้อ 1) และ 2) จะแบ่งข้อมูลที่ใช้ในการทดลองออกเป็น 2 ประเภท คือ การทดลองกับภาพจากฐานข้อมูลมาตรฐานที่ใช้ในการทดสอบวิธีการต่าง ๆ (www.alphamatting.com) และ การทดลองกับภาพที่ถ่ายด้วยกล้อง Kinect+DSLR ซึ่งเป็นกล้อง Kinect ที่มีการติดตั้งกล้องถ่ายภาพแบบ DSLR (Canon EOS รุ่น kiss X4 เลนส์ kit 18-55 ความยาวโฟกัสที่ใช้คือ 18 mm) ไว้เหนือกล้อง Kinect โดยภาพระยะลึกเป็นภาพที่ได้จาก Depth Camera ที่อยู่ภายในตัวกล้อง Kinect ส่วนภาพสีจะใช้ภาพที่ถ่ายด้วยกล้อง DSLR แทนภาพที่ได้จากกล้อง Kinect เพื่อให้ได้คุณภาพของภาพที่ดีกว่าที่ได้จากกล้องสีของตัวอุปกรณ์ Kinect ทั้งนี้ ก่อนทำการทดลอง ได้มีการปรับตั้งค่า (calibration) เพื่อให้ตำแหน่งในภาพระยะลึกตรงกับกับตำแหน่งที่ปรากฏในภาพสี

ในส่วนของข้อมูลจากฐานข้อมูลบนเว็บ www.alphamatting.com มีการให้ข้อมูลภาพไตรแมทและมีข้อมูลภาพแมทที่ถูกต้องมาด้วย ทำให้สามารถคำนวณและเปรียบเทียบค่าที่คำนวณได้กับค่าจริงออกมาเป็นตัวเลขได้ ส่วนภาพที่ถ่ายด้วยกล้อง Kinect จะอาศัยการพิจารณาประเมินคุณภาพจากภาพแมท

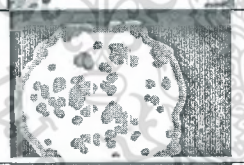
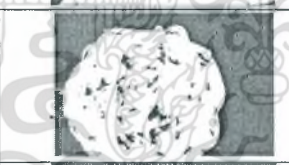
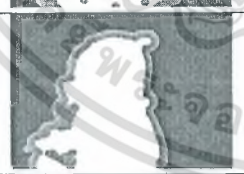
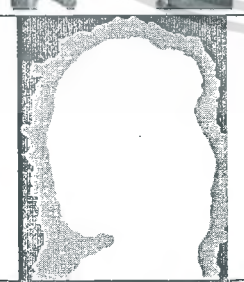
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่คำนวณได้ และเนื่องจากภาพจากฐานข้อมูลฯ ไม่ได้ให้ข้อมูลภาพระยะลึกมาด้วย ในการทดลอง จึงจำลองโดยอาศัยแนวเส้นกึ่งกลางระหว่างขอบเขตของ foreground region และขอบเขตของ background region ตามที่ปรากฏในภาพไตรแมท ที่มีมาให้ในฐานข้อมูลดังกล่าว เพื่อใช้เป็นเส้นขอบเขตที่แบ่ง foreground region และ background region และจากเส้นแบ่งดังกล่าว ทำการจำลองสร้างภาพระยะลึกขึ้นมา โดยให้บริเวณที่อยู่ด้านในของเส้นแบ่ง มีค่าระยะลึกเป็น 0.5 และบริเวณด้านนอกของเส้นแบ่งมีค่าระยะลึกเป็น 1 หลังจากนั้น ทำการจำลองสัญญาณรบกวน เพื่อบวกเข้าไปกับค่าที่ได้ เพื่อใช้เป็นข้อมูลจำลองภาพระยะลึก ที่จำเป็นสำหรับวิธีการต่าง ๆ ที่อาศัยข้อมูลระยะลึกช่วยในการประมวลผล

3) การทดลองปรับปรุงคุณภาพของภาพระยะลึกเพื่อการสร้างแผนภาพไตรแมบอัตโนมัติ

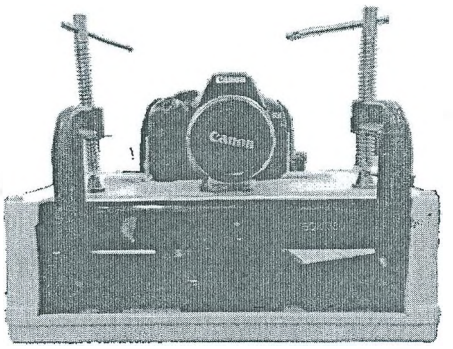
4.1 การทดสอบเทคนิคในกลุ่ม Bayesian-based Matting Methods

4.1.1 การทดสอบเทคนิค Original Bayesian Matting (OBAY) กับภาพจากฐานข้อมูลฯ จำนวน 5 รูป มีค่า Root Means Square Error (RMSE) เฉลี่ยทุกภาพอยู่ที่ 0.1080 และ Sum of Absolute Difference (SAD) ที่ 7451 โดยผลการทดสอบสำหรับแต่ละรูปแสดงดังในภาพที่ 4.1 ส่วนผลการทดลองกับรูปที่ถ่ายด้วยกล้อง Kinect+DSLR (ดูภาพที่ 4.2) จำนวน 3 รูป แสดงดังในภาพที่ 4.3

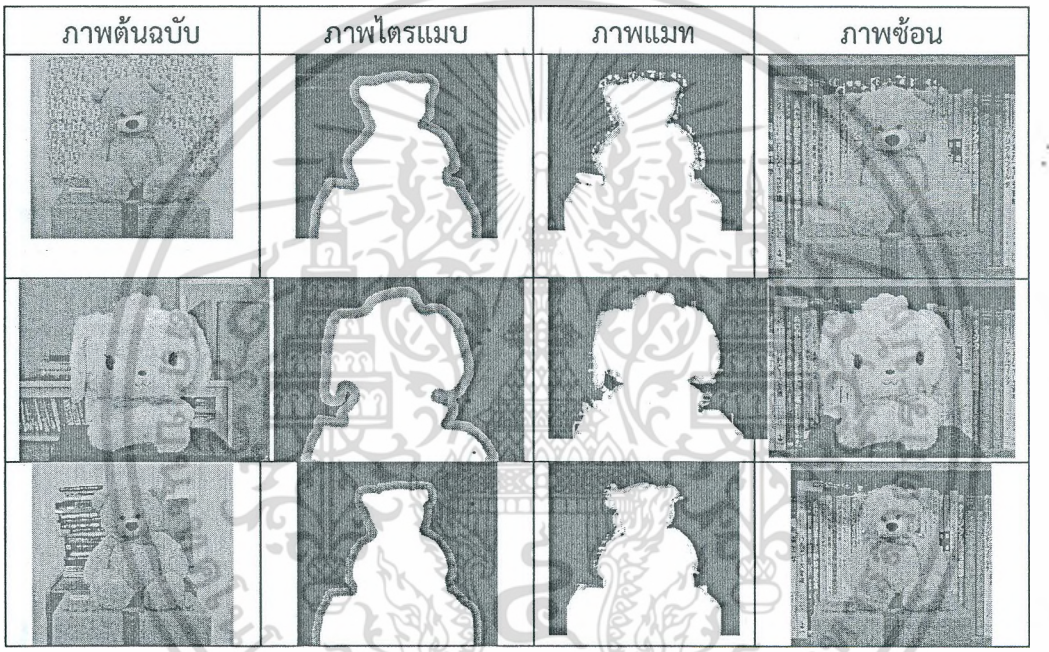
ภาพต้นฉบับ	ภาพไตรแมบ	ภาพแมท	RMSE	SAD
			0.0418	2496
			0.1635	14657
			0.0747	3040
			0.1375	17060
			0.1226	5785

ภาพที่ 4.1 ภาพต้นฉบับ ภาพไตรแมบ และภาพแมทที่สร้างขึ้นด้วยวิธี OBAY และค่า RMSE, SAD ที่ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 4.2 อุปกรณ์ Kinect+DSLR ที่ใช้ในการทดลอง



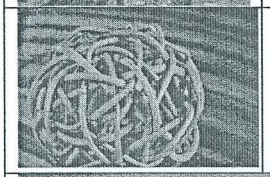
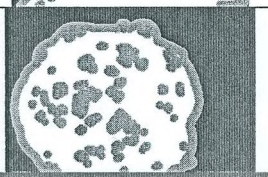
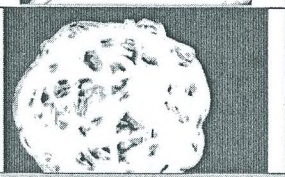
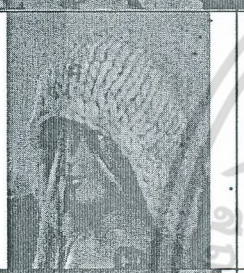
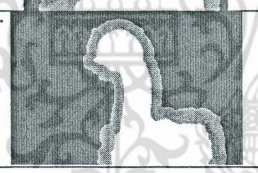
ภาพที่ 4.3 ภาพสีจากกล้อง Kinect ภาพไตรแมบ ภาพแมทและภาพซ้อนที่สร้างขึ้นด้วยวิธีการ OBAY

ผลการทดลองกับภาพถ่ายจริงด้วยวิธี OBAY พบปัญหาบริเวณที่มีค่าสีของวัตถุกับพื้นหลังที่ใกล้เคียงกัน และบริเวณที่พื้นหลังมีลวดลายละเอียดที่มีหลากหลายสี

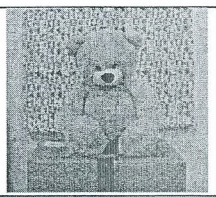
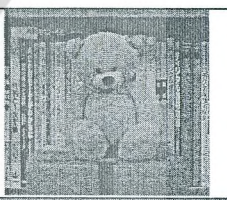
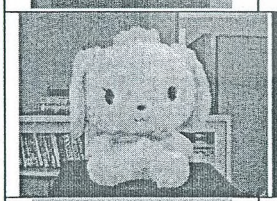
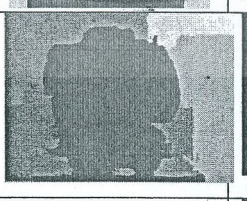
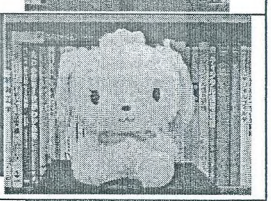
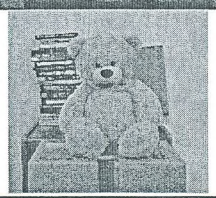
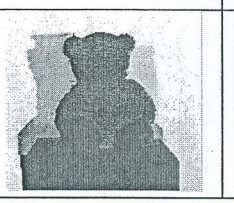
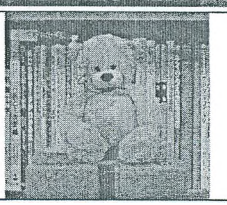
4.1.2 ผลการทดลองกับวิธีการของ Wang (DBAY-W) โดยใช้ภาพจากฐานข้อมูลฯ จำนวน 5 รูป ค่า RMSE เฉลี่ยคือ 0.1040 และค่าเฉลี่ยของ SAD เท่ากับ 7173 โดยผลการทดสอบสำหรับแต่ละรูปแสดงดังในภาพที่ 4.4 ซึ่งเมื่อเปรียบเทียบกับ OBAY แล้ว ในแง่ของค่า RMSE ให้ผลที่ไม่แตกต่างกันแต่อย่างใด

ผลการทดลองวิธี DBAY-W กับภาพจากกล้อง Kinect+DSLR จำนวน 3 รูป แสดงดังในภาพที่ 4.5 จากผลการทดลอง สังเกตว่าสามารถคำนวณภาพแมทได้ถูกต้องมากขึ้น ในบริเวณที่เป็นปัญหาสำหรับ OBAY เช่น บริเวณที่มีสีบางส่วนของพื้นหลังใกล้เคียงกับวัตถุ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพต้นฉบับ	ภาพไตรแมบ	ภาพแมท	RMSE	SAD
			0.0418	2425
			0.1635	14152
			0.0647	3284
			0.1475	16005
			0.1026	4761

ภาพที่ 4.4 ภาพต้นฉบับ ภาพไตรแมบ และภาพแมทที่สร้างขึ้นด้วยวิธี DBAY-W และค่า RMSE, SAD ที่ได้

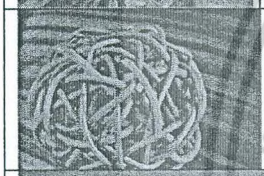
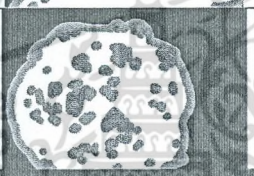
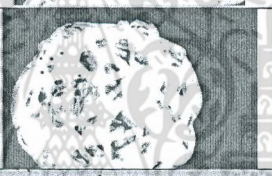
ภาพต้นฉบับ	ภาพไตรแมบ	ภาพระยะลึก	ภาพแมท	ภาพซ้อน
				
				
				

ภาพที่ 4.5 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมบ ภาพแมทและภาพซ้อนด้วยวิธีการ DBAY-W

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

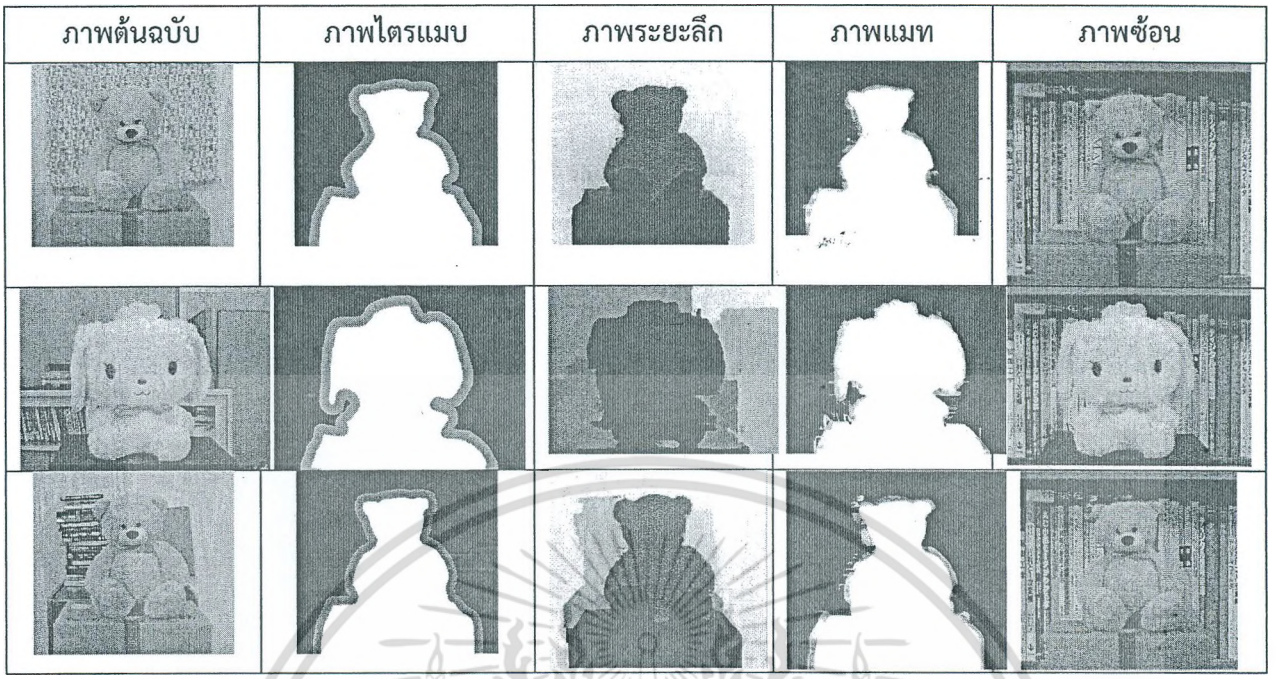
4.1.3 ผลการทดลองด้วยวิธีการของ Lu (DBAY-L) กับภาพจากฐานข้อมูลมาตรฐานจำนวน 5 รูป มีค่า RMSE เฉลี่ยคือ 0.1733 และค่าเฉลี่ยของ SAD เท่ากับ 14831 โดยผลการทดสอบสำหรับแต่ละรูปแสดงดังในภาพที่ 4.6 ซึ่งเมื่อเปรียบเทียบกับ OBAY และ DBAY-W แล้ว สองวิธีการก่อนหน้านี้ให้ผลที่ดีกว่า

ผลการทดลองวิธี DBAY-L กับภาพจากกล้อง Kinect+DSLR จำนวน 3 รูป แสดงดังในภาพที่ 4.7 จากผลการทดลอง พบว่าโดยรวมให้ผลที่ด้อยกว่าสองวิธีการแรก

ภาพต้นฉบับ	ภาพไตรแมบ	ภาพแมท	RMSE	SAD
			0.1509	9971
			0.3815	32820
			0.2201	9984
			0.2329	21379
			0.2529	9610

ภาพที่ 4.6 ภาพต้นฉบับ ภาพไตรแมบ และภาพแมทที่สร้างขึ้นด้วยวิธี DBAY-L และค่า RMSE, SAD ที่ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

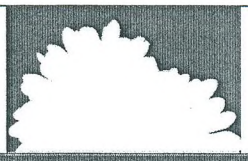
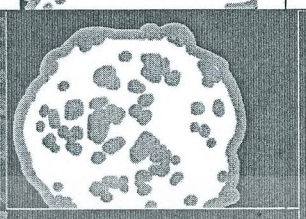
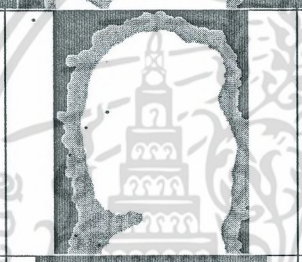


ภาพที่ 4.7 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมบ ภาพแมทและภาพซ้อนด้วยวิธีการ DBAY-L

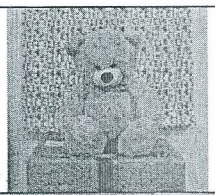
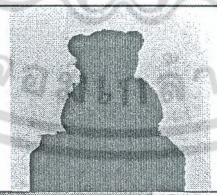
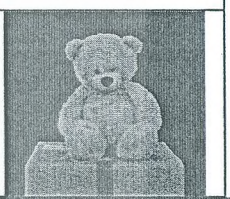
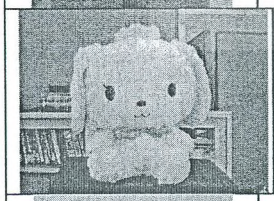
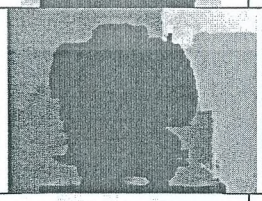
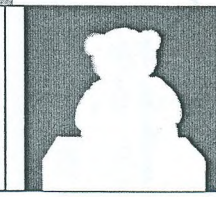
4.2 การทดสอบเทคนิคในกลุ่ม Closed-form Matting Methods

4.2.1 การทดสอบเทคนิค Original Closed-form Matting (OCF) กับภาพจากฐานข้อมูลมาตรฐานจำนวน 5 รูป เช่นเดียวกับที่ใช้ในการทดลองตามข้อ 4.1 มีค่า RMSE ของภาพแมทเฉลี่ยจากทุกภาพอยู่ที่ 0.1513 และค่าเฉลี่ยของ SAD เท่ากับ 3921 โดยผลการทดสอบสำหรับแต่ละรูปแสดงดังในภาพที่ 4.8 นอกจากนี้ในการทดลอง ให้ทำการทดสอบวิธีการดังกล่าว กับรูปภาพจากฐานข้อมูลเดียวกันเพิ่มเติมอีก 10 รูป ได้ค่า RMSE ของภาพแมทเฉลี่ยเท่ากับ 0.1376 และค่าเฉลี่ยของ SAD เท่ากับ 7206 ส่วนผลการทดลองกับรูปที่ถ่ายด้วยกล้อง Kinect+DSLR จำนวน 3 รูป แสดงดังในภาพที่ 4.9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพต้นฉบับ	ภาพไตรแมป	ภาพแมท	RMSE	SAD
			0.1054	1373
			0.2160	6336
			0.1133	1322
			0.1817	9011
			0.1399	1561

ภาพที่ 4.8 ภาพต้นฉบับ ภาพไตรแมป และภาพแมทที่สร้างขึ้นด้วยวิธี OCF และค่า RMSE, SAD ที่ได้

ภาพต้นฉบับ	ภาพไตรแมป	ภาพระยะลึก	ภาพแมท	ภาพซ้อน
				
				
				

ภาพที่ 4.9 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมป ภาพแมทและภาพซ้อนด้วยวิธีการ OCF

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2 การทดสอบเทคนิคของ Pitie (DCF-P) ใช้ภาพจากฐานข้อมูลมาตรฐานจำนวน 5 รูป ที่ใช้ในทุกการทดลองที่ผ่าน และทำการทดสอบเพิ่มเติมกับอีก 10 รูป ซึ่งเป็นภาพชุดเดียวกับที่ใช้ในการทดลอง 4.2.1 โดยกำหนดให้

1. $n_{f,d}$ คือขนาดของ filter ที่ใช้ในการกรองภาพระยะลึกเพื่อให้รอยต่อไม่เปลี่ยนแปลงอย่างทันทีทันใด

2. ϵ_D คือ การกำหนดค่าน้ำหนักให้กับข้อมูลค่าระยะลึก หาก ϵ_D ค่ามากจะมีผลทำให้น้ำหนักของข้อมูลระยะลึกมีผลต่อการคำนวณน้อย และในทางกลับกัน หาก ϵ มีค่าน้อย จะทำให้ค่าระยะลึกมีน้ำหนักต่อการคำนวณมาก

3. s_d คือค่าที่กำหนดขนาดสเกลของค่าระยะลึก หรือกล่าวได้ว่าค่าระยะลึกจะถูกปรับให้มีค่าสูงสุดเท่ากับ $1/s_d$

ตารางที่ 4.1 แสดงค่าเฉลี่ยของ RMSE และ SAD ทั้งกรณีของชุดภาพ 5 รูป และชุดภาพ 10 รูป

ตารางที่ 4.1 ค่าเฉลี่ยของ RMSE และ SAD ที่ค่าพารามิเตอร์ต่าง ๆ โดยใช้วิธี DCF-P

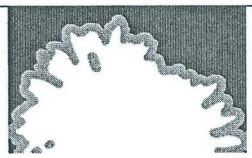
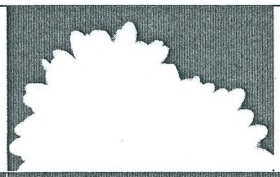
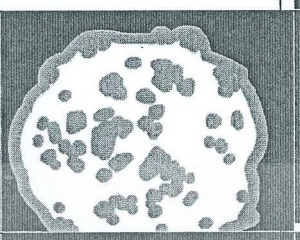
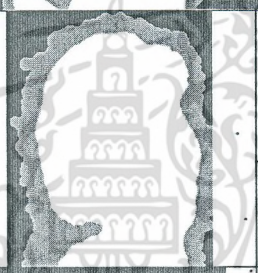
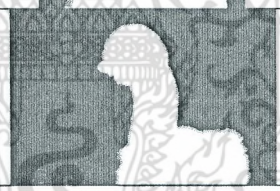
Parameter	RMSE / SAD (5 รูป)	RMSE / SAD (10 รูป)
$n_{f,d}$ = 1, ϵ_D = 1, s_d = 1	0.1377 / 3929	0.1379 / 7220
$n_{f,d}$ = 1, ϵ_D = 10, s_d = 1	0.1371 / 3940	0.1377 / 7226
$n_{f,d}$ = 1, ϵ_D = 20, s_d = 1	0.1370 / 3920	0.1376 / 7222
$n_{f,d}$ = 1, ϵ_D = 1, s_d = 10	0.1376 / 3931	0.1377 / 7218
$n_{f,d}$ = 30, ϵ_D = 10, s_d = 1	0.1382 / 3989	0.1382 / 7241
$n_{f,d}$ = 30, ϵ_D = 1, s_d = 1	0.1381 / 3982	0.1381 / 7227

ภาพที่ 4.10 แสดงผลการทดลองด้วยวิธี DFC-P จำนวนทั้งหมด 5 ภาพโดยตั้งค่าพารามิเตอร์ให้เป็น $n_{f,d}$ = 1, ϵ_D = 20, s_d = 1 ตามลำดับ

จากผลการทดลองด้วยวิธีการของ DFC-P การปรับค่า $n_{f,d}$, ϵ_D, และ s_d ส่งผลต่อความถูกต้องในการคำนวณค่า α ไม่แตกต่างกันมากนัก

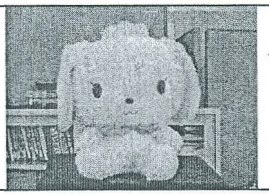
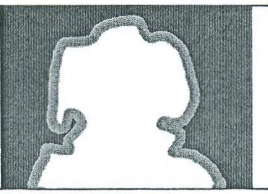
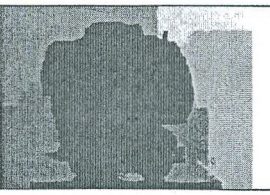
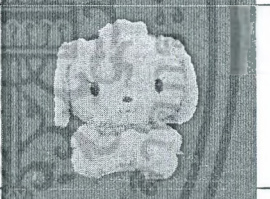
การทดลองวิธีการ DFC-P กับภาพถ่ายด้วยกล้อง Kinect โดยปรับค่าพารามิเตอร์ต่าง ๆ กัน แสดงดังในภาพที่ 4.11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพต้นฉบับ	ภาพไตรแมบ	ภาพแมท	RMSE	SAD
			0.1053	1374
			0. 2160	6337
			0.1138	1326
			0.1819	9023
			0. 0706	1584

ภาพที่ 4.10 ภาพต้นฉบับ ภาพไตรแมบ และภาพแมทที่สร้างขึ้นด้วยวิธี DFC-P และค่า RMSE, SAD ที่ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพต้นฉบับ	แผนภาพไตรแมบ	ภาพระยะลึก
		
ค่าพารามิเตอร์	**ภาพแมท**	**ภาพซ้อน**
$n_{f,d} = 1, \epsilon_D = 1, s_d = 1$		
$n_{f,d} = 1, \epsilon_D = 10, s_d = 1$		
$n_{f,d} = 1, \epsilon_D = 20, s_d = 1$		
$n_{f,d} = 1, \epsilon_D = 1, s_d = 10$		
$n_{f,d} = 30, \epsilon_D = 10, s_d = 1$		
$n_{f,d} = 30, \epsilon_D = 1, s_d = 1$		

ภาพที่ 4.11 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมบ ภาพแมทและภาพซ้อนด้วยวิธีการ DCF-P

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.3 การทดสอบเทคนิค Closed-form Matting โดยใช้ข้อมูลระยะลึกช่วยของ Sun แบบที่ 1 (DCF-S1) กับภาพจากฐานข้อมูลมาตรฐานจำนวน 5 รูปแบบเดียวกับที่ใช้ในการทดสอบ OBAY และอีก 10 รูปเพิ่มเติม อาศัยค่าข้อมูลระยะลึกในการปรับค่าพารามิเตอร์ ϵ_i ซึ่งไม่คงที่ตลอดทั้งภาพ โดยมีพารามิเตอร์ k_1 และ k_2 ควบคุมผลของความแปรปรวนของค่าระยะลึกที่มีผลต่อค่า ϵ_i

ตารางที่ 4.2 ค่าเฉลี่ยของ RMSE และ SAD ที่ค่าพารามิเตอร์ต่าง ๆ โดยใช้วิธี DCF-S1

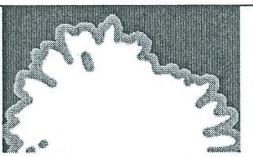
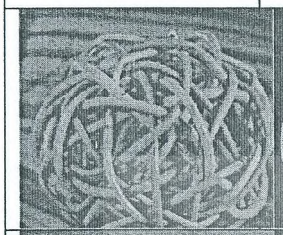
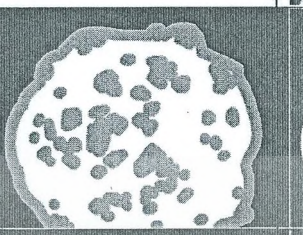
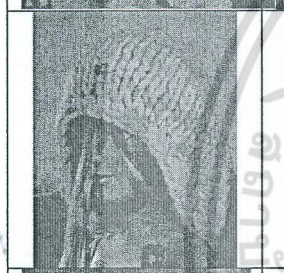
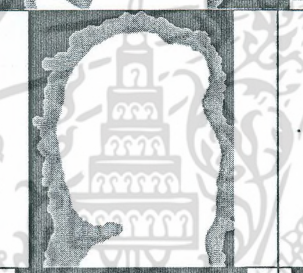
Parameter	RMSE / SAD (5 รูป)	RMSE / SAD (10 รูป)
$n_{f,d}$ = 1, s_d =1, k_1 = 0.1, k_2 = 10^6	0. 2518 / 8697	0. 2916 / 16992
$n_{f,d}$ = 1, s_d =10, k_1 = 0.1, k_2 = 10^6	0.2446 / 8418	0.2857 / 16691
$n_{f,d}$ = 30, s_d =1, k_1 = 0.1, k_2 = 10^6	0. 1872 / 5941	0. 2332 / 13511
$n_{f,d}$ = 30, s_d =10, k_1 = 0.1, k_2 = 10^6	0. 1623 / 4830	0. 1880 / 10065

ภาพที่ 4.12 แสดงผลที่ได้จากการทดลองด้วยวิธี DFC-S1 จำนวนทั้งหมด 5 ภาพโดยตั้งค่าพารามิเตอร์ให้เป็น $n_{f,d}$ = 30, s_d =10, k_1 = 0.1, k_2 = 10^6.

จากผลการทดลองด้วยวิธีการ DCF-S1 พบว่าค่า $n_{f,d}$ และ s_d มีผลต่อภาพแมทที่คำนวณได้โดยค่าที่มากกว่า ส่งผลให้คุณภาพของภาพแมทดีขึ้น (ทั้งโดยการวัดจาก RMSE/SAD และการประเมินจากภาพที่ได้)

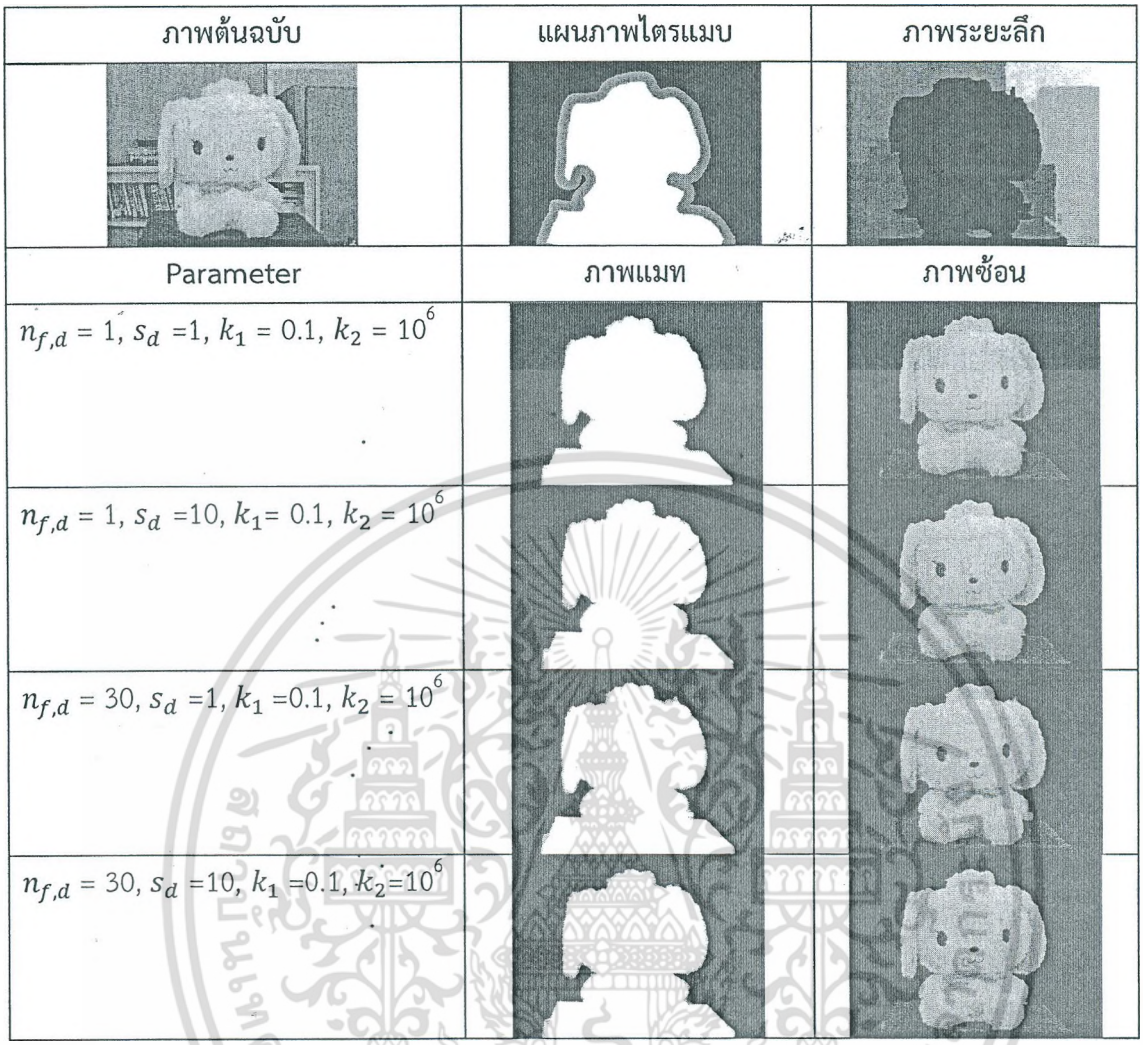
การทดลองวิธีการ DFC-S1 กับภาพถ่ายด้วยกล้อง Kinect โดยปรับค่าพารามิเตอร์ต่าง ๆ กัน แสดงดังในภาพที่ 4.13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพต้นฉบับ	ภาพไตรแมบ	ภาพแมท	RMSE	SAD
			0. 1304	1939
			0. 2230	6811
			0. 1731	2288
			0. 2317	11921
			0. 0533	1192

ภาพที่ 4.12 ภาพต้นฉบับ ภาพไตรแมบ และภาพแมทที่สร้างขึ้นด้วยวิธี DCF-S1 และค่า RMSE, SAD ที่ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 4.13 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมบ ภาพแมทและภาพซ้อนด้วยวิธีการ DCF-S1

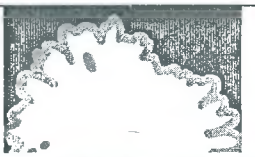
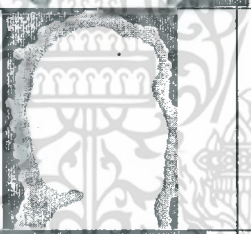
4.2.4. การทดสอบวิธีการของ Sun แบบที่ 2 (DCF-S2) กับภาพจากฐานข้อมูลฯ โดยชุดที่ 1 มีจำนวน 5 รูป และชุดที่ 2 มีจำนวน 10 รูป ผลค่าเฉลี่ยของ RMSE และ SAD แสดงดังตารางที่ 4.3

ตารางที่ 4.3 ค่าเฉลี่ยของ RMSE และ SAD ที่ค่าพารามิเตอร์ต่าง ๆ โดยใช้วิธี DCF-S2

Parameter	RMSE / SAD (5 รูป)	RMSE / SAD (10 รูป)
$n_{f,d}$ = 1, s_d =1, k_1 = 0.1, k_2 = 10^6	0. 1402 / 4075	0. 1444 / 7752
$n_{f,d}$ = 1, s_d =10, k_1 = 0.1, k_2 = 10^6	0. 1383 / 4001	0. 1439 / 7682
$n_{f,d}$ = 30, s_d =1, k_1 = 0.1, k_2 = 10^6	0. 1433 / 4177	0. 1468 / 7806
$n_{f,d}$ = 30, s_d =10, k_1 = 0.1, k_2 = 10^6	0. 1378 / 3895	0. 1444 / 7712

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพต่อไปนี้ แสดงผลที่ได้จากการทดลองด้วยวิธี DFC-S2 จำนวนทั้งหมด 5 ภาพโดยตั้งค่าพารามิเตอร์ให้เป็น $n_{f,d}$ = 30, s_d =10, k_1 = 0.1, k_2 = 10^6

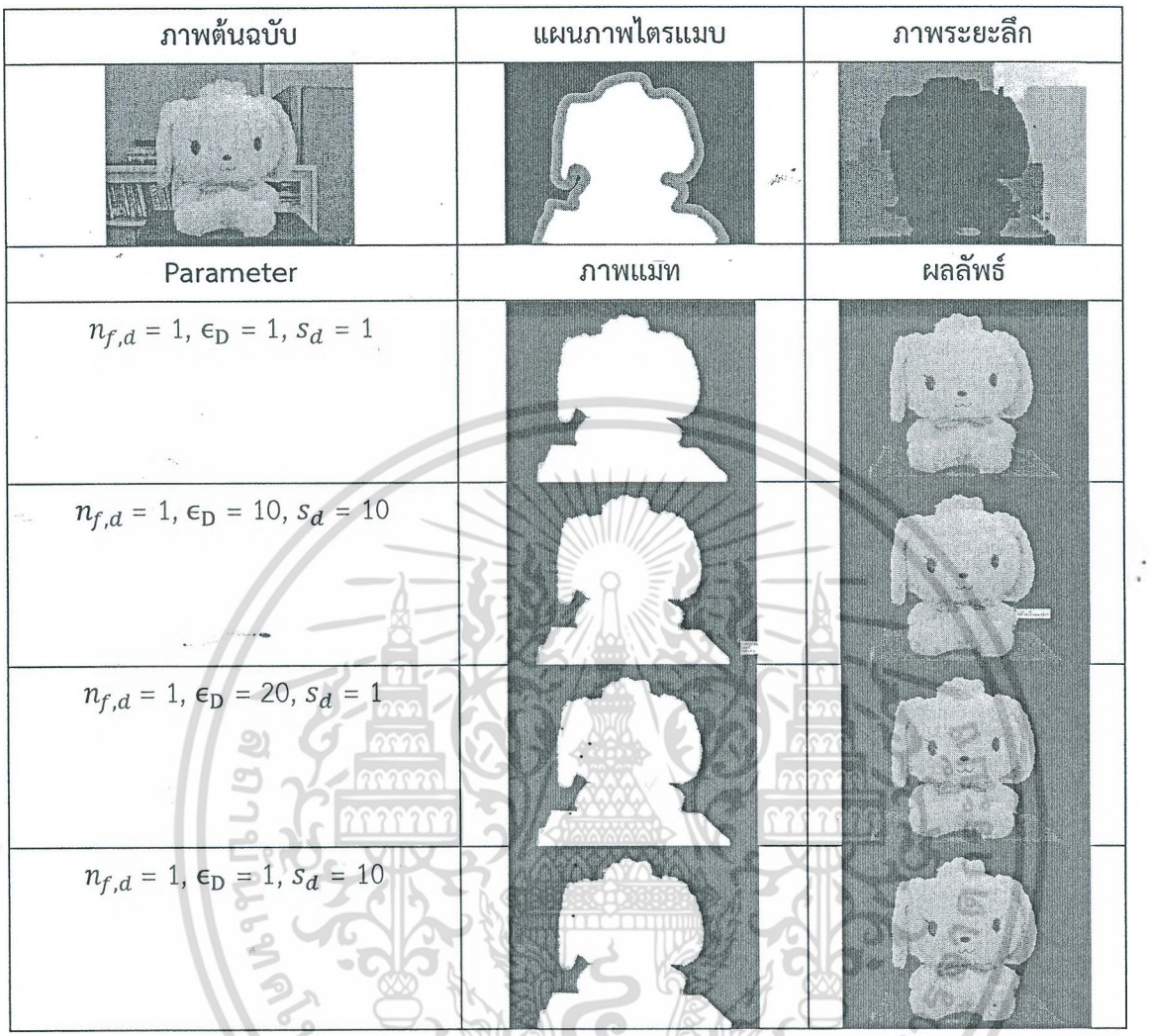
ภาพต้นฉบับ	ภาพไตรแมบ	ภาพแมท	RMSE	SAD
			0. 1086	1461
			0. 2230	6764
			0. 1188	1436
			0. 1781	8884
			0. 0603	1032

ภาพที่ 4.14 ภาพต้นฉบับ ภาพไตรแมบ และภาพแมทที่สร้างขึ้นด้วยวิธี DFC-S2 และค่า RMSE, SAD ที่ได้

จากผลการทดลองปรับค่าพารามิเตอร์ต่าง ๆ ส่งผลต่อประสิทธิภาพในการคำนวณภาพแมทบ้าง แต่มีความแตกต่างกันไม่มาก และให้มีความถูกต้องใกล้เคียงกับ OCF

สำหรับผลการทดลองด้วยวิธีการดังกล่าว กับภาพถ่ายด้วยกล้อง Kinect+DSLR โดยปรับค่าพารามิเตอร์ต่า ๆ แสดงดังภาพที่ 4.15

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



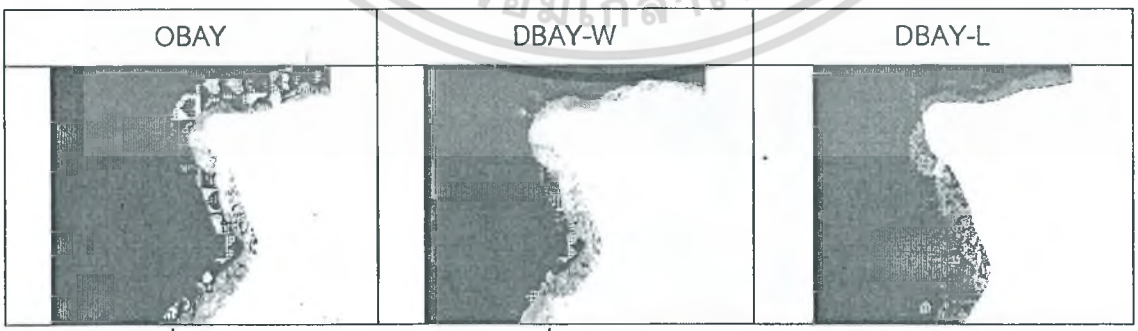
ภาพที่ 4.15 ภาพจากกล้อง Kinect+DSLR ภาพไตรแมบ ภาพแมทและภาพซ้อนด้วยวิธีการ DCF-S2

จากผลการทดลองทั้งหมดตามที่ได้รายงานไปแล้ว สามารถวิเคราะห์เปรียบเทียบเป็นประเด็นต่าง ๆ ได้ดังต่อไปนี้

1. สำหรับกลุ่มภาพจากฐานข้อมูลมาตรฐาน ผลที่ได้กลุ่มของ Bayesian matting methods จะให้ผลบริเวณที่วัตถุมีลักษณะเป็นรูหรือเป็นโพรงภายในบริเวณ foreground region ได้ดีกว่า Closed-form matting methods แต่วิธีหลังให้ค่า SAD ที่ต่ำกว่าเล็กน้อย และโดยการประเมินด้วยการตรวจสอบคุณภาพของภาพแมทบริเวณขอบของวัตถุ จะได้ผลที่ใกล้เคียงภาพจริงมากกว่า เนื่องจากภาพแมทมีแนวโน้มที่จะเปลี่ยนแปลงค่าอย่างต่อเนื่อง อันเป็นผลมาจากสมมติฐานและวิธีการในการคำนวณของเทคนิค Closed-form matting methods

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. วิธีการที่อยู่ในกลุ่มของ Bayesian-based matting techniques พบว่าวิธีการของ Lu (DBAY-L) เมื่อเปรียบเทียบโดยใช้ค่า RMSE ให้ผลที่ด้อยกว่าวิธีการอื่น ๆ (OBAY และ DBAY-W) ในขณะที่วิธี DBAY-W ได้ผลใกล้เคียงกับ OBAY
3. วิธีการที่อยู่ในกลุ่มของ Closed-form matting techniques พบว่าวิธีการที่ใช้ค่าระยะลึกได้ผลใกล้เคียงกับวิธีการเดิม (OCF) สำหรับรูปภาพส่วนใหญ่ โดยการปรับค่าพารามิเตอร์ต่าง ๆ ให้ผลในลักษณะที่สอดคล้องกันทุกวิธี กล่าวคือ การปรับพารามิเตอร์ ที่มีผลต่อการให้น้ำหนักกับค่าระยะลึกมาก จะมีผลทำให้รอยต่อวัตถุกับฉากหลังในส่วนที่เป็นเส้นผมหรือขนขาดหายไป คือจะให้ผลในลักษณะของการทำ hard segmentation ในขณะที่หากปรับพารามิเตอร์เพื่อให้น้ำหนักกับค่าระยะลึกน้อยลง ส่งผลทำให้ผลที่ได้ใกล้เคียงกับวิธีการ OCF มากขึ้น แต่ในขณะเดียวกัน ก็จะทำให้มีส่วนที่คำนวณผิดพลาด (เช่น ส่วนที่ควรจะเป็น foreground ขาดหายไป) เพิ่มมากขึ้นในลักษณะเดียวกับที่พบในผลที่ได้จาก OCF นอกจากนี้ การเพิ่มข้อมูลระยะลึกเข้าไปเป็นลำดับที่ 4 ใน I_i พบว่าส่งผลทำให้ภาพแมทที่ได้มีลักษณะใกล้เคียงกับการทำ hard segmentation กล่าวคือ รายละเอียดในส่วนที่เป็นเส้นผมหรือขนขาดหายไป การกรองภาพระยะลึกก่อนนำไปใช้ในวิธีการ DCF-S2 ที่ใช้ระยะลึกเพื่อการปรับค่าน้ำหนักเพียงอย่างเดียว ส่งผลในลักษณะเดียวกันกับการลดน้ำหนักความสำคัญของข้อมูลระยะลึก ในขณะที่การเพิ่มระดับการกรองภาพสำหรับวิธีการ DCF-P และ DCF-S1 ที่ใช้ข้อมูลระยะลึกเป็นองค์ประกอบที่ 4 เพิ่มเติม พบว่าส่งผลทำให้ขอบรอยต่อของวัตถุมีการไล่ระดับมากขึ้น แต่ก็ยังขาดรายละเอียดในส่วนของเส้นผมหรือขนอยู่
4. การทดสอบกับภาพจริง พบว่าในกรณีของภาพที่ฉากหลังมีลวดลายและสีที่หลากหลาย การใช้ข้อมูลระยะลึกให้ผลที่ดีขึ้น โดยเฉพาะอย่างยิ่งในกรณีของวิธีการที่ใช้ Bayesian methods (ดูภาพที่ 4.16 ประกอบ) แต่สำหรับวิธีการในกลุ่มของ Closed-form methods ผลที่ได้ยังไม่แตกต่างกัน นอกจากนี้ ด้วยเหตุที่ปริมาณข้อมูลที่ใช้ในการทดลองในส่วนดังกล่าวยังน้อยอยู่ ทำให้ยังไม่สามารถสรุปผลได้อย่างมีน้ำหนักเพียงพอ



ภาพที่ 4.16 ภาพเปรียบเทียบภาพแมท ที่คำนวณได้ของ OBAY , DBAY-W และ DBAY-L

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 การทดลองปรับปรุงคุณภาพของภาพระยะลึกเพื่อการสร้างแผนภาพไตรแมบอัตโนมัติ

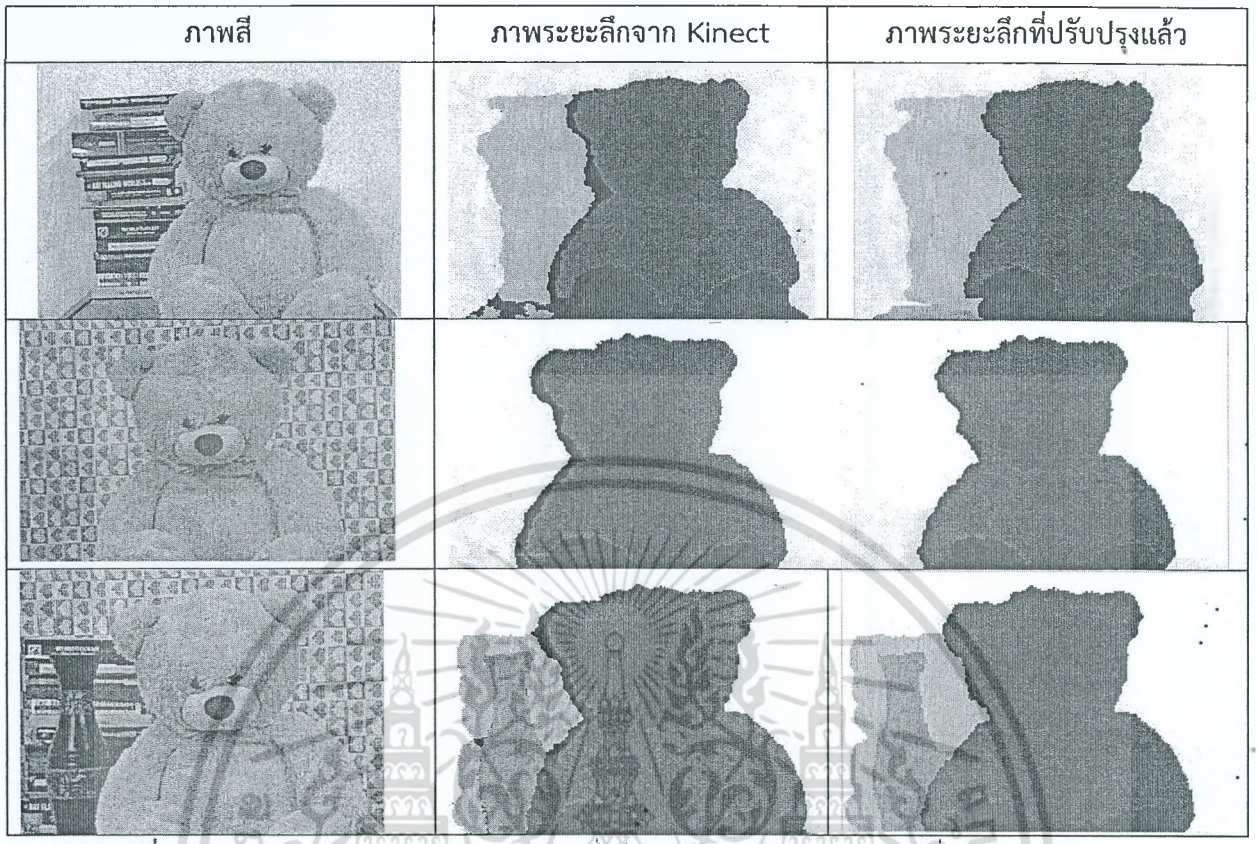
สำหรับการทดลองปรับปรุงคุณภาพของภาพระยะลึกเพื่อการสร้างแผนภาพไตรแมบอัตโนมัติจะสามารถแบ่งออกได้เป็น 2 ส่วนด้วยกันคือ

4.3.1 ส่วนการแก้ไขข้อมูลรูปภาพระยะลึกที่เกิดจากเงาของวัตถุหรือการที่วัตถุไม่สะท้อนแสงโดยที่บริเวณไม่มีข้อมูลระยะลึกเป็นสีดำ จะสามารถตรวจสอบได้จากการทำงานของ Kinect ที่ใช้หลักการ Structure light ในการคำนวณหาระยะลึกตามหัวข้อที่ 3.3.3 เนื่องจากการแก้ไขข้อมูลระยะลึกที่หายไปจากการเฉลี่ยค่าความลึกบริเวณข้างเคียงจะทำให้ข้อมูลระยะลึกที่ได้นั้นผิดพลาดไปจากความเป็นจริง ซึ่งจะส่งผลต่อการกำหนดขอบเขตของไตรแมบที่เหมาะสมเพราะบริเวณค่าระยะลึกที่แก้ไขนั้นกลายเป็นบริเวณที่ไม่แน่ชัดว่าเป็นค่าระยะลึกของวัตถุหรือพื้นหลัง

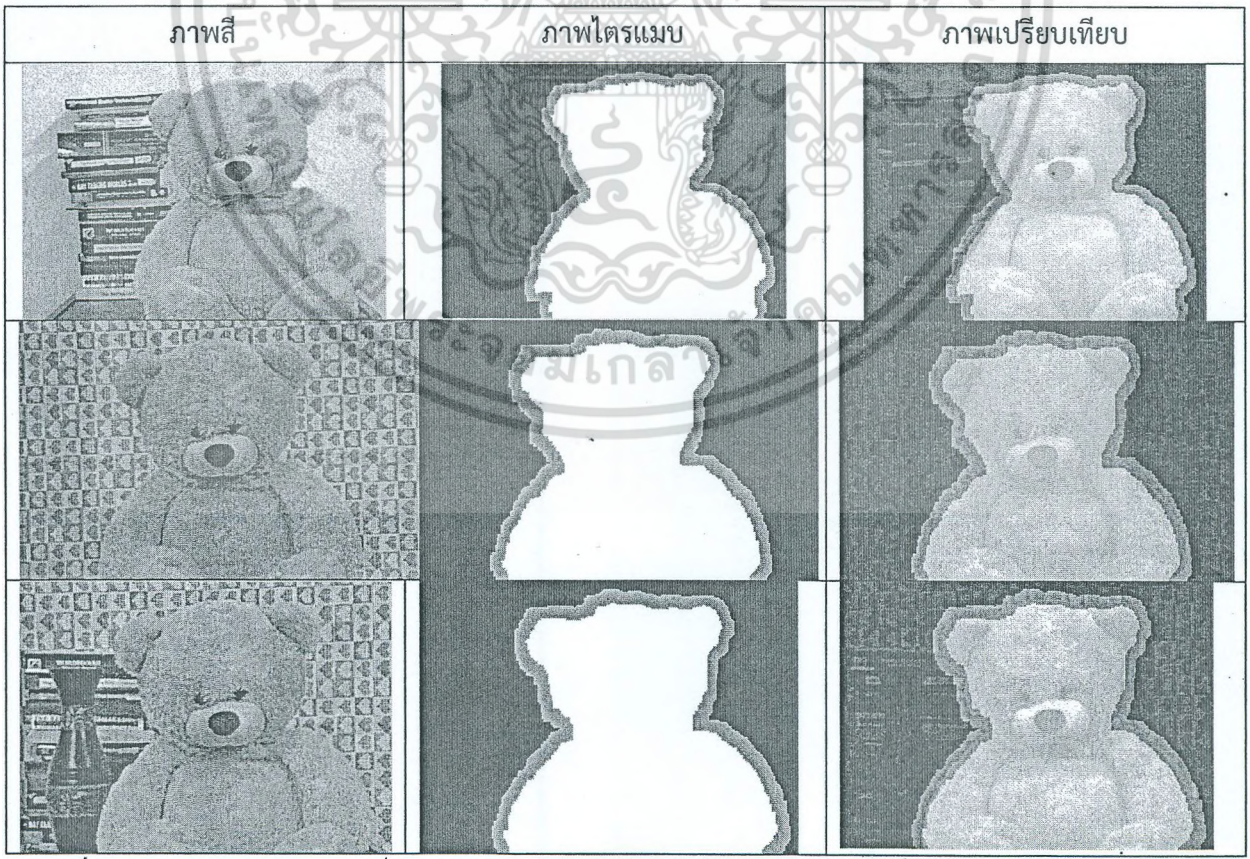
จากผลการทดลองการปรับปรุงคุณภาพของภาพระยะลึกจะสามารถแก้ไขบริเวณที่ไม่ทราบค่าที่เกิดจากการบังแสงของวัตถุโดยอาศัยหลักการคำนวณหาค่าระยะลึกของ Kinect โดยในการตรวจสอบความถูกต้องเบื้องต้นจะตรวจสอบจากผู้ใช้เป็นคนตัดสินผลลัพธ์ที่ได้ โดยดูว่าบริเวณที่ไม่ทราบข้อมูลระยะลึกที่ควรจะเป็นเงาของวัตถุ สามารถตรวจสอบถูกต้องและได้รับการแก้ไขจากข้อมูลระยะลึกของพื้นหลังข้างเคียง จากผลการทดลองกับภาพจริงจำนวน 3 ภาพจะสามารถตรวจสอบบริเวณที่เกิดจากเงาของวัตถุและแก้ไขปรับปรุงข้อมูลความลึกของภาพได้

4.3.2 การสร้างไตรแมบแบบอัตโนมัติจากค่าความผิดพลาดหรือความคลาดเคลื่อนของ Kinect ที่ได้จากการวัดในแต่ละระยะ เนื่องจากข้อมูลภาพระยะลึกของ Kinect จะมีความผิดพลาดเพิ่มสูงขึ้นเมื่อวัตถุมีระยะที่ห่างจากตัวกล้องเพิ่มมากขึ้นทำให้การกำหนดขอบเขตของไตรแมบในแต่ละระยะมีความแตกต่างกันออกไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 4.17 เปรียบเทียบ ภาพสี ภาพระยะลึกที่ได้จากกล้อง และภาพระยะลึกที่ปรับปรุงแล้ว



ภาพที่ 4.18 ภาพสี ภาพไตรแมบที่สร้างจากข้อมูลภาพระยะลึก และภาพระยะลึกเมื่อซ้อนทับกับภาพสี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากภาพผลการทดลองในการสร้างไตรแมบโดยการใช้ข้อมูลความคลาดเคลื่อนของการตัวอุปกรณ์ Kinect ที่เกิดจากการวัดมาใช้ในการกำหนดขอบเขตของ unknown region ให้เหมาะสมครอบคลุมภาพวัตถุและบริเวณข้อมูลภาพระยะลึกที่ผิดพลาดคลาดเคลื่อนจากการซ้อนกันกับภาพสี

สำหรับการตรวจสอบและประเมินผลการสร้างขอบเขตของ unknown region ทำโดยการเปรียบเทียบระหว่างภาพไตรแมบที่กำหนดด้วยมือกับภาพไตรแมบที่สร้างขึ้นจากข้อมูลภาพระยะลึก โดยที่ภาพไตรแมบที่สร้างด้วยมือ กำหนดให้แทนภาพไตรแมบที่มีความถูกต้อง (ground truth) ในการประเมินอาศัยการวัดค่าต่าง ๆ ดังต่อไปนี้

1. ค่าประสิทธิผล (effectiveness factor) คือ ค่าที่บ่งบอกว่าผลลัพธ์ตรงตามที่คาดหวังหรือไม่ โดยในที่นี้หมายถึงบริเวณที่ unknown region ของภาพไตรแมบที่สร้างจากภาพระยะลึกนั้น ครอบคลุมบริเวณ unknown region ตามที่ปรากฏในภาพไตรแมบที่กำหนดด้วยมือ หรือไม่มากน้อยเพียงไร โดยค่าประสิทธิผลจะมีค่าอยู่ช่วง 0-1 โดยที่ 1 คือผลลัพธ์ที่ดีที่สุดซึ่งจะหมายความว่าบริเวณของ unknown region ที่คาดหวังนั้นถูกครอบคลุมทั้งหมด สำหรับค่าประสิทธิผลคำนวณได้จากสมการต่อไปนี้

$$f = \frac{B - A}{B} \tag{4.1}$$

โดยที่

f คือค่าประสิทธิผล

B คือบริเวณของ unknown ที่ได้จากการกำหนดด้วยมือ

A คือบริเวณของ unknown region ที่ถูกต้อง ที่ครอบคลุมโดยแผนภาพไตรแมบที่สร้างจากภาพระยะลึก (ดูภาพที่ 4.19)

ภาพที่ 4.19 ตัวอย่างบริเวณ unknown region ในภาพไตรแมบที่กำหนดด้วยมือ ซึ่งไม่ครอบคลุมโดยภาพไตรแมบที่สร้างจากภาพระยะลึก

2. ค่าประสิทธิภาพ (efficiency factor) คือ ค่าที่ใช้ในบ่งบอกถึงประสิทธิภาพการทำงานหรือผลลัพธ์ที่ได้ โดยในที่นี้จะหมายถึง ค่าที่ใช้ในการบ่งบอกว่าขอบเขตของ unknown region ที่สร้างจากภาพระยะลึก มีประสิทธิภาพมากน้อยเท่าไร โดยค่าประสิทธิภาพจะมีค่าอยู่ในช่วง 0-1 ซึ่งหากค่าประสิทธิภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มีค่าคือ 1 แสดงว่าการกำหนดขอบเขตของ unknown region ด้วยมือ กับที่ได้จากการคำนวณเท่ากันพอดี ค่าประสิทธิภาพสามารถหาได้จากสมการต่อไปนี้

$$g = \frac{B - A}{C} \tag{4.2}$$

กำหนดให้

g คือค่าประสิทธิภาพ

B คือบริเวณของ unknown region ที่แท้จริงที่ได้จากการกำหนดด้วยมือ

A คือบริเวณของ unknown region ที่ถูกต้อง ที่ครอบคลุมโดยแผนภาพไตรแมบที่สร้างจากภาพระยะลึก

C คือบริเวณของunknown region ที่ได้จากการคำนวณโดยใช้ข้อมูลระยะลึก

ผลการทดลองเปรียบเทียบการสร้างไตรแมบด้วยมือกับการสร้างไตรแมบจากภาพระยะลึกตามวิธีที่อธิบายในงานวิจัยนี้ แสดงในภาพต่อไปนี้

ภาพไตรแมบจากการวาด	ภาพไตรแมบจากการวัด	$C-(B-A)$	A	$B-A$	ประสิทธิผล/ประสิทธิภาพ
					0.824 /0.711
					0.718 /0.794
					0.801/ 0.888

ภาพที่ 4.20 แสดงผลการสร้างภาพไตรแมบจากภาพระยะลึก เปรียบเทียบกับภาพที่กำหนดโดยใช้การวาดด้วยมือ

จากผลการทดลองเปรียบการสร้างไตรแมบด้วยมือกับการสร้างไตรแมบจากค่าความคลาดเคลื่อนของตัวอุปกรณ์ จากภาพถ่ายทั้ง 3 ภาพจะได้ค่าเฉลี่ย ประสิทธิผลที่ 0.781 และประสิทธิ์ภาพคือ 0.797

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลการดำเนินงานและข้อเสนอแนะ

5.1 สรุปผลการดำเนินงาน

ในงานวิจัยนี้ ได้ทำการศึกษาและพัฒนาเทคนิคการซ้อนภาพฉากหลังธรรมชาติ โดยมุ่งศึกษาแนวทางการนำเอาข้อมูลระยะลึกที่ได้จากกล้องถ่ายภาพประเภท RGB-D camera มาใช้ให้เป็นประโยชน์ ซึ่งในงานวิจัยนี้ ได้แบ่งแนวทางการใช้ประโยชน์จากข้อมูลดังกล่าว ออกเป็น 2 ลักษณะ ได้แก่ การใช้ข้อมูลระยะลึกเป็นส่วนหนึ่งในกระบวนการคำนวณภาพแมท และการใช้ข้อมูลระยะลึกในการคำนวณหรือสร้างภาพไตรแมบ โดยในส่วนแรกงานวิจัยนี้ได้ทำการศึกษาเปรียบเทียบวิธีการต่าง ๆ ทั้งมีใช้ค่าระยะลึกและไม่ได้ใช้ค่าระยะลึกที่แตกต่างกัน รวม 7 วิธีด้วยกัน โดยทำการทดสอบกับภาพจากฐานข้อมูลภาพที่เผยแพร่เพื่อใช้ในการเปรียบเทียบประสิทธิภาพของวิธีการสร้างภาพแมทต่าง ๆ และโดยการถ่ายภาพจริงด้วยกล้อง Kinect+DSLR ผลจากการศึกษามีข้อสรุปดังนี้

1. วิธีการที่ใช้ข้อมูลระลึกทุกวิธีที่ทดสอบไม่ได้ให้ผลที่ดีขึ้นอย่างมีนัยสำคัญ เมื่อเปรียบเทียบกับวิธีการเดียวกันที่ไม่ได้อาศัยข้อมูลระยะลึก ยกเว้นบางภาพที่ถ่ายจากกล้อง Kinect-DSLR ที่มีฉากหลังที่มีลวดลายและหลากสีสัน ที่พบว่าการใช้ข้อมูลระยะลึกกับวิธีการ Bayesian-matting method ให้ผลที่ดีขึ้นเมื่อเทียบกับวิธีการเดียวกันที่ไม่ได้ใช้ข้อมูลดังกล่าวประกอบในการคำนวณ ทำให้มีความเป็นไปได้ว่าในกรณีของวัตถุหรือฉากหลังที่มีลวดลายที่ซับซ้อน การใช้ภาพระยะลึกอาจจะส่งผลต่อคุณภาพของการคำนวณในทางที่ดีขึ้นได้ ทั้งนี้ จำเป็นต้องการทำการทดลองเพิ่มเติมกับภาพในลักษณะดังกล่าว เพื่อให้ได้ข้อสรุปที่เชื่อถือได้ต่อไป

2. จากผลการทดลองโดยรวม พบว่าผลที่ได้ขึ้นอยู่กับลักษณะของภาพเป็นสำคัญ ด้วยเหตุที่ภาพส่วนใหญ่ที่ใช้ในการทดลอง มีลวดลายค่อนข้างน้อย ทั้งวัตถุด้านหน้าและฉากหลัง ทำให้เทคนิค Closed-form method ได้ผลที่ค่อนข้างดี นอกจากนี้ เมื่อนำค่า RMSE มาเปรียบเทียบกับผลการประเมินด้วยสายตาจากภาพแมทที่ได้ บางส่วนมีความไม่สอดคล้องกัน จึงมีความเป็นไปได้ที่การใช้มาตรวัดด้วย RMSE อาจไม่ได้สะท้อนคุณภาพของภาพได้ดีพอ ดังเช่นที่ปรากฏเป็นข้อสังเกตใน [9]

3. ผลการทดลองปรับปรุงคุณภาพของภาพระยะลึกเพื่อการสร้างแผนภาพไตรแมบอัตโนมัติ สามารถแก้ไขบริเวณที่ไม่ได้เกิดจากเงาของวัตถุและสร้างไตรแมบที่จากการวัดค่าความคลาดเคลื่อนของ Kinect ให้มี unknown region ครอบคลุมภาพวัตถุได้ถูกต้อง โดยในงานวิจัยนี้ได้พัฒนาวิธีการปรับปรุงภาพระยะลึก และการกำหนดความกว้างของ unknown band นอกจากนี้ ในรายงานฉบับนี้ได้นำเสนอการประเมินในแง่ประสิทธิภาพและประสิทธิผล ซึ่งเป็นรูปแบบการประเมินที่พัฒนาขึ้นในงานวิจัยนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ข้อเสนอแนะ

แม้ว่าจะผลการทดลองบางส่วน จะบ่งบอกว่าการใช้ข้อมูลระยะลึกอาจเป็นประโยชน์ต่อการสร้างภาพแมทในกรณีที่วัตถุหรือฉากหลังมีลวดลายและสีสันที่หลากหลาย แต่ด้วยจำนวนภาพในการทดลองที่ยังน้อยอยู่ ทำให้มีความจำเป็นในการทดสอบเพิ่มเติมกับข้อมูลภาพที่มากและหลากหลายกว่านี้ นอกจากนี้ ด้วยเหตุที่ว่าข้อมูลระยะลึกไม่ได้มีความสัมพันธ์โดยตรงกับค่าแมท ทำให้การประยุกต์ใช้ข้อมูลดังกล่าวโดยตรงในการคำนวณค่าแมทขาดการสนับสนุนทางทฤษฎีที่เพียงพอ ประกอบกับผลจากการทดลองที่ได้จากการวิจัยครั้งนี้ ทำให้แนวทางการใช้ข้อมูลระยะลึกเพื่อการซ้อนภาพฉากหลังธรรมชาติ ที่ควรพิจารณาในการวิจัยในโอกาสต่อไป คือการปรับปรุงการใช้ข้อมูลดังกล่าว เพื่อการสร้างแผนภาพไตรแมบที่มีความถูกต้องและมีประสิทธิภาพมากขึ้น อันจะส่งผลต่อความถูกต้องในการสร้างภาพแมทตามไปด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] A. R. Smith and J. F. Blinn, "Blue screens matting", *SIGGRAPH 96*, pp. 259-268, 1996.

[2] Y. Chuang et. al., "A Bayesian approach to digital matting", *IEEE Int. Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 264-271, 2001.

[3] O. Wang et. al., "Automatic Natural Video Matting with Depth", *Pacific Conf. Computer Graphics and Applications*, pp. 469-472, 2007.

[4] A. Levin et. al., "A Closed-Form Solution to Natural Image Matting", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 30, issue 2, pp. 1-15, 2008.

[5] F. Pitie and A. Kokaram, "Matting with a Depth Map", *IEEE Int. Conf. Image Processing*, pp. 21 – 24, 2010.

[6] W. Sun et al, "Adaptive Depth Map Assisted Matting in 3D Video", *IEEE Int. Conf. Multimedia and Expo*, pp. 1-6, 2011.

[7] D. Herrera et. al., "Joint depth and color camera calibration with distortion correction", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 34, issue 10, pp. 2058-2064, 2012.

[8] T. Lu, and S. Li, "Image Matting with Color and Depth Information" *Int. Conf. Pattern Recognition*, pp. 3787 – 3790, 2012.

[9] C. Rhemann et. al., "A Perceptually Motivated Online Benchmark for Image Matting", *IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1826 – 1833, 2009.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้