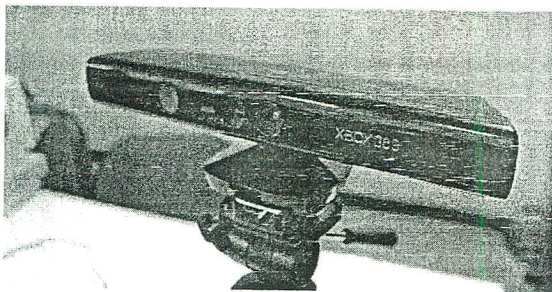


บทที่ 3

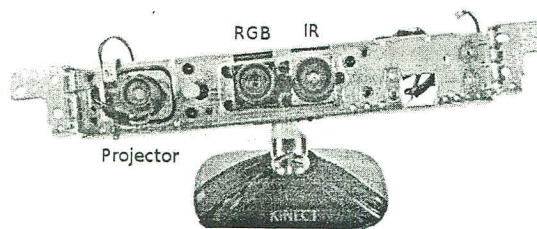
การแยกพื้นหลังธรรมชาติจากภาพหลายมุมมองด้วยอุปกรณ์ Kinect

3.1 การถ่ายภาพหลายมุมมองด้วยอุปกรณ์ Kinect

Kinect คืออุปกรณ์ที่มีความสามารถในการถ่ายภาพหลายมุมมองซึ่งภายในจะประกอบไปด้วย 3 ส่วนหลักคือ กล้องอินฟราเรด (Infra-red, IR) , ส่วนโปรเจคเตอร์ (Projector) และกล้องสี (RGB Camera) ดังภาพที่ 3.1 สำหรับการทำงานจะเริ่มจากการฉายแสงอินฟราเรดออกจากตัว Kinect ซึ่งเป็นแสงที่ไม่สามารถมองเห็นด้วยตาเปล่า แสงที่ฉายจะมีลักษณะเป็นจุดๆตามแนวตั้ง 480 จุด และแนวนอน 640 จุด ซึ่งกล้อง IR จะวัดระดับความลึกโดยการเปรียบเทียบความแตกต่างระหว่างจุดต่างๆในภาพที่ถ่ายด้วยกล้องอินฟราเรดกับภาพที่ฉายด้วย Projector ทำให้สามารถบอกระยะของวัตถุได้โดย และกล้องสีจะเป็นตัวถ่ายภาพแบบปกติ



(a)



(b)

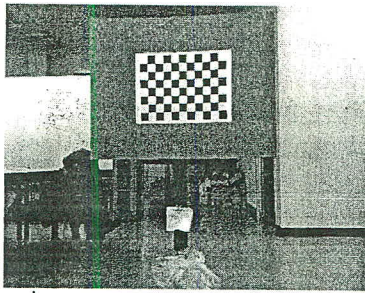
ภาพที่ 3.1 ภาพอุปกรณ์ Kinect

(a) ภาพอุปกรณ์ Kinect, (b) องค์ประกอบภายในอุปกรณ์ Kinect

3.2 การปรับค่าอุปกรณ์ Kinect

อุปกรณ์ Kinect โดยทั่วไปจะมีค่า Focal length และขนาด sensor ที่แตกต่างกัน ซึ่งเกิดจากสายการผลิตหรือความคลาดเคลื่อนในจัดทำ ทำให้ส่งผลต่อประสิทธิภาพการใช้งาน จึงจำเป็นต้องทำการปรับค่าตัวอุปกรณ์ Kinect (Calibration)

การ Calibration วัดเปรียบเทียบเพื่อหาค่าของอุปกรณ์ที่แท้จริง กระทำได้โดยการถ่ายภาพจากตัวอุปกรณ์ Kinect กับภาพ Checkerboard จากมุมมองต่างๆ ดังภาพที่ 3.2 นำภาพที่จากการถ่ายมาหาค่าพารามิเตอร์ต่างๆโดยในงานวิจัยนี้ได้ใช้ toolbox ของ Daniel Herrera C [13] .ในการปรับค่า



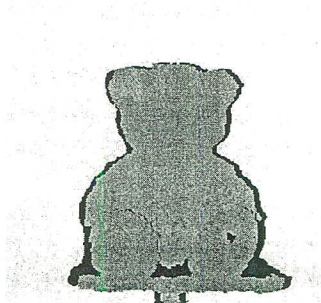
ภาพที่ 3.2 ตัวอย่างภาพ Checkerboard

3.3 การนำข้อมูลความลึกจากอุปกรณ์ Kinect มาใช้ประโยชน์ในการ Matting

3.3.1 การสร้าง Trimap จากข้อมูลความลึก

การสร้างภาพ Trimap โดยปกติต้องมีผู้ใช้งานต้องทำการกำหนดขอบเขตของภาพเอง เพื่อบอกอาณาเขตบริเวณของวัตถุที่อยู่เบื้องหน้า ภาพระยะลึกที่ได้จากอุปกรณ์ Kinect สามารถนำมาใช้ในการกำหนดขอบเขต Trimap ตามกระบวนการดังต่อไปนี้

1. ปกติภาพระยะความลึกที่จากอุปกรณ์ Kinect นั้นจะมีส่วนที่ไม่สามารถวัดค่าได้เช่น . วัตถุนั้นไม่สะท้อนแสงอินฟราเรด หรือวัตถุนั้นบังแสงอินฟราเรดดังภาพที่ 3.3 ที่บริเวณขอบของวัตถุ จะข้อมูลที่อ่านจากตัวอุปกรณ์จะมีค่าระยะลึกจะเป็น 0 หรือที่ทราบ(NaN) ทำให้ต้องมีการทดแทนประมาณค่าระยะลึกที่หายไปยังบริเวณดังกล่าว



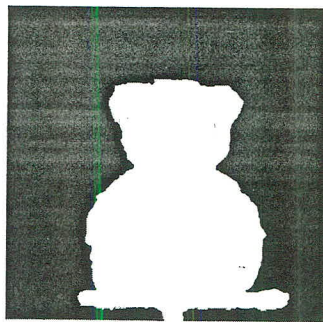
ภาพที่ 3.3 ตัวอย่างภาพระยะลึกที่มีบริเวณไม่ทราบค่า

สำหรับการประมาณค่าระยะลึกจะทำได้โดยการประมาณค่าจากบริเวณรอบๆ จากการให้ค่าน้ำหนักกับบริเวณที่อยู่ใกล้มีค่าน้ำหนักมากกว่าบริเวณที่อยู่ไกล เนื่องจากค่าระยะลึกที่อยู่ใกล้มีความเป็นไปได้สูงที่มีระยะลึกจะเท่ากันหรือใกล้เคียงกัน ซึ่งในงานวิจัยนี้ใช้ฟังก์ชัน Gaussian ในการให้น้ำหนักกับจุดต่างๆที่อยู่รอบจุดที่ไม่ทราบค่า



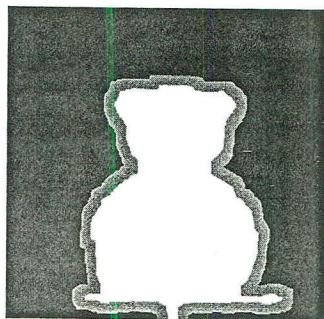
ภาพที่ 3.4 ตัวอย่างภาพระยะลึกที่ผ่านการประมาณค่าจากบริเวณใกล้เคียง

2. ทำการหาขอบของวัตถุโดยใช้ค่าเทรชโวล์(Threshold) ของระยะลึกที่ต่างกันระหว่างฉากหน้ากับพื้นหลังซึ่งในที่นี้ค่าเทรชโวล์ดังกล่าววัดได้จากการใช้อัลกอริทึม Kmean มาช่วยในการหากลุ่มค่าระยะลึกที่มากที่สุดมาบวกกับกลุ่มที่มีค่าระยะลึกน้อยที่สุดแล้วหาร 2 ภาพที่ 3.5 แสดงตัวอย่างของภาพที่ได้จากวิธีการดังกล่าว



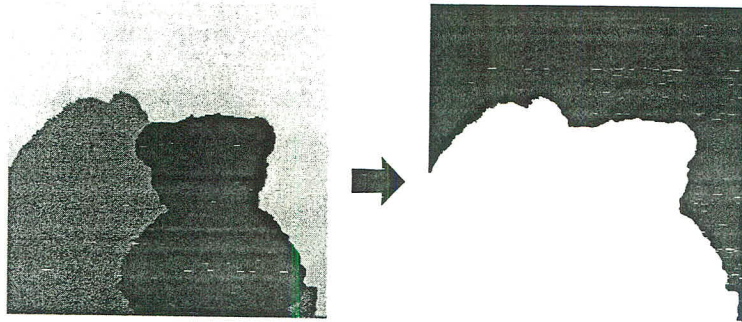
ภาพที่ 3.5 ภาพที่ได้จากการทำ Threshold

3. ทำการปรับเพิ่มความหนาของขอบภาพที่ได้โดยการทำ Dilation เพื่อระบุส่วนของภาพที่เป็น Unknown region และสร้างเป็นแผนภาพ Trimap ดังตัวอย่างแสดงในภาพที่ 3.6



ภาพที่ 3.6 ภาพ Trimap โดยใช้ระยะลึก

การสร้างภาพ Trimap โดยใช้ระยะความลึกมาช่วยนั้น อาจยังไม่ดีเท่าที่ควรหากพื้นหลังนั้นมีหลายระยะจนเกินไปจะทำให้การหาขอบภาพเกิดความผิดพลาดและส่งผลต่อการสร้างภาพ Trimap ดังภาพที่ 3.7 ด้วยเหตุนี้จึงจำเป็นต้องอาศัยผู้ช่วยปรับแต่งขอบเขตบริเวณของวัตถุอีกที



ภาพที่ 3.7 ตัวอย่างการหาขอบที่ผิดพลาดเพราะมีหลายระยะเกินไป

3.3.2 การใช้ข้อมูลระยะลึกมาช่วยเพิ่มประสิทธิภาพการทำ Bayesian matting

Bayesian matting [9] คือ การนำกฎความน่าจะเป็นแบบมีเงื่อนไขของ Baye มาช่วยในการประมาณค่าความเป็นได้สูงสุดที่สุดของค่าสีวัตถุเบื้องหน้า (Foreground), วัตถุเบื้องหลัง (Background) และแอลฟา (Alpha) ที่ควรจะเป็นของบริเวณ Unknown ซึ่งสามารถเขียนเป็นสมการได้ดังนี้

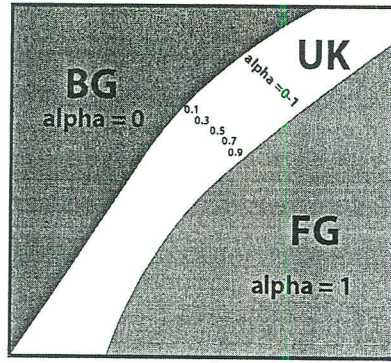
$$\begin{aligned} \arg \max_{F, B, \alpha} P(F, B, \alpha | c) \\ &= \arg \max_{F, B, \alpha} P(C | F, B, \alpha) P(F) P(B) P(\alpha) P(C) \\ &= \arg \max_{F, B, \alpha} L(C | F, B, \alpha) + L(F) + L(B) + L(\alpha) \end{aligned} \quad (3.1)$$

$\arg \max_{F, B, \alpha}$ คือหาค่าของ F , B และ α ที่ให้ค่าผลลัพธ์ความเป็นไปได้สูงสุด โดยที่ F คือ Foreground, B คือ Background และค่า α สำหรับ $P()$ คือค่าความน่าจะเป็นของแต่ละกลุ่ม ซึ่งสามารถเขียนในรูปแบบ log likelihood ได้ดังนี้ $L(_) = \log P(_)$, สำหรับในส่วนของสมการที่ (1) ในเทอมแรกจะมาจากความแตกต่างระหว่างค่าสีที่ได้จากการประมาณ F , B และ α ซึ่งสามารถหาได้จากสมการที่ (2)

$$L(C | F, B, \alpha) = -\|C - \alpha F - (1 - \alpha)B\|^2 / \sigma_c^2 \quad (3.2)$$

สำหรับขั้นตอนการทำ Bayesian matting จะมีหลักการเบื้องต้นดังนี้คือ

1. ทำการประมาณค่า Alpha ให้กับ Unknown จากบริเวณที่ทราบค่าโดยรอบ สำหรับในครั้งแรกจะประมาณจากระยะห่างระหว่าง Foreground และ Background ดังภาพที่ 3.8

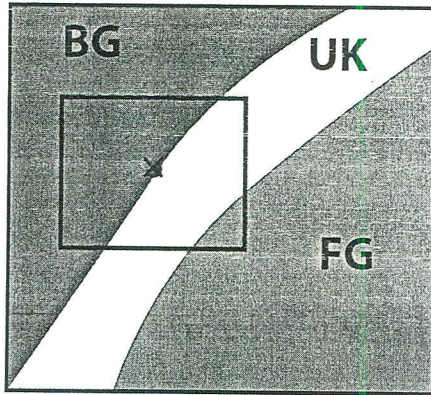


BG = Background
FG = Foreground
UK = Unknown

ทำการประมาณค่า Alpha ให้ Uk โดยดูจาก
ระยะห่างระหว่าง FG กับ BG ซึ่ง Alpha
ในส่วน Uk จะมีค่าระหว่าง 0 ถึง 1

ภาพที่ 3.8 ตัวอย่างการประมาณค่า Alpha

2. ในการคิดคำนวณแต่ละจุดใน Unknown จะใช้การหาค่าที่ทราบอยู่ก่อนจากบริเวณโดยรอบ โดยการสร้าง Window ครอบคลุมไปยังบริเวณโดยรอบเพื่อนำค่ามาคิดคำนวณดังภาพที่ 3.9



BG = Background
FG = Foreground
UK = Unknown

ตำแหน่ง Pixel ที่คิดคำนวณ



Window ที่ไว้ใช้ในการหาค่าจาก
บริเวณที่ทราบโดยรอบตามขนาด
Window



Windows ที่ได้ทำการไปดึงค่าบริเวณ
โดยรอบ

ภาพที่3.9 ตัวอย่างการใช้งาน Window ในการหาค่า

3. ทำการถ่วงน้ำหนักตามความสัมพันธ์ของแต่ละ Pixel ใน Window ให้ค่าถ่วงน้ำหนักของ Alpha ในแต่ละpixel มีค่าเป็น α_i^2 โดยใช้ Gaussian ที่ค่า $\sigma = 8$ สำหรับในส่วน of ค่าสี RGB สมการที่ใช้ในการถ่วงน้ำหนักค่าสีจะเป็น

$$w_i = \alpha_i^2 g_i \quad (3.3)$$

- 4 ทำการแบ่งกลุ่ม (Cluster) สีในส่วน of Window แล้วหาค่าเฉลี่ย, ความแปรปรวนและค่า log likelihoods ของแต่ละกลุ่มจากสมการที่ (3.4), (3.5) และ (3.6)

$$\bar{F} = \frac{1}{T} \sum_{i \in N} w_i F_i \quad (3.4)$$

$$\Sigma_F = \sum_{i \in N} w_i (F_i - \bar{F}) (F_i - \bar{F})^T \quad (3.5)$$

$$L(F) = -(F - \bar{F})^T \Sigma_F^{-1} (F - \bar{F}) / 2 \quad (3.6)$$

โดยที่ F_i จะประกอบค่าสี RGB คือ $F_i = \begin{bmatrix} R_i \\ G_i \\ B_i \end{bmatrix}$ (3.7)

และ

$$T = \sum_{i \in N} w_i \quad (3.8)$$

5. ในส่วนของ Background จะคำนวณเหมือนกันแตกต่างกันเพียงส่วนการคำนวณค่า Alpha โดยจะให้ Alpha มีค่าเท่ากับ $(1 - \alpha_i)^2$
6. ทำการหาค่าสีของ Foreground และ Background ตามสมการที่ (3.9) จากค่า Alpha ที่ได้จากการประมาณก่อนหน้านี้

$$\begin{aligned} & \begin{bmatrix} \Sigma_F^{-1} + I\alpha/\sigma_C^2 & I\alpha(1 - \alpha)/\sigma_C^2 \\ I\alpha(1 - \alpha)/\sigma_C^2 & \Sigma_B^{-1} + I(1 - \alpha^2)/\sigma_C^2 \end{bmatrix} \begin{bmatrix} F \\ B \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_F^{-1}\bar{F} + C\alpha/\sigma_C^2 \\ \Sigma_B^{-1}\bar{B} + C(1 - \alpha)/\sigma_C^2 \end{bmatrix} \end{aligned} \quad (3.9)$$

7. นำค่า F และ B มาทำการหา Alpha ใหม่ ตามสมการที่ (3.10)

$$\alpha = \frac{(C-B) \cdot (F-B)}{\|F - B\|^2} \quad (3.10)$$

8. นำค่าที่ได้ของแต่ละกลุ่มทำการหาค่าความเป็นไปได้สูงสุดในสมการที่ (3.1)
9. วนรอบทำซ้ำเพื่อให้ได้ผลลัพธ์ที่ดีที่สุด

สำหรับการใช้ข้อมูลระยะลึกมาช่วยเพิ่มประสิทธิภาพการทำ Bayesian matting ซึ่งอาศัยหลักการเดียวกับ [14] สามารถทำได้ดังนี้

1. เพิ่มค่าระยะลึกเข้าไปในส่วน Channel สี ที่เดิมมีค่าคือ RGB เป็น RGBD เขียนเป็นสมการได้ดังนี้

$$F_{d,i} = \begin{bmatrix} R_i \\ G_i \\ B_i \\ D_i \end{bmatrix} \quad (3.11)$$

2. จากสมการที่ (3.3) ที่มีการถ่วงค่าน้ำหนักของค่าสี w_i เมื่อมีการนำค่าระยะลึกมาคิดจะสามารถคิดค่าน้ำหนักของค่าระยะลึก ได้จากสมการที่ (3.12)

$$w_{d,i} = |2\alpha_i - 1| \quad (3.12)$$

3. จากสมการที่ (3.4) เมื่อทำการเพิ่มค่าในส่วนของระยะลึกจะเขียนสมการได้ดังนี้

$$\bar{F}_d = W_T \sum_{i \in N} W_i F_{d,i} \quad (3.13)$$

โดยที่ W_T และ W_i จะมีค่าดังต่อไปนี้

$$W_i = \begin{bmatrix} w_i & 0 & 0 & 0 \\ 0 & w_i & 0 & 0 \\ 0 & 0 & w_i & 0 \\ 0 & 0 & 0 & w_{d,i} \end{bmatrix} \quad (3.14)$$

$$W_T = \begin{bmatrix} \frac{1}{T} & 0 & 0 & 0 \\ 0 & \frac{1}{T} & 0 & 0 \\ 0 & 0 & \frac{1}{T} & 0 \\ 0 & 0 & 0 & \frac{1}{T_d} \end{bmatrix} \quad (3.15)$$

โดยที่ T_d มีค่าเท่ากับ

$$T_d = \sum_{i \in N} w_{d,i} \quad (3.16)$$

4. จากสมการที่ (3.5) เมื่อทำการเพิ่มค่าระยะลึกจะเขียนสมการได้ดังนี้

$$\Sigma_{F_d} = \left(\sum_{i \in N} W_i^{\frac{1}{2}} (F_{d,i} - \bar{F}_d) (F_{d,i} - \bar{F}_d)^T W_i^{\frac{1}{2}} \right) \otimes U \quad (3.17)$$

โดยที่

$$W_i^{\frac{1}{2}} = \begin{bmatrix} \sqrt{w_i} & 0 & 0 & 0 \\ 0 & \sqrt{w_i} & 0 & 0 \\ 0 & 0 & \sqrt{w_i} & 0 \\ 0 & 0 & 0 & \sqrt{w_{d,i}} \end{bmatrix} \quad (3.18)$$

$$U = \begin{bmatrix} u1 & u1 & u1 & u2 \\ u1 & u1 & u1 & u2 \\ u1 & u1 & u1 & u2 \\ u2 & u2 & u2 & u3 \end{bmatrix} \quad (3.19)$$

$$u1 = \frac{1}{\sum_{i \in N} w_i} \quad (3.20)$$

$$u2 = \sqrt{w_i} \sqrt{w_{d,i}} \quad (3.21)$$

และ

$$u3 = \frac{1}{\sum_{i \in N} w_{d,i}} \quad (3.22)$$

จากสมการ (3.17) เครื่องหมาย $\otimes$ แทน Element-wise Matrix Product

5. ในส่วนสมการที่ (3.6) , (3.7) และ (3.8) เมื่อนำค่าระยะลิคมาใช้ในการคำนวณจะสามารถเขียนได้ดังนี้

$$L(F_d) = -(F_d - \bar{F}_d)^T \Sigma_{F_d}^{-1} (F_d - \bar{F}_d) / 2 \quad (3.23)$$

$$\begin{bmatrix} \Sigma_{F_d}^{-1} + I\alpha^2/\sigma_C^2 & I\alpha(1-\alpha)/\sigma_C^2 \\ I\alpha(1-\alpha)/\sigma_C^2 & \Sigma_{B_d}^{-1} + I(1-\alpha^2)/\sigma_C^2 \end{bmatrix} \begin{bmatrix} F_d \\ B_d \end{bmatrix} = \begin{bmatrix} \Sigma_{F_d}^{-1} \bar{F}_d + C_d \alpha / \sigma_C^2 \\ \Sigma_{B_d}^{-1} \bar{B}_d + C_d (1-\alpha) / \sigma_C^2 \end{bmatrix} \quad (3.24)$$

$$\alpha = \frac{(C_d - B_d) \cdot (F_d - B_d)}{\|F_d - B_d\|^2} \quad (3.25)$$

โดยที่ค่าของ σ_C กับ σ_d ไม่ได้เป็นค่าตัวเดียวกัน

6. นำค่าที่ได้ของแต่ละกลุ่มทำการหาค่าความเป็นไปได้สูงสุดในสมการที่ (3.1) ซึ่งเมื่อนำค่าระยะลิคมาใช้จะสามารถเขียนสมการได้ว่า

$$\begin{aligned} \arg \max_{F, B, \alpha} P(F_d, B_d, \alpha | C_d) \\ &= \arg \max_{F, B, \alpha} P(C_d | F_d, B_d, \alpha) P(F_d) P(B_d) P(\alpha) P(C) \\ &= \arg \max_{F, B, \alpha} L(C_d | F_d, B_d, \alpha) + L(F_d) + L(B_d) + L(\alpha) \end{aligned} \quad (3.26)$$

7. วนรอบทำซ้ำเพื่อให้ได้ผลลัพธ์ที่ดีที่สุด