

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

2.1. Maximally Stable Extremal Regions (MSER)

การตรวจจับพื้นที่ที่มีการเปลี่ยนแปลงของ ค่าความเข้มสี (Intensity) ที่คงที่และอยู่ในเกณฑ์ที่กำหนดด้วยวิธีการ MSER ซึ่งเป็นส่วนหนึ่งในงานวิจัยของ Matas และคณะ [6] โดยวิธี MSER นั้นสามารถนำใช้งานกับงานวิจัยนี้ได้ เนื่องจากพื้นที่ตัวอักษรแต่ละตัวนั้นมีค่าความเข้มสีที่คงที่ที่ต่างกับพื้นที่รอบๆตัวเองอย่างชัดเจนคงที่ และยังสามารถทนทานต่อมุมมองที่แตกต่างกัน ขนาดที่แตกต่างกัน และการหมุน โดยสามารถอธิบายนิยามการทำงานได้ดังนี้

รูปภาพ (Image, I) ให้ I และ D เป็นฟังก์ชันความสัมพันธ์ โดยที่ $D \subset Z^2 \rightarrow S$ และ S จะเป็นพื้นที่ Extremal ก็ต่อเมื่อ

- S ถูกเรียงลำดับได้ (มีคุณสมบัติ Antisymmetric และ Transitive binary relations)
- ให้ A มีความสัมพันธ์ที่ไม่มีคุณสมบัติสะท้อนและมีคุณสมบัติสมมาตรและ $A \subset D \times D$

พื้นที่ (Region, Q) คือพื้นที่ที่ต่อเนื่องกันโดยเป็นสับเซตของ D โดยแต่ละจุด p และ q เป็นสมาชิกใน Q ซึ่งมีลำดับสมาชิกเป็น $p, a_1, a_2, \dots, a_n, q$ และ $pAa_1, a_1Aa_{i+1}, \dots, a_nAq$ หรืออาจบอกได้ว่าทุกจุดหรือพิกเซลนั้นติดกัน

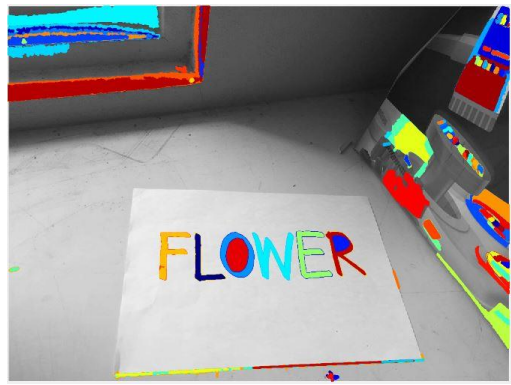
พื้นที่ขอบเขตรอบนอก ((Outer)Region Boundary, ∂Q) สามารถอธิบายนิยามได้คือ $\partial Q = \{q \in D \setminus Q : \exists p \in Q : qAp\}$ กล่าวคือขอบนอกที่เป็นพิกเซลที่ติดกับพิกเซลภายในของพื้นที่ Q

พื้นที่เอ็กซ์ตรีมอล (Extremal Region, $Q \subset D$) เป็นพื้นที่ที่มีค่าความเข้มสูงที่สุดในพื้นที่ และทุก ๆ พิกเซลของ p ที่เป็นสมาชิกของ Q และ q ที่เป็นสมาชิกของ ∂Q โดยที่ $I(p) > I(q)$ หรือพื้นที่ที่มีค่าความเข้มต่ำที่สุดในพื้นที่โดยที่ $I(p) < I(q)$ ซึ่งพื้นที่ที่สนใจคือ $I(p)$

พื้นที่เอ็กซ์ตรีมอลที่คงที่มากที่สุด (Maximally Stable Extremal Region) กำหนดให้ $Q_1, \dots, Q_{i-1}, Q_i, \dots$ เป็นลำดับของพื้นที่เอ็กซ์ตรีมอลที่ซ้อนทับโดยที่ $Q_i \subset Q_{i-1}$ โดยที่พื้นที่ Q_i^* เป็นพื้นที่เอ็กซ์ตรีมอลที่คงที่มากที่สุดก็ต่อเมื่อ $q(i) = |Q_{i+\Delta} \setminus Q_{i-\Delta}| / |Q_i|$ มีค่าต่ำสุดสัมพัทธ์ที่ i^* (เมื่อ $|\cdot|$ คือจำนวนสมาชิกของเซตนั้นๆ) และ $\Delta \in S$ คือพารามิเตอร์ที่ใช้ในการเปลี่ยนแปลงค่า



(ก)



(ข)



(ค)



(ง)



(จ)



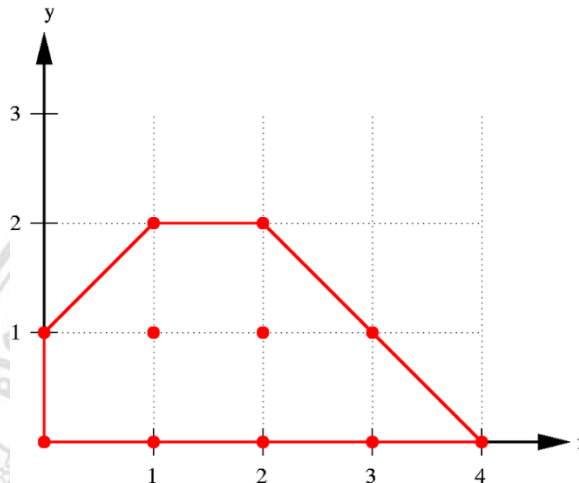
(ฉ)

รูปที่ 2.1 ตัวอย่างการสกัด MSER (ก) เป็นรูปตั้งต้น และ(ข) ถึง (ฉ) รูปตัวอย่างที่ใช้ค่าเกณฑ์ความเปลี่ยนแปลงที่ต่างกันจากน้อยไปมาก

รูปที่ 2.1 (ก) เป็นรูปก่อนการตรวจจับพื้นที่ด้วย MSER ซึ่งตัวอย่างได้ทำการแสดงผลลัพธ์จากการเปลี่ยนค่าเกณฑ์ความเปลี่ยนแปลง โดยพื้นที่ที่ตรวจจับได้จะดูเททับด้วยสีอื่น ๆ นอกจากสีโทนขาว-ดำ ซึ่งถ้าค่าเกณฑ์ความเปลี่ยนแปลงมีค่าน้อยจะสามารถตรวจจับพื้นที่ได้มากเนื่องจากพื้นที่ที่มีการเปลี่ยนแปลงเพียงเล็กน้อยตามเกณฑ์จะถูกตรวจจับและเมื่อเพิ่มค่าเกณฑ์ความเปลี่ยนแปลงขึ้นจึงตรวจจับได้น้อยลงตามลำดับ

2.2. คอนเวกซ์ฮัลล์ (Convex Hull) [19]

คอนเวกซ์ฮัลล์ คือเซตของจุดรอบรูปเหลี่ยมใดๆที่มีมุมภายในทุกมุมมีขนาดไม่เกิน 180 องศา และการลากเส้นตรงระหว่างจุดทุกคู่จะอยู่ภายในพื้นที่รูปเหลี่ยมใดๆ โดยคอนเวกซ์ฮัลล์ที่มีขนาดเล็กที่สุดคือล้อมรอบจุดๆหนึ่ง ซึ่งจะมีเซตของจุดที่ล้อมรอบอยู่อย่างน้อยสามจุด ดังตัวอย่างในรูปที่ 9



รูปที่ 2.2 ตัวอย่างคอนเวกซ์ฮัลล์

สามารถคำนวณพื้นที่ของคอนเวกซ์ฮัลล์ได้ดังสมการที่ 1

$$\begin{aligned} \text{พื้นที่รูปเหลี่ยมใด ๆ} = & x_1 \times y_2 + x_2 \times y_3 + x_3 \times y_4 + \dots + x_{n-1} \times y_n + x_n \times y_1 \\ & - x_2 \times y_1 - x_3 \times y_2 - x_4 \times y_3 - \dots - x_n \times y_{n-1} - x_1 \times y_n \end{aligned} \quad (1)$$

2.3. การลดรูปหลายเหลี่ยมให้อยู่ในรูปอย่างง่าย (Polygonal Simplification) [20]

การลดรูปหลายเหลี่ยมให้เป็นรูปอย่างง่ายเป็นการลดเหลี่ยมของรูปหลายเหลี่ยมลงโดยเรียงลำดับตามความสำคัญของแต่ละจุด ซึ่งความสำคัญของจุดแต่ละจุดนั้นจะขึ้นอยู่กับผลคูณระหว่างมุม (Angle between vector, θ) กับผลคูณความยาว (Multiply by Segment Lengths, *MSL*) ระหว่างจุดก่อนหน้าและจุดถัดไปที่เชื่อมต่อกัน ยังมีมุมที่ขนาดกว้างและมีค่าผลคูณความยาวที่มากจุดๆนั้นก็จะมีค่าสำคัญสูง

กำหนดให้ จุด v เป็นจุดมุมที่สนใจ และ vp เป็นจุดมุมตำแหน่งถัดไป และ vm เป็นจุดมุมตำแหน่งก่อนหน้า สามารถคำนวณผลคูณความยาวระหว่างจุดก่อนหน้าและจุดถัดไปที่เชื่อมต่อกันได้จาก

$$MSL = \sqrt{(x_v - x_{vm})^2 + (y_v - y_{vm})^2} \times \sqrt{(x_{vp} - x_v)^2 + (y_{vp} - y_v)^2} \quad (2)$$

ส่วนการหามุมนั้นสามารถหาได้จากสมการที่ 3

$$\theta = \arccos \left(\frac{((x_v - x_{vm}) \times (x_{vp} - x_v)) + ((y_v - y_{vm}) \times (y_{vp} - y_v))}{MSL} \right) \quad (3)$$

จากสมการที่ 2 และ 3 ทำให้สามารถคำนวณค่าความสำคัญของแต่ละจุดมุมแทนด้วย IMP หากจาก

$$IMP = MSL \times \theta \quad (4)$$

2.4. Histograms of Oriented Gradients (HOG) [5]

คือหนึ่งในวิธีการที่นิยมใช้ในการตรวจจับวัตถุ สามารถดึงลักษณะรูปร่างของวัตถุโดยใช้การกระจายตัวของความเข้มเกรเดียนท์ หรือทิศทางของเส้นขอบ โดยการทำงานจะแบ่งภาพออกไปส่วนย่อยๆ โดยส่วนย่อยแต่ละส่วนนั้นจะถูกเรียกว่า เซลล์ (Cell) ซึ่งแต่ละเซลล์นั้นจะรวบรวมฮิสโตแกรมของทิศทางของขอบภายในเซลล์ การนำฮิสโตแกรมในแต่ละเซลล์มารวมกันนั้นจะทำให้เกิดคุณลักษณะของวัตถุนั้นๆ โดยวิธีการสกัดคุณลักษณะมีขั้นตอนดังนี้

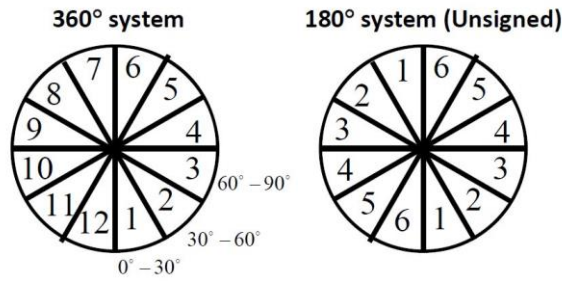
1. คำนวณหาค่าเกรเดียนท์ทั้งใน แกน x และ แกน y ของเส้นขอบจากสมการที่ 5 จากนั้นคำนวณหาขนาด (Magnitude, M) และทิศทางของเกรเดียนท์ (Direction, α) จากสมการที่ 6 และ 7 ตามลำดับ

$$\nabla f = \begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} f(x+1,y) - f(x,y) \\ f(x,y+1) - f(x,y) \end{bmatrix} \quad (5)$$

$$M(x,y) = \sqrt{f_x^2 + f_y^2} \quad (6)$$

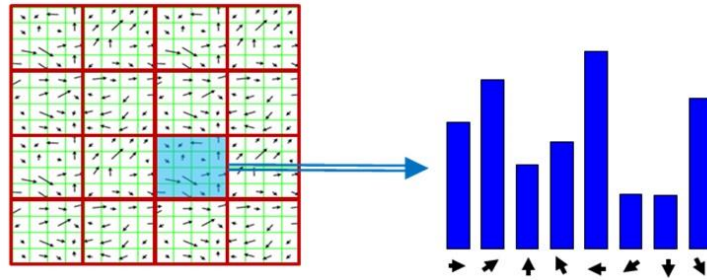
$$\alpha(x,y) = \tan^{-1} \left(\frac{f_x}{f_y} \right) \quad (7)$$

2. นำทิศทางของเกรเดียนท์ที่จำแนกกลุ่มตามทิศทางโดยระบบการแบ่งกลุ่มตามทิศทางมีอยู่ 2 ระบบ คือแบบ 360 องศา มีทั้งหมด 12 กลุ่ม และ 180 องศา มีทั้งหมด 6 กลุ่ม แสดงได้ดังรูปที่ 10



รูปที่ 2.3 ระบบการจัดกลุ่มทิศทางเกรเดียนท์

3. เมื่อจัดการจัดกลุ่มตามทิศทางเกรเดียนท์ในแต่ละพิกเซลแล้วจึงสร้างฮิสโตแกรมจากจำนวนกลุ่มของทิศทางเกรเดียนท์ซึ่งความถี่ในแต่ละกลุ่มนั้นจะเกิดจากผลรวมของขนาดของเกรเดียนท์ในกลุ่มนั้น ๆ จากนั้นทำการนอร์มอลไลซ์ (Normalization) เพื่อเพิ่มประสิทธิภาพความถูกต้องจากการปรับขนาด



รูปที่ 2.4 ตัวอย่างการสร้างฮิสโตแกรมจากทิศทางและขนาดของเกรเดียนท์

2.5. Homography Estimation

เป็นขั้นตอนของการคำนวณค่าตำแหน่งเชิง 3 มิติ ของมาร์คเกอร์เมื่อเทียบกับภาพจากกล้อง ซึ่งเป็นค่าที่ถูกแสดงอยู่ในรูปเมทริกซ์ที่เรียกว่า โฮโมกราฟีเมทริกซ์ (Homography Matrix) ซึ่งจะระบุความสัมพันธ์ระหว่างจุดที่ใช้อ้างอิงตำแหน่งใด ๆ บนภาพที่เป็นจุดเดียวกันของรูปภาพ 2 รูป [9] เช่นการระบุตำแหน่งใด ๆ บนภาพจากกล้อง (Camera Coordinated Frame) และ จุดที่อ้างอิงตำแหน่งใด ๆ ของมาร์คเกอร์ (Marker Coordinated Frame) ซึ่งจะสนใจเฉพาะที่ระนาบการฉายภาพ (Projection Plane) อยู่ในระนาบแกน Z เท่ากับ 0

เพื่อให้ได้รูป 2 มิติอยู่ในรูปการคูณกันของเมทริกซ์ได้นั้นจะต้องแทนตำแหน่งพิกัดคาร์ทีเซียน (x, y) ด้วยพิกัดเอกพันธ์ (Homogeneous coordinate) ที่มีสมาชิกเป็น $(\tilde{u}, \tilde{v}, h)$ โดยที่

$$x = \frac{\tilde{u}}{h}, y = \frac{\tilde{v}}{h} \quad (8)$$

ดังนั้นการเขียนแทนพิกัดเอกพจน์อาจสามารถเขียนได้เป็น (hx, hx, h) ซึ่งค่าพารามิเตอร์ h สามารถเป็นค่าใดก็ได้ที่ไม่เท่ากับ 0 ทำให้รูปแบบพิกัดเอกพจน์จะมีค่าเท่ากับจำนวนนับไม่ถ้วน สำหรับการเขียนแทนจุดพิกัดใดๆ ซึ่งตัวเลือกที่จะทำให้สะดวกในการคำนวณคือการตั้งค่า h เท่ากับ 1 ดังนั้นแต่ละตำแหน่ง 2 มิติจะถูกเขียนในพิกัดเอกพจน์ด้วย $(\tilde{u}, \tilde{v}, 1)$ การแสดงตำแหน่งนพิกัดเอกพจน์ช่วยให้สามารถแสดงสมการการแปลงรูปทางเลขาคณิตทั้งหมดได้ในรูปการคูณกันของเมทริกซ์ โดยสมการการแปลงรูปสามารถแสดงได้ดังสมการที่ 8

$$p_i^{(2)} = H_{21} p_i^{(1)} \quad (9)$$

โดยที่ H แทนโฮโมกราฟฟีเมทริกซ์ และ p แทนจุดที่ตำแหน่งใด ๆ ที่ตรงกันของทั้งสองรูปซึ่งแสดงแบบย่อได้ดังสมการที่ 9

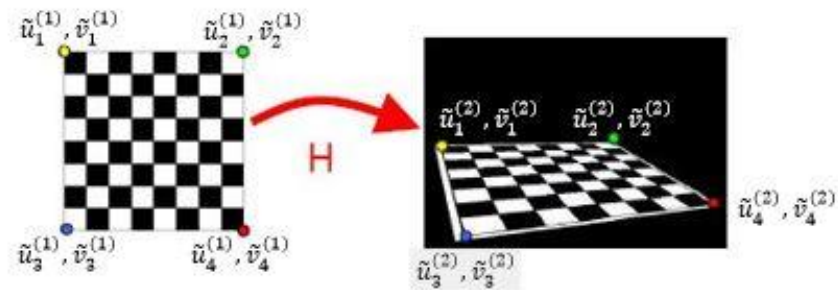
$$\begin{bmatrix} \tilde{u}^{(2)} \\ \tilde{v}^{(2)} \\ 1 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \begin{bmatrix} \tilde{u}^{(1)} \\ \tilde{v}^{(1)} \\ 1 \end{bmatrix} \quad (10)$$

เพื่อที่จะประมาณตำแหน่งของจุดใด ๆ นั้น จำเป็นต้องมีจุดที่ตรงกันอย่างน้อย 4 จุดเพื่อทำการจับคู่ลักษณะเด่นของทั้ง 2 รูป โดยสามารถใช้สมการเชิงเส้นเพื่อสร้างโฮโมกราฟฟีเมทริกซ์ได้ดังนี้

$$\begin{aligned} -\tilde{u}_1^{(1)} h_1 - \tilde{v}_1^{(1)} h_2 - h_3 + \tilde{u}_1^{(1)} \tilde{u}_1^{(2)} h_7 + \tilde{v}_1^{(1)} \tilde{u}_1^{(2)} h_8 + \tilde{u}_1^{(2)} &= 0 \\ -\tilde{u}_1^{(1)} h_4 - \tilde{v}_1^{(1)} h_5 - h_6 + \tilde{u}_1^{(1)} \tilde{v}_1^{(2)} h_7 + \tilde{v}_1^{(1)} \tilde{v}_1^{(2)} h_8 + \tilde{v}_1^{(2)} &= 0 \\ -\tilde{u}_2^{(1)} h_1 - \tilde{v}_2^{(1)} h_2 - h_3 + \tilde{u}_2^{(1)} \tilde{u}_2^{(2)} h_7 + \tilde{v}_2^{(1)} \tilde{u}_2^{(2)} h_8 + \tilde{u}_2^{(2)} &= 0 \\ -\tilde{u}_2^{(1)} h_4 - \tilde{v}_2^{(1)} h_5 - h_6 + \tilde{u}_2^{(1)} \tilde{v}_2^{(2)} h_7 + \tilde{v}_2^{(1)} \tilde{v}_2^{(2)} h_8 + \tilde{v}_2^{(2)} &= 0 \end{aligned} \quad (11)$$

$$\begin{aligned} -\tilde{u}_k^{(1)} h_1 - \tilde{v}_k^{(1)} h_2 - h_3 + \tilde{u}_k^{(1)} \tilde{u}_k^{(2)} h_7 + \tilde{v}_k^{(1)} \tilde{u}_k^{(2)} h_8 + \tilde{u}_k^{(2)} &= 0 \\ -\tilde{u}_k^{(1)} h_4 - \tilde{v}_k^{(1)} h_5 - h_6 + \tilde{u}_k^{(1)} \tilde{v}_k^{(2)} h_7 + \tilde{v}_k^{(1)} \tilde{v}_k^{(2)} h_8 + \tilde{v}_k^{(2)} &= 0 \end{aligned}$$

เพื่อที่สมการที่ 10 จะทำให้สามารถสร้างเมทริกซ์โฮโมกราฟฟีจากการหาค่า h_1 ถึง h_8 จากการแก้สมการ ทำให้สามารถประมาณจุดและมุมมองที่มีความสัมพันธ์กันระหว่าง 2 รูปได้ถูกต้อง สามารถแสดงตัวอย่างการจับคู่ดังรูปที่ 12



รูปที่ 2.5 ตัวการประมาณจุดที่สัมพันธ์กันระหว่าง 2 รูปด้วยโฮโมกราฟฟีเมทริกซ์



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved