บทที่ 2

ทฤษฎีที่เกี่ยวข้องกับงานวิจัย

ในชีวิตประจำวันเรารับรู้ข่าวสารจากโลกภายนอกผ่านการมองเห็นของดวงตาเท่า ๆ กับการใช้ประสาทรับรู้ส่วนอื่นๆ ดังนั้นภาพจึงนับว่าเป็นข้อมูลพื้นฐานที่สำคัญมาก ทั้งนี้เพราะภาพเพียงหนึ่งภาพสามารถบรรจุข่าวสารได้อย่างมหาศาลสมดังสุภาษิตจีนโบราณที่ว่า “ภาพเพียงหนึ่งภาพมีค่าเท่ากับคำนับพันคำ” โดยที่ดวงตาทำหน้าที่เป็นหน่วยรับข้อมูลประเภทนี้ หลังจากที่สมองประมวลผลแล้วก็จะสั่งงานอวัยวะที่เกี่ยวข้องให้ตอบสนองกับโลกภายนอก แต่อย่างไรก็ตามประสาทสัมผัสด้านนี้ยังมีข้อจำกัดด้านต่างๆ เช่น ความสามารถและความเร็วในการรับรู้ ดังนั้นจึงมีความจำเป็นที่จะต้องขยายขอบเขตการรับรู้ให้กว้างและมีประสิทธิภาพยิ่งขึ้น ประกอบกับวิทยาการทางด้านการสื่อสารและคอมพิวเตอร์ก้าวหน้าอย่างรวดเร็ว จึงจำเป็นที่จะต้องทำความเข้าใจเกี่ยวกับธรรมชาติของภาพและการใช้คอมพิวเตอร์เข้ามาช่วยในการประมวลผลภาพ

2.1 ภาพเชิงกายภาพ (Physical Image)

ภาพเป็นผลผลิตที่ได้จากการกระทำในกระบวนการทางแสง ซึ่งเกิดจากพลังคลื่นแม่เหล็ก ไฟฟ้าในแต่ละช่วงของความถี่ เช่น แสงธรรมดา รังสีเอกซเรย์ รังสีอินฟราเรด เป็นต้น ตกกระทบวัตถุแล้วสะท้อนมาสู่ประสาทรับรู้ของมนุษย์ แต่เนื่องจากข้อจำกัดของประสาทรับรู้ของมนุษย์ทำให้ไม่สามารถรับรู้พลังงานบางช่วงความถี่ เช่น รังสีเอกซเรย์ รังสีอินฟราเรด เป็นต้น ดังนั้นจึงมีความจำเป็นต้องใช้อุปกรณ์ตรวจจับ (Sensor) ที่เหมาะสมเพื่อแปลงพลังงานข้างต้นให้อยู่ในรูปแบบที่มนุษย์สามารถมองเห็นได้

พิจารณาภาพที่เกิดจากคลื่นแม่เหล็กไฟฟ้าช่วงที่ตาสามารถมองเห็น ซึ่งก็คือแสงสีขาวนั้นเอง โดยแสงสีขาวนี้แท้จริงแล้วประกอบด้วยแสงสีต่างๆ เช่นที่มองเห็นในรุ้งกินน้ำ แสงสีต่างๆ มีความยาวคลื่นที่ต่างกัน กล่าวคือ แสงสีแดงมีความยาวคลื่นยาวที่สุด ซึ่งหมายถึงมีพลังงานน้อยที่สุด ในทางตรงกันข้ามกันแสงสีม่วงมีความยาวคลื่นสั้นที่สุด ซึ่งหมายถึงมีพลังงานมากที่สุด สีของวัตถุที่มนุษย์มองเห็น คือ สีของแสงสะท้อนออกจากวัตถุนั้น ๆ ในขณะที่สีอื่นถูกดูดกลืน

2.2 ความหมายและนิยามของภาพในระบบดิจิตอล

ภาพ คือ ฟังก์ชัน 2 มิติ $f(x,y)$ ของค่าความเข้มของแสง โดยที่ x และ y คือค่าแสดงตำแหน่งในระบบพิกัดฉากและค่าของฟังก์ชัน f ณ ตำแหน่ง (x,y) ใด ๆ จะเป็นสัดส่วนกับความ

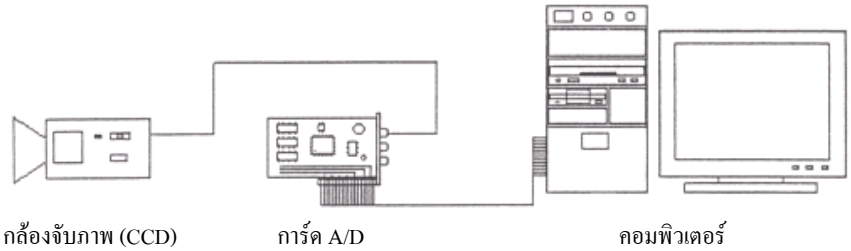
เข้มแสง ณ ตำแหน่งนั้น

พิจารณาตำแหน่งใดๆ ภายในภาพของภาพประกอบที่ 2-1 พบว่าแต่ละตำแหน่งจะมีค่าความเข้มแสงต่างกันออกไป งานในลำดับถัดไปคือแปลงค่าความเข้มของภาพนี้ให้เป็นจำนวนที่คอมพิวเตอร์สามารถคำนวณได้ ซึ่งภาพที่ได้หลังจากการแปลงคือ “ภาพเชิงดิจิตอล”

ภาพประกอบที่ 2-1 ภาพในความหมายของฟังก์ชัน 2 มิติ $f(x,y)$ ของค่าความเข้มของแสง ณ ตำแหน่ง (x,y)

ตามปกติฟังก์ชัน $f(x,y)$ จะเป็นปริมาณสกาลาร์ในกรณีที่กำลังพิจารณาภาพขาว-ดำ หรือภาพโทนสีเทา และฟังก์ชัน $f(x,y)$ จะเป็นปริมาณเวกเตอร์ในกรณีที่ค่าความเข้มแสง ณ ตำแหน่งใดๆ มีมากกว่าหนึ่งองค์ประกอบ เช่น ภาพสี ซึ่งมี 3 องค์ประกอบ คือ สีแดง สีเขียว และสีเหลือง

2.3 การแปลงภาพให้เป็นภาพในเชิงดิจิตอล



ภาพประกอบที่ 2-2 อุปกรณ์พื้นฐานสำหรับการเก็บข้อมูลภาพ

ภาพเชิงกายภาพเป็นฟังก์ชันต่อเนื่อง (Continuous Function) ถ้าต้องการที่จะนำภาพนี้มาประมวลผลโดยใช้คอมพิวเตอร์ จำเป็นอย่างยิ่งที่จะต้องดำเนินการแปลงภาพจากภาพเชิงกายภาพ

ให้เป็นภาพเชิงดิจิตอลซึ่งเป็นฟังก์ชันแบบเต็มหน่วย (Discrete Function) และเรียกกระบวนการนี้ว่า Image Digitization ซึ่งประกอบด้วย 3 ขั้นตอนดังนี้

2.3.1 การบันทึกภาพ (Image Acquisition)

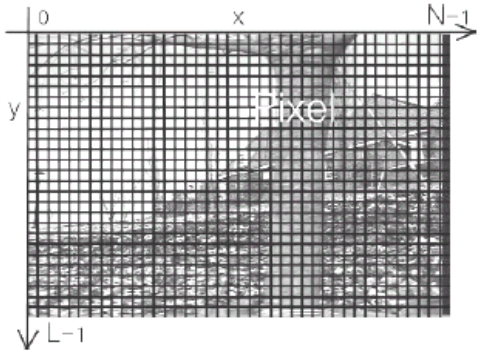
การบันทึกภาพด้วยอุปกรณ์เชิงแสง (Optical Device) เป็นการแปลงภาพเชิงต่อเนื่อง 3 มิติซึ่งเป็นภาพที่สามารถมองเห็นในธรรมชาติรอบตัวเราให้เป็นภาพเชิงต่อเนื่อง 2 มิติ เช่น ภาพบนฟิล์ม รูปถ่ายบนกระดาษ ภาพบนจอภาพของคอมพิวเตอร์

การบันทึกภาพแบบดั้งเดิม (แบบอนาลอก) เป็นกระบวนการที่เกิดจากการทำปฏิกริยาเคมีระหว่างแสงกับสารเคมีบนแผ่นฟิล์ม เนื่องจากสารเคมีนี้ไวต่อแสงต่างกันทำให้สิ่งที่ปรากฏบนแผ่นฟิล์มหรือกระดาษอัดรูปเป็นสัดส่วนกับความเข้มแสงที่ตกกระทบ

การบันทึกด้วยอุปกรณ์อิเล็กทรอนิกส์ที่ให้ผลลัพธ์เป็นปริมาณดิจิตอล เป็นกระบวนการแปลงความเข้มแสงให้เป็นกระแสไฟฟ้าด้วยอุปกรณ์ตรวจจับที่เรียกว่า Charge-Coupled Device (CCD) Sensor

2.3.2 การสุ่มเลือกทางตำแหน่ง (Spatially Image Sampling)

หลังจากได้ภาพซึ่งเป็นภาพต่อเนื่อง 2 มิติแล้ว กระบวนการถัดไปคือการแปลงภาพนี้ให้เป็นภาพเชิงดิจิตอลแบบเต็มหน่วยโดยการสุ่มเลือกทางตำแหน่ง (Spatially Sampling) โดยจะสุ่มเลือกวัดค่าของ $f(x,y)$ เฉพาะบางจุดบนภาพทั้งแนวแกน x และ y ค่าความเข้มแสงของแต่ละตัวอย่างที่ถูกสุ่มเลือกมาจะสมนัยกับค่าความเข้มแสงที่ตกกระทบบนแต่ละหน่วยที่มีรูปร่างเป็นสี่เหลี่ยมเล็ก ๆ บนภาพ (เพื่อให้ง่ายต่อการทำความเข้าใจ อาจแสดงให้เห็นได้ด้วยพื้นที่ที่ถูกแบ่งด้วยเส้นตรงทั้งสองแนวดังแสดงในภาพประกอบที่ 2-3) และเรียกแต่ละจุดนี้ว่าจุดภาพ (Picture Element หรือ Pixel) ผลที่ได้นี้สามารถแทนด้วยเมตริกซ์ของปริมาณเชิงจำนวนเต็ม (Integer) ขนาด L x N โดย L และ N จำนวนจุดภาพในแนวแกนดิ่งและระดับตามลำดับ และจุดกำเนิดของระบบอ้างอิงที่ใช้ในภาพจะอยู่ที่มุนบนด้านซ้ายของภาพเสมอ



ภาพประกอบที่ 2-3 ภาพอุปมาอุปมัยของการสุ่มเลือกทางตำแหน่งของภาพ (อนึ่ง ขณะปฏิบัติจริงการสุ่มเลือกทั้งสองแกนจะมีความถี่มากกว่าในภาพนี้)

$$\mathbf{F} = \begin{bmatrix} f(0,0) & f(0,1) & f(0,2) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & f(1,2) & \dots & f(1,N-1) \\ - & - & - & - & - \\ f(L-1,0) & f(L-1,1) & f(L-1,2) & & f(L-1,N-1) \end{bmatrix}$$

ดังนั้น $f(2,3)$ แทนจุดภาพ ณ ตำแหน่งแถวที่ 2 คอลัมน์ที่ 3 และ $f(L-1,N-1)$ แทนจุดภาพ ณ ตำแหน่งแถวสุดท้าย คือ L-1 คอลัมน์สุดท้าย คือ N-1 ในทางปฏิบัติค่าของ L และ N นิยมเป็นค่าจำนวนเต็มที่เป็นกำลังสอง (Power of Two) เช่น 1024(2^{10}), 512(2^9), 256(2^8) เป็นต้น ดังนั้น

$L = 2^k, N = 2^n$ โดยที่ k และ n เป็นค่าคงที่ใด ๆ

อนึ่งคุณภาพของภาพเชิงดิจิตอลที่ได้นี้ขึ้นอยู่กับความถี่ในการสุ่มเลือกทางตำแหน่งทั้งสองแนวแกน กล่าวคือ ถ้าหากสุ่มเลือกอย่างละเอียดก็จะได้จำนวนจุดภาพมาก นั่นหมายถึง คุณภาพที่ดีของภาพที่ได้ แต่ผลกระทบที่ตามมาคือความต้องการเนื้อที่เพื่อจัดเก็บภาพในหน่วยความจำของคอมพิวเตอร์มากขึ้น และในทางกลับกันถ้าหากสุ่มอย่างหยาบก็จะได้จำนวนจุดภาพน้อย ภาพที่ได้มีคุณภาพไม่ดีแต่ใช้เนื้อที่เพื่อจัดเก็บน้อยกว่า

2.3.3 การประมาณค่าความเข้มแสงของแต่ละจุดภาพ (Quantization)

หลังจากได้สุ่มเลือกจุดภาพมาแล้วขั้นตอนถัดไปคือกำหนดค่าให้กับความเข้มแสงของแต่ละจุดภาพนั้น โดยในกรณีของภาพโทนสีขาวจะถูกแทนด้วยค่า 255 และสีดำถูกแทนด้วยค่า 0 ในขณะที่ระดับโทนสีจะเข้มขึ้นเรื่อย ๆ จากสีขาวจนถึงสีดำดังแสดงในภาพประกอบที่ 2-4 ค่าเหล่านี้จะถูกแทนด้วยจำนวนเต็มที่มีค่าระหว่างค่าทั้งสองนี้ โดยจะแบ่งออกเป็น 256 ระดับสี หรือ 2^8 โดย 8 คือ จำนวนบิตในหน่วยความจำที่ใช้เพื่อเก็บค่านี้หนึ่งค่า ดังนั้นสีดำจะถูกแทนด้วย 00000000 และสี ขาวจะถูกแทนด้วย 11111111 (ซึ่งใช้เนื้อที่ในหน่วยความจำ 8 บิตหรือ 1ไบต์)

ภาพประกอบที่ 2-4 การแทนค่าความเข้มแสงด้วยจำนวนเต็มที่มีค่าอยู่ระหว่าง 0 และ 255

ในกรณีที่เป็นภาพขาว-ดำ ค่าความเข้มแสง ณ จุดภาพใด ๆ จะมีค่าที่เป็นไปได้เพียง 2 ($2=2^1$) ค่าเท่านั้นคือ 0 หรือไม่ก็ 255 ดังนั้นหนึ่งจุดภาพจะต้องการพื้นที่เพื่อเก็บข้อมูลที่มีความยาวเพียง 1 บิตเท่านั้น

ถ้าให้ G แทนระดับสีทั้งหมดของภาพที่กำลังพิจารณา จะได้ $G=2^b$ เมื่อ b คือ จำนวนบิตของหน่วยความจำที่ใช้เก็บภาพหนึ่งจุดภาพ และถ้าให้ M คือ ขนาดหน่วยความจำที่ต้องการเพื่อเก็บภาพทั้งภาพ (มีหน่วยเป็นบิต) จะได้

$$M = L \times N \times b \text{ หรือ } M = N^2 \times b \text{ ถ้า } L=N \qquad (2\text{-}1)$$

ตัวอย่างเช่น ภาพโทนสีเทาที่ถูกแทนด้วย 256 ระดับสี ขนาดความกว้าง 128 จุดภาพ และยาว 128 จุดภาพ ต้องการหน่วยความจำขนาด 128 x 128 = 131,072 บิต

ทำนองเดียวกับความละเอียดของการสุ่มเลือกทางตำแหน่งที่มีผลต่อคุณภาพของภาพที่ได้ ระดับสีที่ประมาณจากค่าความเข้มแสงก็มีผลต่อคุณภาพของภาพเช่นกัน กล่าวคือรายละเอียดของภาพจะลดลงเมื่อจำนวนระดับสีลดลง

2.4 ลักษณะของภาพสี (Color Image Characteristics)

โดยทั่วไปภาพสีเชิงดิจิตอลจะถูกเก็บในลักษณะเป็นค่าเวกเตอร์ กล่าวคือ แต่ละจุดภาพจะถูกแทนด้วยค่า 3 ค่า ซึ่งเป็นค่าความเข้มแสงของสีแดง สีเขียว และสีน้ำเงิน ผลรวมของค่าทั้งสามใช้แทนสีผสมอื่นๆ อนึ่งเพื่อที่จะทำให้ง่ายต่อการทำความเข้าใจจะแทนภาพสีด้วยภาพ 3 ระนาบดังภาพประกอบที่ 2-5 การประมวลภาพสีจึงไม่แตกต่างจากการประมวลผลภาพโทนสีเทามากนัก เพียงแต่ประยุกต์ใช้กระบวนการนี้บนแต่ละระนาบสี หรือทำการแปลงแบบจำลองสีจากระบบ RGB ดังกล่าวข้างต้นไปสู่ระบบอื่น เช่น Yc_rc_b ก่อนแล้วจึงประมวลผลบนองค์ประกอบที่เหมาะสมของแบบจำลองสีระบบใหม่นั้น เช่น องค์ประกอบ Y เท่านั้น

(a) (b) (c) (d)

ภาพประกอบที่ 2-5 ภาพสี (a) ซึ่งมี 3 องค์ประกอบคือ สีแดง (b), สีเขียว (c), และสีน้ำเงิน (d)

หลังจากภาพถูกแปลงเป็นปริมาณดิจิตอลซึ่งคอมพิวเตอร์สามารถเข้าใจได้แล้ว งานในลำดับถัดไปก็คือการประมวลผลที่ได้นี้เพื่อวัตถุประสงค์ต่าง ๆ ซึ่งขึ้นอยู่กับงานที่จะนำภาพไปประยุกต์ใช้

2.5 ความสัมพันธ์เบื้องต้นระหว่างจุดภาพ (Fundamental Relation Between Pixels)

ความสัมพันธ์ระหว่างจุดภาพเป็นการกล่าวถึงระบบเพื่อนบ้านรวมถึงการเชื่อมโยงกันระหว่างจุดภาพ ซึ่งความสัมพันธ์จะถูกใช้เป็นพื้นฐานสำหรับการประมวลผลแต่ละวิธี

เป็นที่ทราบกันโดยทั่วไปว่า มีความน่าจะเป็นสูงมากที่จุดภาพ 2 จุดภาพที่อยู่ติดกัน จะอยู่ในบริเวณของวัตถุเดียวกัน ดังนั้นการศึกษาถึงความสัมพันธ์ระหว่างจุดภาพที่เป็นเพื่อนบ้านซึ่งกันและกันจึงเป็นสิ่งที่จำเป็นและควรกระทำ

2.5.1 ระบบเพื่อนบ้านของจุดภาพ (Neighbor System of a Pixel)

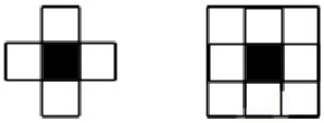
เมื่อพิจารณาจุดภาพใดๆ จำเป็นอย่างยิ่งที่จะต้องพิจารณาอิทธิพลของเพื่อนบ้านรอบข้างด้วย เพราะโดยธรรมชาติของภาพแล้ว จุดภาพที่อยู่ใกล้เคียงกันมีความเป็นไปได้สูงมากที่จะถูกครอบคลุมโดยวัตถุต้นกำเนิดเดียวกัน

4-Neighbors System ($N_4(p)$) ของจุดภาพ $f(x,y)$ ใดๆ คือ จุดภาพรอบข้างที่อยู่ในแนวราบ (ซ้าย-ขวา) และในแนวดิ่ง (บน-ล่าง) ดังแสดงภาพประกอบที่ 2-6 a. นั่นคือ ประกอบด้วยจุดภาพ ณ ตำแหน่งต่าง ๆ ดังต่อไปนี้ $f(x-1,y)$, $f(x+1,y)$, $f(x,y-1)$ และ $f(x,y+1)$ โดยในที่นี้จุดภาพที่ระบายสีทึบ (p) คือ จุดภาพที่กำลังพิจารณา

Diagonal Neighbors ($N_D(p)$) ของจุดภาพ f(x,y) ใดๆ คือ จุดภาพรอบข้างที่อยู่ในแนวทแยงมุมทั้งสองลักษณะ (ซ้าย-ขวา) ประกอบด้วยจุดภาพ ณ ตำแหน่งต่าง ๆ ดังต่อไปนี้ $f(x-1,y-1)$, $f(x+1, y-1)$, $f(x-1,y+1)$ และ $f(x+1,y+1)$

8-Neighbors System ($N_8(p)$) ของจุดภาพ $f(x,y)$ ใด ๆ คือ จุดภาพรอบข้างตามลักษณะของ 4-Neighbors System ข้างต้นและจุดภาพที่อยู่ในแนวทแยง ดังแสดงในภาพประกอบที่ 2-6 b. นั่นคือประกอบไปด้วยจุดภาพ ณ ตำแหน่งต่าง ๆ ดังต่อไปนี้ $f(x-1,y)$, $f(x+1,y)$, $f(x,y-1)$, $f(x,y+1)$, $f(x-1,y-1)$, $f(x+1,y-1)$, $f(x-1,y+1)$ และ $f(x+1,y+1)$ โดยในที่นี้จุดภาพที่ระบายสีทึบ (p) คือ จุดภาพที่กำลังพิจารณา

.



ภาพประกอบที่ 2-6 แสดงระบบเพื่อนบ้านของจุดภาพใด ๆ (ระบายสีดำ) โดย a คือระบบ 4-Neighbors และ b คือระบบ 8-Neighbors ของจุดภาพนี้

2.5.2 ความเชื่อมโยงกัน (Connectivity)

อาศัยลักษณะระบบเพื่อนบ้านทั้งสองที่กล่าวมาแล้วในหัวข้อก่อนนี้ เราสามารถนิยามคำจำกัดความต่าง ๆ ได้ดังนี้

1. ความเชื่อมโยงกัน (Connectivity)

ถ้าจุดภาพ p และ q เชื่อมโยงกัน จุดภาพทั้งสองต้องเป็นสมาชิกของระบบเพื่อนบ้านข้างต้นและมีค่าความเข้มแสงเท่ากันหรือใกล้เคียงกัน (ภายในขอบเขตที่กำหนด) นั้นคือ

4-Connectivity : จุดภาพ p และ q เชื่อมโยงกันแบบ 4-Connected ก็ต่อเมื่อจุดภาพทั้งสอง (1). มีค่าความเข้มแสงอยู่ในเซต V และ (2). ต้องเป็นสมาชิกของระบบเพื่อนบ้าน 4-Neighbors

8-Connectivity : จุดภาพ p และ q เชื่อมโยงกันแบบ 8-Connected ก็ต่อเมื่อจุดภาพทั้งสอง (1). มีค่าความเข้มแสงอยู่ในเซต V และ (2). ต้องเป็นสมาชิกของระบบเพื่อนบ้าน 8-Neighbors

m-Connectivity : จุดภาพ p และ q เชื่อมโยงกันแบบ m-Connected ก็ต่อเมื่อจุดภาพทั้งสอง (1). มีค่าความเข้มแสงอยู่ในเซต V (2). ต้องเป็นสมาชิกของระบบเพื่อนบ้าน 4-Neighbors หรือ (3). เป็นสมาชิกของระบบเพื่อนบ้าน D- Neighbors และ $N_4(P) \cap N_4(q) \in \varnothing$

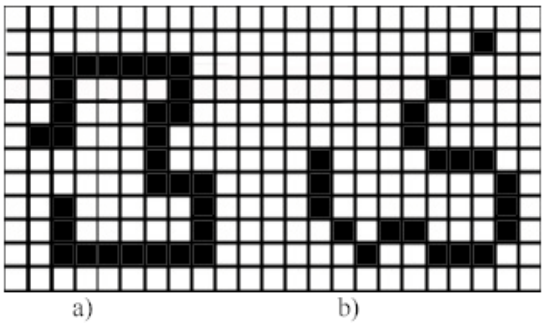
โดยเซต V คือเซตของค่าความเข้มแสง ถ้าภาพที่กำลังพิจารณาเป็นภาพขาว-ดำ V={เซตของค่าความเข้มแสงของส่วนที่เป็น Foreground} ถ้าภาพที่กำลังพิจารณาเป็นภาพโทนสีเทา V={เซตของค่าความเข้มแสงช่วงที่เป็น Foreground ของส่วนภาพที่กำลังพิจารณา}

2. เส้นทางเดิน (Path) คือ ลำดับของจุดภาพที่เรียงต่อเนื่องกันในลักษณะที่จุดภาพหนึ่งเป็นเพื่อนบ้านในระบบใดระบบหนึ่งข้างต้นกับจุดภาพข้างเคียงตลอดเส้นทางนี้

ถ้าสอดคล้องกับระบบ 4-Connected เส้นทางเดินจะเป็น 4-Path

ถ้าสอดคล้องกับระบบ 8-Connected เส้นทางเดินจะเป็น 8-Path

ถ้าสอดคล้องกับระบบ m-Connected เส้นทางเดินจะเป็น m-Path

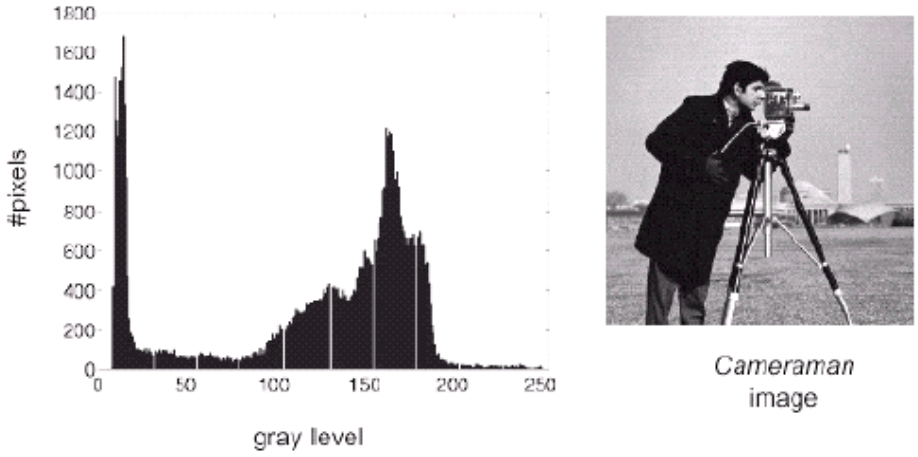


ภาพประกอบที่ 2-7 แสดงเส้นทางเดิน โดย (a) เป็นเส้นทางเดินแบบ 4-Path และ (b) เป็นเส้นทางเดินแบบ 8-Path

3. องค์ประกอบที่เชื่อมโยงถึงกัน (Connected Component) จุดภาพใดๆ ที่อยู่ในบริเวณภาพ S เป็นองค์ประกอบที่เชื่อมโยงถึงกันของภาพ S ก็ต่อเมื่อจุดภาพนี้เชื่อมโยงกับจุดภาพ P ใดๆ ที่กำหนดให้

2.6 ฮิสโตแกรม (Histogram)

ฮิสโตแกรมเป็นฟังก์ชั่นที่แสดงว่าในแต่ละระดับความเข้มมีจำนวนจุดภาพในภาพที่มีค่าความเข้มนี้เท่าไร เมื่อนำมาเขียนเป็นกราฟ แกนนอนคือระดับความเข้ม (Gray Level) และแกนยืนคือความถี่ของความเข้มนั้น ๆ (Number of Pixels) ฮิสโตแกรมเป็นการนำเสนอภาพในลักษณะรวม ๆ (Global Characteristics) ภาพประกอบที่ 2-8 แสดงฮิสโตแกรมของภาพทางขวามือ



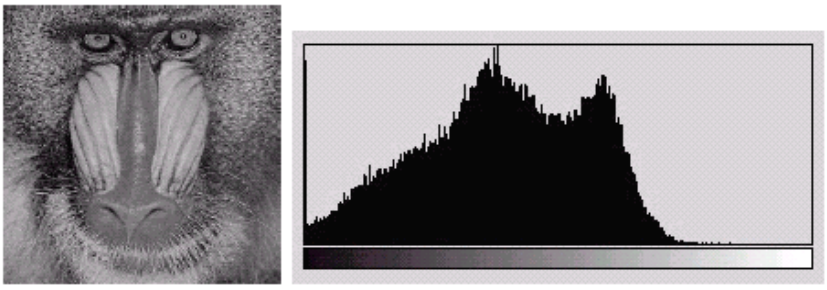
ภาพประกอบที่ 2-8 ภาพทางซ้ายมือแสดงฮิสโตแกรมของภาพชุด Cameraman ทางขวามือ

ให้ h_i เป็นฮิสโตแกรมของระดับ i ซึ่งสามารถคำนวณได้จากค่าร้อยละของจำนวนจุดภาพที่มีความเข้มระดับนี้ (n_i) เทียบกับจำนวนจุดภาพทั้งหมด (กว้าง L จุดภาพ และยาว N จุดภาพ)

$$h_i = n_i / LN \text{ สำหรับ } 0 \leq i \leq 255 \qquad (2\text{-}2)$$

ลักษณะที่สำคัญของฮิสโตแกรมคือ

1. ค่าเฉลี่ยของความเข้ม ฮิสโตแกรมที่มีค่าความเฉลี่ยต่ำ (เอียงมาทางค่าความเข้มต่ำ) จะแทนภาพที่มืด และฮิสโตแกรมที่มีค่าเฉลี่ยสูง (เอียงมาทางค่าความเข้มสูง) จะแทนภาพที่สว่าง

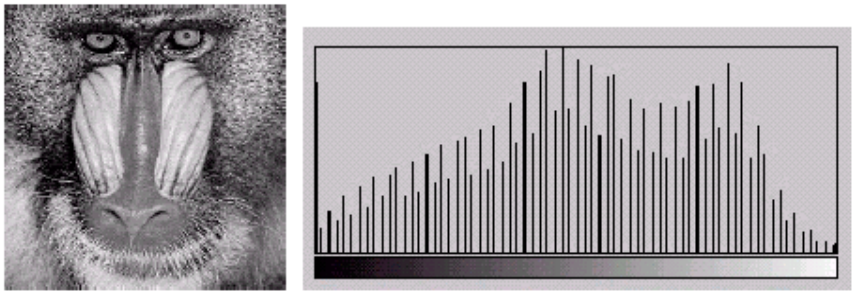


(a)

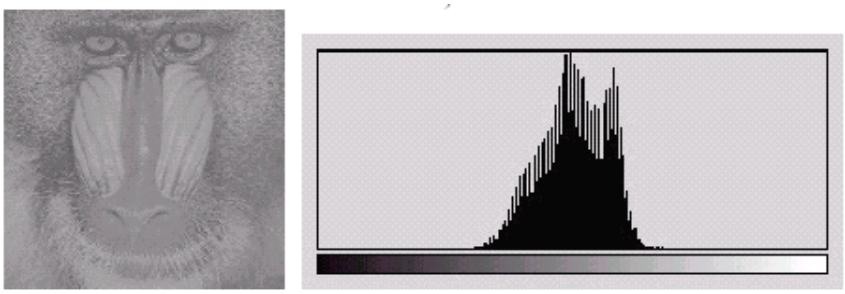
(b)

ภาพประกอบที่ 2-9 ค่าเฉลี่ยของค่าความเข้มของภาพ (a) ฮิสโตแกรมที่มีค่าเฉลี่ยของค่าความเข้มต่ำจะแทนภาพที่มืด (b) ฮิสโตแกรมที่มีค่าเฉลี่ยของค่าความเข้มสูงจะแทนภาพที่สว่าง

2. ลักษณะของการกระจายของฮิสโตแกรม จะแสดงความคมชัดของภาพ นั่นคือ ฮิสโตแกรมที่มีการกระจายเป็นแถบกว้างจะแทนภาพที่มีความคมชัดสูง และฮิสโตแกรมที่มีการกระจายเป็นแถบแคบจะแทนภาพที่มีความคมชัดต่ำ



(a)



(b)

รูปที่ 2-10 ความคมชัดของภาพ (a) ฮิสโตแกรมที่มีการกระจายเป็นแถบกว้างจะแทนภาพที่มีความคมชัดสูง (b) ฮิสโตแกรมที่มีการกระจายเป็นแถบแคบจะแทนภาพที่มีความคมชัดต่ำ

2.7 การปรับปรุงคุณภาพของภาพ (Image Enhancement)

การปรับปรุงคุณภาพของภาพเป็นกระบวนการที่ทำให้คุณภาพของภาพดีขึ้นเพื่อจุดหมายเฉพาะ โดยการกระทำที่ทำให้คุณลักษณะของภาพ เช่น ความมืด ความสว่าง ขอบของวัตถุในภาพ ให้ปรากฏชัดเจนขึ้น รวมทั้งกำจัดสิ่งรบกวน การปรับปรุงภาพแบ่งออกเป็นประเภทของการประมวลผลได้เป็น 4 แบบ คือ

1. Pixel-based Image Enhancement เป็นการปรับภาพด้วยวิธีที่กระทำครั้งละหนึ่งจุดภาพจนหมด ซึ่งได้แก่ การกระทำทางพีชคณิต การกระทำเชิงตรรก การกระทำทางเรขาคณิต

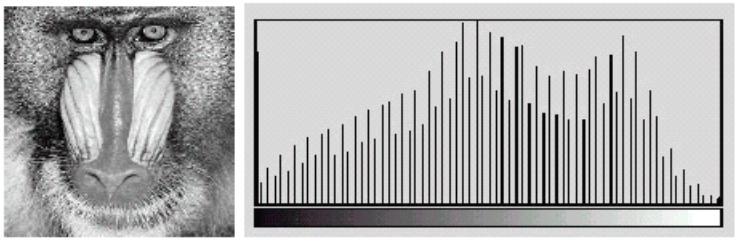
2. Histogram-based Image Enhancement เป็นการปรับปรุงคุณภาพของภาพโดยการกระทำบนฮิสโตแกรมของภาพที่ต้องการปรับปรุงคุณภาพนั้น

3. Spatial-filtering-based Image Enhancement เป็นการปรับปรุงคุณภาพของภาพโดยใช้กระบวนการ spatial convolution

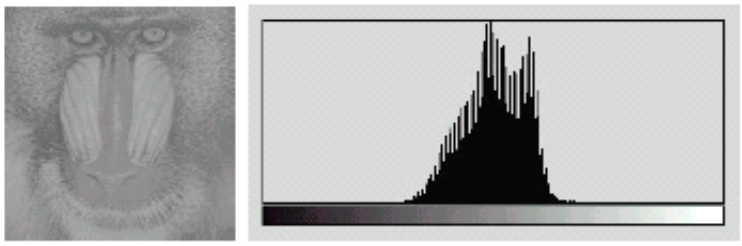
4. Frequency-based Image Enhancement เป็นการปรับปรุงคุณภาพของภาพโดยกระทำบน Frequency domain โดยใช้การแปลงฟูเรียร์

2.7.1 ความคมชัดของภาพ (Contrast)

ความคมชัด คือ ลักษณะความเด่นชัดของเส้นและขอบในภาพ โดยวัดได้ด้วยขอบเขตของค่าของระดับความเข้มแสงที่ปรากฏในภาพ ภาพที่มีระดับความเข้มแสงเป็นขอบเขตกว้าง ดังภาพประกอบที่ 2-11 a. ซึ่งสังเกตจากฮิสโตแกรมพบว่า ภาพนี้มีค่าความเข้มแสงตั้งแต่ 0 ถึง 255 ในขณะที่ภาพประกอบที่ 2-11 b. มีขอบเขตความเข้มแสงที่แคบกว่า



(a)



(b)

ภาพประกอบที่ 2-11 ภาพเปรียบเทียบความคมชัดของภาพ โดยภาพ (a) เป็นภาพที่มีความคมชัดสูง ภาพ (b) เป็นภาพที่มีความคมชัดต่ำ

ค่าความคมชัดจึงสามารถคำนวณได้จาก

$$\text{contrast} = c = \sqrt{\frac{1}{MN}\sum_{y=0}^{M-1}\sum_{x=0}^{N-1}[i(x, y) - B]^2} \qquad (2\text{-}3)$$

2.7.2 การปรับปรุงภาพโดยกระทำการกรองทางตำแหน่ง (Spatial filtering-based image enhancement)

1. การกรองแบบเฉลี่ย (Mean Moving Average Filtering)

ลักษณะ เป็นการกรองโดยใช้ spatial convolution ซึ่งมีค่าน้ำหนักเท่ากัน คือ $1/s^2$ เมื่อ s คือขนาดของ kernel เช่น น้ำหนักแต่ละช่องจะเป็น 1/9 ถ้า kernel มีขนาด 3x3 เป็นต้น การกรองชนิดนี้เป็นแบบเชิงเส้น และเป็นชนิดที่ยอมให้ความถี่ต่ำผ่านได้ (low-pass filter)

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

ดังนั้น $g(x, y) = \frac{1}{N_t}\sum_{i,}\sum_{j\in H} i(x-i, y-i)$ ถ้า i(x,y) เป็นภาพต้นฉบับที่ปนเปื้อนสัญญาณรบกวน g(x,y) เป็นภาพหลังจากการทำการกรองแล้ว N_t เป็นจำนวนช่องของ kernel H

ข้อดี สามารถกำจัดสัญญาณรบกวนแบบ Gaussian noise ได้ดี

ข้อเสีย ทำให้ภาพเบลอ เส้นและขอบไม่ชัดเจน กำจัดสัญญาณรบกวนแบบ salt-pepper noise ได้ไม่ดี

2. การกรองแบบ Gaussian

ลักษณะ เป็นการกรองโดยใช้ spatial convolution ในลักษณะที่เป็นสมมาตรในเชิงวงกลม ซึ่งค่าน้ำหนักจะมีค่าแปรผันตามลักษณะการกระจายแบบ Gaussian distribution คือ

$$h(x, y) = \ell^{-\frac{\pi(x^2+y^2)}{a^2}} \qquad (2\text{-}4)$$

เมื่อ a คือความกว้างของตัวกรองชนิดนี้เป็นแบบเชิงเส้น และเป็นชนิดที่ยอมให้ความถี่ต่ำผ่านได้เช่นกัน

3/25	6/25	8/25	6/25	3/25
6/25	14/25	19/25	14/25	6/25
8/25	19/25	25/25	19/25	8/25
6/25	14/25	19/25	14/25	6/25
3/25	6/25	8/25	6/25	3/25

ข้อดี สามารถกำจัดสัญญาณรบกวนแบบ Gaussian noise ได้ดี

ข้อเสีย ทำให้ภาพเบลอ กำจัดสัญญาณรบกวนแบบ salt-pepper noise ได้ไม่ดี

3. การกรองแบบ Median (Order Statistic Filtering)

ลักษณะ เป็นการกรองโดยใช้ spatial convolution ซึ่งมีน้ำหนักเป็น 1 เท่ากันตลอด หลังจากทำ convolution แล้วให้นำผลลัพธ์มากระทำ median โดยเลือกค่าตรงกลาง นั้นคือ

$$g(i,j) = \text{med}\,(x(i+r,j+s),(r,s)\in A,(i,j)\in Z^2) \qquad (2\text{-}5)$$

การกรองชนิดนี้เป็นแบบไม่เป็นเชิงเส้น และเป็นชนิดที่ยอมให้ความถี่ต่ำผ่านได้เช่นกัน

1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1

ข้อดี สามารถกำจัดสัญญาณรบกวนแบบ salt-pepper noise (ซึ่งเป็นสัญญาณรบกวนแบบ implusive หรือ outlier noise) และ Laplacian noise ได้ดี อีกทั้งยังไม่ทำให้ภาพเบลอ ภาพและขอบยังคงชัดเจน

ข้อเสีย ทำให้ภาพเบลอ กำจัดสัญญาณรบกวนแบบ Gaussian noise ได้ไม่ดี

4. การกรองแบบ Max / Min (2d-Order Statistic Filtering)

ตัวกรองแบบ Max เป็นตัวกรองที่ทำการขยายจุดภาพที่สว่าง (bright spots) ส่วนตัวกรองแบบ Min เป็นตัวกรองที่ทำการขยายจุดภาพที่มืด (dark spots) ดังภาพประกอบที่ 2-12

(a) (b) (c)

ภาพประกอบที่ 2-12 Cameraman Image (a) Original (b) ผลของ Max Filtering (c) ผลของ Min Filtering

2D Order Statistic Filtering จะแทนค่าแต่ละส่วนใน 3*3 neighbourhood ด้วยส่วน maximum หรือส่วน minimum ใน nieghbourhood

สมการ Max/Min Filtering

$$\hat{f}(x, y) = \max_{(s,t)\in s_{xy}}\{g(s, t)\} \quad (2\text{-}6)$$

$$\hat{f}(x, y) = \min_{(s,t)\in s_{xy}}\{g(s, t)\} \quad (2\text{-}7)$$

5. การกรองแบบ High-pass (Sharpening Edge Crispening)

ลักษณะ เป็นการกระทำที่มุ่งเน้นทำให้ภาพคมชัดมากกว่าจะกรองสัญญาณรบกวนออกจากภาพ ทำให้รายละเอียดของภาพปรากฏเด่นชัดขึ้น ทำได้โดยการใช้ spatial convolution ซึ่งมีค่าน้ำหนักตรงกลางเป็นค่าบวก และค่าน้ำหนักในช่องอื่น ๆ เป็นค่าลบ และ/หรือศูนย์ เพื่อที่จะทำให้ผลรวมของค่าน้ำหนักทุกช่องบน Kernel รวมกันเป็น 1 การกรองชนิดนี้ยังคงเป็นแบบเชิงเส้น ตัวอย่าง kernel ขนาด 3x3 ที่ถูกใช้บ่อย ๆ เช่น

0	-1	0
-1	5	1-
0	-1	0

1	-2	1
-2	5	-2
1	-2	1

-1	-1	-1
-1	9	-1
-1	-1	-1

ข้อดี ทำให้รายละเอียดของภาพคมชัดขึ้น

ข้อเสีย ทำให้สัญญาณรบกวนเด่นชัดขึ้นด้วย

Unsharp Filtering

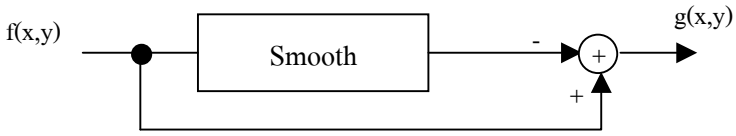
Unsharp filtering เป็นโอเปอเรเตอร์สำหรับปรับความคมชัดที่ง่าย ชื่อมาจากความสามารถในการปรับเพิ่มความเด่นชัดของขอบภาพ (และส่วนอื่นๆ ของภาพที่มีความถี่สูง) โดยนำภาพต้นแบบ(Original image) มาลบกับ Unsharp หรือ smooth

Unsharp filtering เป็นเทคนิคพื้นฐานที่ใช้ในงานกราฟฟิกและการพิมพ์ crispening edge ในงานอุตสาหกรรม

การทำ Unsharp จะให้ผลของขอบภาพ g(x,y) จากภาพอินพุท $f(x,y)$ โดย

$$g(x,y) = f(x,y) - f_{smooth}(x,y) \quad (2\text{-}8)$$

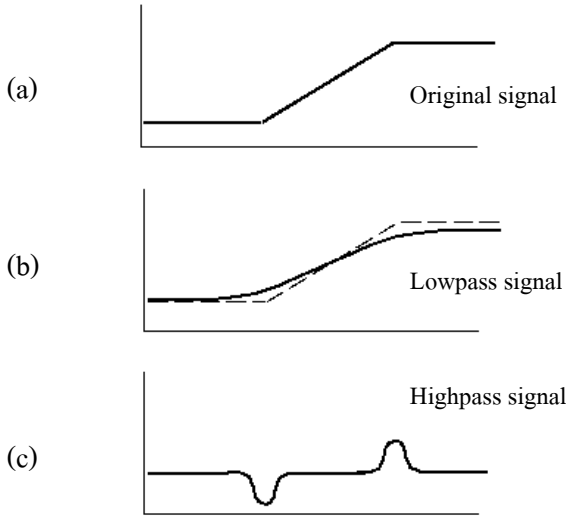
เมื่อ $f_{smooth}(x,y)$ คือ smooth เวอร์ชันของ $f(x,y)$ ดังภาพประกอบที่ 2-13



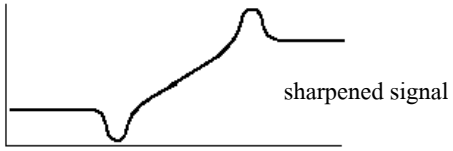
ภาพประกอบที่ 2-13 spatial sharpening

เราสามารถเข้าใจการทำงานของ Unsharp filtering ได้ดีขึ้น ด้วยการทดสอบคุณสมบัติการตอบสนองทางความถี่ ถ้าเรานำสัญญาณดังภาพประกอบที่ 2-14 (a) มาลบส่วน lowpass ของสัญญาณออกไป (ภาพประกอบที่ 2-14(c)) ผลของ highpass หรือขอบสามารถแสดงให้เห็นได้ดังภาพประกอบที่ 2-14(c)

เมื่อนำสัญญาณนี้รวมเข้ากับสัญญาณอินพุท (Original image) จึงสามารถปรับเพิ่มความคมชัดของขอบภาพ

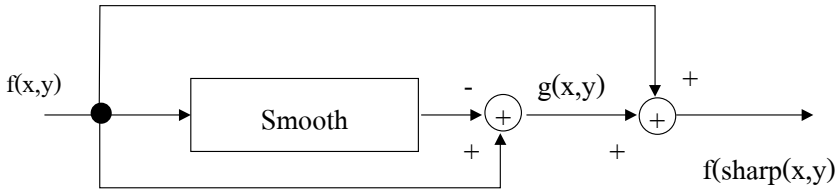


ภาพประกอบที่ 2-14 การคำนวณขอบภาพสำหรับ Unsharp filtering



ภาพประกอบที่ 2-15 Sharpening the original signal using the edge image.

ดังนั้น Unsharp Sharpening โอเปอเรเตอร์ที่สมบูรณ์ แสดงได้ดังภาพประกอบที่ 2-16



ภาพประกอบที่ 2-16 The Complete Unsharp Filtering Operator.

สมการคือ

$$f_{sharp}(x,y) = f(x,y) + k * g(x,y) \qquad (2\text{-}9)$$

เมื่อค่า k เป็นค่าคงที่สกาลาร์และเปลี่ยนแปลงอยู่ในช่วง 0.2-0.7 ค่า k ที่เพิ่มขึ้นเป็นการเพิ่มความคมชัด

2.8 การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพ (Morphological Image Processing)

Morphological Image Processing เป็นการประมวลผลภาพโดยการเปลี่ยนแปลงลักษณะรูปร่างหรือโครงสร้างของภาพ โอเปอเรชั่นพื้นฐานโดยทั่วไปได้แก่ การ Dilation Erosion และ Skeleton โดยการ Dilation คือการขยายภาพโดยมีสัดส่วนเท่ากันทั่วทั้งภาพ การ Erosion คือการย่อภาพ ส่วนการทำ Skeleton เป็นการหาโครงสร้างหลักของวัตถุ

2.81 เทคนิคของการ Hit และ Miss

โอเปอเรชั่นพื้นฐานสำหรับการกระทำกับรูปร่างหรือโครงสร้างของภาพ ไม่ว่าจะเป็น

การย่อหรือการขยายภาพ จำเป็นที่จะต้องมีการนำเอาเทคนิคการ Hit และ Miss มาใช้ แนวคิดนี้คือ การกำหนดให้มีเมตริกซ์ (Template) ที่มีขนาดเล็ก ๆ และเป็นจำนวนคี่ โดยทั่วไปจะมีค่าเท่ากับ 3x3 ไปเปรียบเทียบกับข้อมูลภาพ โดยการเปรียบเทียบจะทำตลอดทั้งภาพจนถึงท้ายภาพ ถ้าข้อมูลของภาพมีลักษณะเหมือนกับเมตริกซ์ดังกล่าวเอ้าพุทที่ได้ขึ้นอยู่กับพิกเซลที่เป็นศูนย์กลางของเมตริกซ์ ซึ่งจะถูกกำหนดให้เป็นค่าตามต้องการ (0 หรือ 1) แต่ถ้าข้อมูลในเมตริกซ์ไม่เหมือนกับข้อมูลภาพ ข้อมูลเอ้าพุทที่ได้จะมีค่าตรงกันข้าม

2.8.2 โอเปอเรชั่นพื้นฐานสำหรับรูปร่างหรือโครงสร้างพื้นฐาน

พิจารณาข้อมูลภาพจะเป็นลักษณะดังนี้

1 * 1 * 1
* 1 * 1 *
1 * 1 * 1

ภาพประกอบที่ 2-17 ข้อมูลภาพ

จากข้อมูลภาพตามภาพประกอบที่ 2-17 สามารถแทนด้วยเซตในทางคณิตศาสตร์ได้ดังนี้คือ {(0,0),(0,2), (0,4), (1,1), (1,3), (2,0), (2,2), (2,4)}

เนื่องจากเราสามารถแทนลักษณะภาพได้ดังภาพประกอบที่ 2-17 ดังนั้นเราสามารถกำหนดให้มีข้อมูลภาพสำหรับการทำโอเปอเรชั่นได้ดังนี้คือ

A=
1 * 1 * 1
* 1 * 1 *
1 * 1 * 1

B=
* * * 1 1
* * * 1 1
* * * 1 1

A ยูเนียน B =
1 * 1 1 1
* 1 * 1 1
1 * 1 1 1

A อินเตอร์เซคชั่น B =
* * * * 1
* * * 1 *
* * * * 1

ข้อมูลภาพดังภาพประกอบที่ 2-18 แสดงถึงพิกเซลที่เราทราบ (ค่าเท่ากับ 1) และค่าที่เราไม่ทราบ (แสดงด้วย *) ชุดของข้อมูลภาพจะขยายออกไปทางด้านบน ล่าง ซ้าย ขวา แบบไม่จำกัด ข้อมูลภาพตามรูปที่ 2-17 สามารถเขียนได้ดังนี้

```
* * * * * * * * * * *
* * * * * * * * * * *
* * * * * * * * * * *
* * * 1 * 1 * 1 * * *
* * * * 1 * 1 * * * *
* * * 1 * 1 * 1 * * *
* * * * * * * * * * *
* * * * * * * * * * *
* * * * * * * * * * *
```

ภาพประกอบที่ 2-18 ข้อมูลภาพ

วงกลมที่ล้อมรอบพิกเซล 1 ตามภาพประกอบที่ 2-18 แสดงตำแหน่งเริ่มต้นของภาพ (Original)

การขยายภาพ (Dilate)

การขยายภาพในที่นี้จะพิจารณาสำหรับข้อมูลที่เป็นแบบไบนารี่โดยการใช้เทคนิคการ Hit และ Miss การขยายภาพจะทำได้โดยกำหนด Template (ซึ่งสามารถสร้างได้จาก * และ 1 โดยมีจุดเริ่มต้นที่กำหนดโดยวงกลม) และนำ Template นี้สแกนไปบนข้อมูลภาพตามลำดับตลอดทั้งภาพ ซึ่งในขณะที่จุดเริ่มต้น (Origin) ของ Template ตรงกับตำแหน่งข้อมูลที่พิกเซลมีค่าเท่ากับ 1 นั้นก็จะทำการยูเนียน Template นี้เข้ากับข้อมูลภาพดังตัวอย่าง

ตัวอย่าง ข้อมูลภาพ Template

```
* * * * * * 1 * * 1 *
* * * * * * 1 * * * 1
* * * * * 1 1 * 1 1 *
* * * * 1 1 1 1 1 1 1        1 *
* * * * 1 1 1 1 1 * 1        1 1
* * * * 1 1 1 1 1 1 1
* * * * 1 1 1 1 1 1 1
```

ข้อมูลแถวแรกของภาพเป็นดังนี้

* * * * * * 1 * * 1 *

เมื่อทำการยูเนียนกับ Template ณ ตำแหน่งข้อมูลภาพที่พิกเซลเท่ากับ 1 ในแถวแรก

* * * * * * 1 * * * *
* * * * * * 1 1 * 1 1

และเมื่อยูเนียนกับ Template เข้ากับพิกเซลที่มีค่าเท่ากับ 1 ณ ตำแหน่งพิกเซลที่สองในแถวแรก

* * * * * * 1 * * * *
* * * * * * 1 1 * 1 1

และเมื่อทำการยูเนียนทั้งภาพจะได้ภาพสุดท้ายดังนี้

* * * * * * 1 * * 1 * *
* * * * * * 1 1 * 1 1 *
* * * * * 1 1 1 1 1 1 1
* * * * 1 1 1 1 1 1 1 1
* * * * 1 1 1 1 1 1 1 1
* * * * 1 1 1 1 1 1 1 1
* * * * 1 1 1 1 1 1 1 1
* * * * 1 1 1 1 1 1 1 1

การย่อภาพ (Erosion)

การย่อภาพเป็นลักษณะของการลบข้อมูลภาพบริเวณขอบของภาพ การย่อภาพมีลักษณะคล้ายกับการขยายภาพโดยการสร้าง Template ขึ้นแล้วนำ Template ไปสแกนตามข้อมูลภาพ

สำหรับทุกตำแหน่งที่เลื่อน Template ไปบนภาพก็จะมีการเปรียบเทียบกับข้อมูลภาพ ถ้าข้อมูลภาพมีค่าเหมือนกับ Template จะทำการกำหนดค่าข้อมูลในตำแหน่งที่ตรงกับจุดเริ่มต้น (Origin) ของ Template ถูกกำหนดให้มีค่าเท่ากับ 1

ตัวอย่าง ข้อมูลภาพ Template

*	*	*	*	*	*	1	*	*	1	*
*	*	*	*	*	*	1	*	*	*	1
*	*	*	*	*	1	1	*	1	1	*
*	*	*	*	1	1	1	1	1	1	1
*	*	*	*	1	1	1	1	1	*	1
*	*	*	*	1	1	1	1	1	1	1
*	*	*	*	1	1	1	1	1	1	1

1	*
1	1

ผลที่ได้จะมีเพียง 3 ตำแหน่งเท่านั้นที่มีค่าเหมือนกับ Template

*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	1	*	*	1	*
*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	1	*	*
*	*	*	*	*	*	*	*	*	*	*

ภาพประกอบที่ 2-19 แสดงผลการย่อภาพ

ผลที่ได้ตามภาพประกอบที่ 2-19 เป็นข้อมูลภาพที่ผ่านการทำโอเปอเรชั่นกับ Template แล้วพบว่ามีข้อมูลของภาพเพียง 3 ตำแหน่งเท่านั้นที่เหมือนกับ Template ถ้ามีการเปลี่ยน Template เป็น $\begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix}$ ผลที่ได้จะมีลักษณะดังนี้คือ

*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	1	1	*	1	*	*
*	*	*	*	1	1	1	1	1	*	*
*	*	*	*	1	1	1	1	*	*	*
*	*	*	*	1	1	1	1	1	1	*

ผลที่ได้จะเป็นการย่อขนาดของภาพและสามารถย่อได้น้อยกว่านี้เมื่อใช้ $\begin{matrix}1 & *\\ 1 & 1\end{matrix}$ Template ซึ่งได้ผลเป็นที่น่ายอมรับมากกว่า ดังนั้นในการเลือก Template เป็นสิ่งที่สำคัญอย่างหนึ่งในการย่อและขยายภาพ

โอเปอเรชั่นการ Opening

กำหนดให้ OPEN(I,T) เป็นการกระทำ Opening ของภาพ I โดยใช้ Template T ซึ่งมีลักษณะดังสมการต่อไปนี้

$$OPEN(I,T) = D(E(I))$$

จากสมการจะได้ว่าการทำโอเปอเรชั่น OPEN คือการนำข้อมูลภาพ I ผ่านการทำขยายภาพ (Dilation) แล้วตามด้วยการย่อภาพ (Erosion) โดยใช้ Template ชุดเดียวกันคือ T

ตัวอย่าง Original								Opened With							
													1		
													1		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	4	0	0	0	0	0	0	1	4
0	0	0	0	1	2	4	5	0	0	0	0	1	2	4	4
0	0	1	1	2	5	5	4	0	0	0	0	1	2	4	4
0	0	0	0	0	1	2	3	0	0	0	0	0	1	2	3
0	0	0	0	0	0	1	3	0	0	0	0	0	0	1	3
0	0	0	0	0	1	2	3	0	0	0	0	0	1	2	3
0	0	0	0	1	2	2	3	0	0	0	0	0	1	2	3

โอเปอเรชั่นการ Closing

กำหนดให้ CLOSE(I,T) เป็นการกระทำแบบ Closing ของภาพ I โดยใช้ Template T ซึ่งมีลักษณะดังสมการต่อไปนี้

$$CLOSE(I,T) = E(D(I))$$

จากสมการจะได้ว่าการทำโอเปอเรชั่น CLOSE คือการนำข้อมูลภาพ I ผ่านการทำการย่อภาพ (Erosion) แล้วตามด้วยการขยายภาพ (Dilation) โดยใช้ Template ชุดเดียวกันคือ T

ตัวอย่าง	Original								Closed With							
														1		
														1		
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	1	4	0	0	0	0	0	0	1	4
	0	0	0	0	1	2	4	5	0	0	0	0	1	2	4	5
	0	0	1	1	2	5	5	4	0	0	1	1	2	5	5	4
	0	0	0	0	0	1	2	3	0	0	0	0	0	1	2	3
	0	0	0	0	0	0	1	3	0	0	0	0	0	1	2	3
	0	0	0	0	0	1	2	3	0	0	0	0	0	1	2	3
	0	0	0	0	1	2	2	3	0	0	0	0	1	2	2	3

2.9 การแยกข้อมูลภาพ (Image Segmentation)

การทำการ Segmentation จะทำให้สามารถแยกข้อมูลภาพของส่วนที่ต้องการออกมาได้ วิธีการพื้นฐานสำหรับการ Segmentation คือการพิจารณา Image amplitude (ได้แก่การพิจารณาความสว่างของภาพสำหรับภาพแบบ Gray scale และความต่างกันของสีสำหรับภาพสี) นอกจากนี้ขอบของภาพและลักษณะของ Texture ก็เป็นองค์ประกอบหนึ่งที่จะทำให้สามารถทำการ Segmentation ได้สะดวกขึ้น

2.9.1 การแยกข้อมูลภาพด้วยการพิจารณาความสว่างของภาพ

เป็นการ Segmentation โดยการพิจารณาความเข้มของจุดต่าง ๆ ภายในจุดภาพ (pixel) ซึ่งผลของการ Segment จะขึ้นอยู่กับวิธีการ Threshold ของส่วนประกอบที่เป็นความเข้มหรือสีของภาพ

2.9.2 พื้นฐานในการแบ่งกลุ่มของระดับความเข้ม

การแบ่งกลุ่มระดับความเข้มจะแบ่งออกเป็น 2 กลุ่ม นั่นคือเป็นการแปลงจากภาพที่เป็นระดับความเข้ม (Gray Scale Image) ให้เป็นภาพ 2 ระดับ (bilevel : โมโนโครมหรือขาวกับดำ) ซึ่งภาพ 2 ระดับนี้สามารถบอกเราได้ถึง จำนวน ตำแหน่ง และรูปร่างของวัตถุต่าง ๆ ที่อยู่ในภาพ แล้วจากนั้นเราจะนำข้อมูลเหล่านี้ไปใช้ในการแบ่งกลุ่มของระดับความเข้มในภาพออกเป็นกลุ่มมากขึ้นได้ โดยมีการสมมติเบื้องต้นว่า จุดภาพที่มีระดับความเข้มใกล้เคียงกัน น่าจะแสดงถึงขอบเขตของวัตถุที่เป็นชนิดเดียวกัน

ทฤษฎีพื้นฐานที่ใช้ในการแปลงภาพจากภาพที่เป็นระดับความเข้ม (Gray Scale) ให้เป็นภาพ 2 ระดับ (bilevel) ทำได้โดยการกำหนดระดับความเข้มขึ้นมาระดับหนึ่ง และกำหนดให้ทุกจุด

ภาพในภาพที่มีค่าระดับความเข้มต่ำกว่าค่าระดับความเข้มที่ตั้งไว้ ให้จุดภาพนั้นเป็นสีดำ (ระดับความเข้ม 0) และกำหนดให้ทุกจุดภาพในภาพที่มีระดับความเข้มมากกว่าหรือเท่ากับค่าของระดับความเข้มที่ตั้งไว้ ให้จุดภาพนั้นเป็นสีขาว (ระดับความเข้มเป็น 255) ทำให้เราสามารถแปลงจากภาพที่เป็น Gray level ให้เป็น bilevel ได้ และเราเรียกค่าของระดับความเข้มที่กำหนดขึ้นนี้ว่าค่า Threshold T นั่นคือ

ถ้ากำหนดให้กลุ่มแรกเป็นกลุ่มจุดภาพที่มีสีดำ เราสามารถกำหนดกลุ่มนี้ได้จาก

$$I(i,j)<T \tag{2-10}$$

และเราสามารถกำหนดกลุ่มของจุดภาพที่อยู่นอกเหนือกลุ่มแรกเป็นกลุ่มของจุดภาพที่เป็นสีขาวโดย

$$I(I,j)\geq T \tag{2-11}$$

การแปลงภาพที่เป็น Gray level ให้เป็น bilevel ในลักษณะนี้เราเรียกว่าการกำหนดค่า Single Threshold คือเป็นการกำหนดค่าระดับความเข้มให้เป็นค่า Threshold เพียงหนึ่งค่า ซึ่งทำให้เราสามารถแยกภาพในส่วนของวัตถุออกจากส่วนของพื้นหลังได้ และหลังจากนั้นเราสามารถนำทฤษฎีพื้นฐานมาใช้ในการประยุกต์เพื่อให้สามารถแบ่งกลุ่มของระดับความเข้มออกเป็นหลาย ๆ กลุ่มต่อไป

ดังนั้นเราจึงต้องเริ่มจากการหาค่าของ Threshold ที่เหมาะสมออกมาให้ได้ก่อน ซึ่งค่า Threshold ที่เหมาะสมจะต้องเป็นค่าที่ทำให้เราสามารถแยกส่วนที่เป็นวัตถุและส่วนที่เป็นพื้นหลังของภาพออกจากกันได้อย่างสมบูรณ์ที่สุด วิธีในการกำหนดค่า Threshold มีหลายวิธีดังนี้

1. การกำหนดค่า Threshold จากค่าเฉลี่ย (mean) ของระดับความเข้มทั้งหมดในภาพ

เป็นการนำค่าเฉลี่ยของระดับความเข้มของจุดภาพทั้งหมดมากำหนดเป็นค่า Threshold ซึ่งหมายถึงการกำหนดให้จำนวนจุดภาพในภาพ ครึ่งหนึ่งเป็นสีดำและอีกครึ่งหนึ่งเป็นสีขาว

เราสามารถกำหนดค่า Threshold ได้จาก

$$T = \frac{\sum_{i=1}^{n} g[i]}{n} \tag{2-12}$$

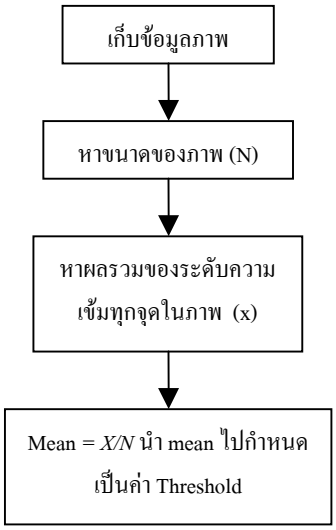
กำหนดให้

T คือ ค่า Threshold

g[i] คือ ค่าระดับความเข้มของจุดภาพ i

n คือ จำนวนจุดภาพทั้งหมด

แสดงขั้นตอนในรูปของ Flowchart ได้ดังนี้



ภาพประกอบที่ 2-20 แสดง Flowchart ของการกำหนดค่า Threshold จากค่า mean ของระดับความเข้มทั้งหมดในภาพ

พบว่าการกำหนดค่า Threshold จากค่าเฉลี่ยของระดับเทาแบบนี้ไม่เหมาะสมนัก เนื่องจากมีภาพจำนวนน้อยที่จะมีจำนวนจุดภาพที่เป็นสีดำอยู่ครึ่งหนึ่งหรือก็คือมีจำนวนจุดภาพที่เป็นส่วนของวัตถุเท่ากับจำนวนที่เป็นส่วนของพื้นหลัง

2. การกำหนดค่า Threshold โดยการกำหนดเปอร์เซ็นต์ของจุดภาพที่เป็นสีดำ

โดยส่วนใหญ่แล้วจำนวนของจุดภาพที่เป็นสีขาวกับสีดำจะเป็นอัตราส่วนกัน ดังนั้นเราสามารถหาค่า Threshold ได้โดยการกำหนดเปอร์เซ็นต์ของจำนวนจุดภาพที่เป็นสีดำตามความเหมาะสม ขั้นตอนในการกำหนดค่า Threshold โดยวิธีนี้ทำได้ดังนี้

ขั้นที่ 1 หาจำนวนจุดภาพที่เป็นสีดำ ตามเปอร์เซ็นต์ที่ต้องการ

$$M = Nx(per) \qquad (2\text{-}13)$$

เมื่อ

M คือ จำนวนจุดภาพที่ต้องการให้เป็นสีดำ

N คือ จำนวนจุดภาพทั้งหมด

Per คือ เปอร์เซ็นต์ของจุดภาพสีดำในภาพ

ขั้นที่ 2 หาค่าของจำนวนจุดภาพสะสมตั้งแต่ระดับความเข้มที่ 0 จนกระทั่งถึงระดับความเข้มที่ทำให้ผลรวมของจำนวนจุดภาพสะสม มากกว่าหรือเท่ากับจำนวนจุดภาพที่ต้องการให้เป็นสีดำ

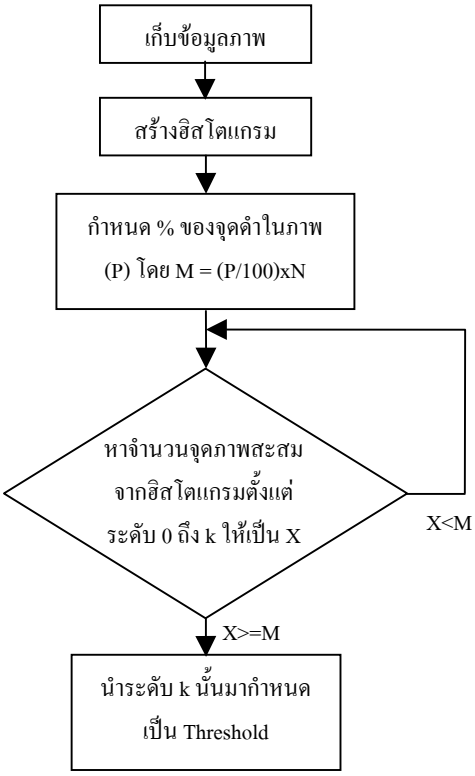
$$\text{if} \quad \sum_{i=0}^{255} h[i] \geq M \qquad ; \text{Then } T = i \qquad (2\text{-}14)$$

เมื่อ

h[i] คือ จำนวนจุดภาพที่มีระดับความเข้มเป็น i

นั่นคือ ที่ระดับความเข้มที่ทำให้เงื่อนไขเป็นจริง เราจะนำค่าที่ระดับความเข้มนั้นมาเป็นค่า threshold

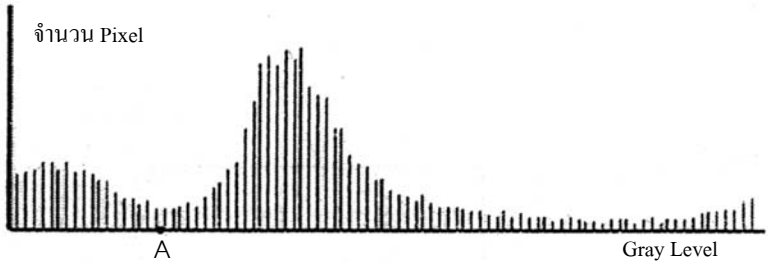
แสดงขั้นตอนในรูปของ Flowchart ได้ดังนี้



ภาพประกอบที่ 2-21 แสดง Flowchart ของการกำหนดค่า Threshold โดยการกำหนดเปอร์เซ็นต์ของจุดภาพที่เป็นสีดำ

3. การกำหนดค่า Threshold โดยใช้จุดต่ำสุดของจุดสูงสุด 2 จุด ในฮิสโตแกรม (Twopeaks Method)

ฮิสโตแกรม (Histogram) คือ กราฟที่แสดงความสัมพันธ์ระหว่างจำนวนของจุดภาพที่ระดับความเข้มต่าง ๆ นั่นคือ จะบอกถึงว่า ที่ระดับความเข้มต่าง ๆ มีจำนวนจุดภาพกี่จุดในภาพที่เป็นระดับความเข้มเท่านั้น โดยการกำหนดค่าของ Threshold โดยวิธีนี้ใช้ได้ดีกับภาพที่มีลักษณะของฮิสโตแกรมเป็นดังภาพประกอบที่ 2-22



ภาพประกอบที่ 2-22 แสดงฮิสโตแกรมที่มีลักษณะเป็น Twopeaks

นั่นคือ ภาพที่ให้ลักษณะฮิสโตแกรมเป็นลักษณะที่มีกลุ่มของข้อมูลที่สูงสุดอยู่ 2 กลุ่ม เราสามารถกำหนดให้ระดับเทาที่ต่ำสุดในช่วงของจุดยอดทั้งสองเป็นค่า Threshold ซึ่งขั้นตอนในการหาค่า Threshold โดยวิธีนี้อธิบายได้ดังนี้

ขั้นที่ 1 หาฮิสโตแกรม

ขั้นที่ 2 หาระดับความเข้มที่มีจำนวนจุดอยู่มากที่สุด

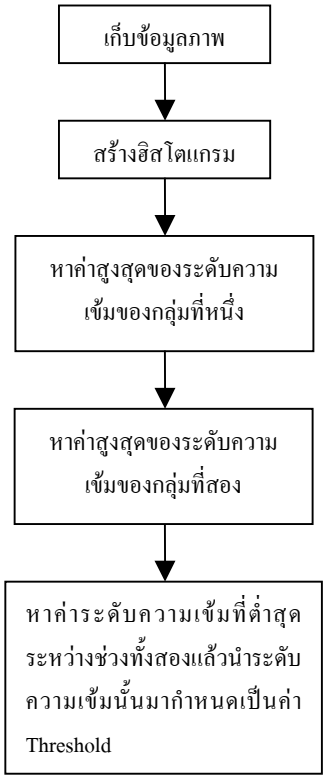
ขั้นที่ 3 หาระดับความเข้มที่มีจำนวนจุดสูงสุดเป็นอันดับที่สองจาก

$$\max\{((k-j)^2 h[k])\} \quad ;(0\leq k\leq 255) \qquad (2\text{-}15)$$

ขั้นที่ 4 หาระดับเทาที่ต่ำสุดในช่วงระหว่างจุดสูงสุดทั้งสอง แล้วนำระดับเทานั้นมาเป็น Threshold

พบว่าการกำหนดค่า Threshold โดยวิธีนี้ไม่สามารถใช้ได้กับภาพที่มีลักษณะของฮิสโตแกรม ที่มีจำนวน peaks มาก ๆ ได้ แต่จะใช้เป็นการกำหนด Threshold อย่างง่ายได้

แสดงเป็นขั้นตอนในรูปของ Flowchart ได้ดังนี้



ภาพประกอบที่ 2-23 แสดง Flowchart ของการกำหนดค่า Threshold โดยใช้จุดต่ำสุดของจุดสูงสุด 2 จุด ในฮิสโตแกรม

4. การกำหนดค่า Threshold โดยวิธี Iterative Selection

จากวิธีในการกำหนดค่า Threshold จากค่าเฉลี่ยของระดับความเข้มในภาพ พบว่าค่า Threshold ที่ได้ไม่เหมาะสมกับการนำไปใช้กับภาพส่วนใหญ่ ดังนั้นเราสามารถทำการปรับปรุงการกำหนดค่า โดยจากเดิมที่เรานำค่าเฉลี่ยของระดับความเข้มของภาพมาเป็นค่า Threshold เลย แต่ในวิธีนี้มีขั้นตอนการทำดังนี้

ขั้นที่ 1 หาค่าเฉลี่ยของระดับความเข้มของภาพ จากสมการที่ (2-12) ให้เป็นระดับ k

ขั้นที่ 2 จากนั้นทำการหาค่าเฉลี่ยของระดับความเข้มในช่วงที่ต่ำกว่าระดับ k จาก

$$T_b = \frac{\sum_{i=0}^{k} i.h[i]}{\sum_{i=0}^{k} h[i]} \qquad (2\text{-}16)$$

เมื่อ

T_b คือ ค่าเฉลี่ยของระดับความเข้มในช่วงที่ต่ำกว่าระดับ k

i คือ ระดับความเข้มที่ i

h[i] คือ จำนวนจุดภาพที่ระดับความเข้มเป็น i

ขั้นที่ 3 หาค่าเฉลี่ยของระดับความเข้มในช่วงที่สูงว่าระดับ k (To) จาก

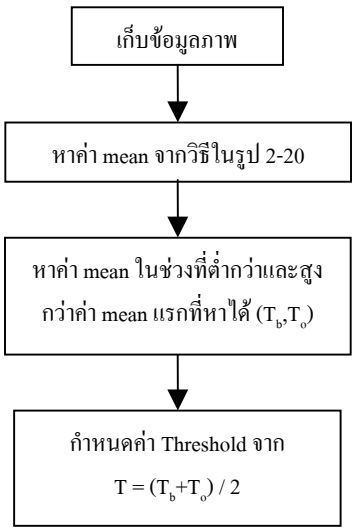
$$T_o = \frac{\sum_{i=k+1}^{255} i.h[i]}{\sum_{i=k+1}^{255} h[i]} \qquad (2\text{-}17)$$

ขั้นที่ 4 จากนั้นนำค่า T_b และ T_o มากำหนดค่า Threshold จาก

$$T = (T_b+T_o) / 2 \qquad (2\text{-}18)$$

นั่นก็คือ เราไม่นำค่าเฉลี่ยที่หาได้จากทั้งภาพมาเป็นค่า Threshold เลยแต่เราจะนำค่าเฉลี่ยของระดับความเข้มในช่วงที่ต่ำกว่าและสูงกว่ามาทำการหาค่าเฉลี่ยอีกครั้ง แล้วจึงนำค่าเฉลี่ยนั้นมากำหนดเป็น Threshold

แสดงขั้นตอนในรูปของ Flowchart ได้ดังนี้



ภาพประกอบที่ 2-24 แสดง Flowchart ของการกำหนดค่า Threshold โดยวิธี Iterative Selection

2.10 การหาขอบภาพ (Edge Detection)

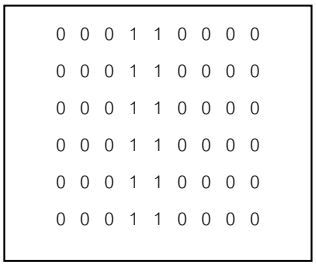
การหาขอบภาพเป็นการหาขอบเขตของวัตถุภายในภาพซึ่งขอบเขตของวัตถุจะเป็นที่เด่น และมีความสำคัญมากที่สุดที่จะนำไปสู่การรู้จักวัตถุนั้น ๆ โดยคอมพิวเตอร์จะเห็นได้ว่าการหาขอบภาพเป็นการหาจุดเด่นของวัตถุนั่นเอง ลักษณะที่เด่นของวัตถุที่เรามองเห็นโดยทั่วไปก็คือ ส่วนที่เป็นสัน หรือส่วนที่เป็นขอบของวัตถุ เมื่อมีแสงมากระทบก็จะมีความสว่างมากกว่าส่วนอื่น โดยทั่วไปลักษณะของขอบภาพที่ดีจะต้องมีลักษณะดังต่อไปนี้

- มีความบางคือ จุดที่เป็นขอบภาพจะต้องมีความกว้างเพียงจุดเดียวเท่านั้น

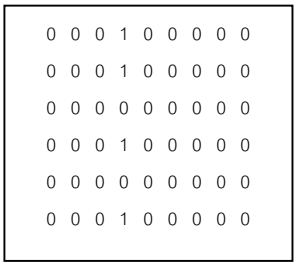
- มีความต่อเนื่อง โดยจุดของขอบภาพในวัตถุเดียวกันมีความต่อเนื่องกัน ถ้าจุดที่เป็นขอบภาพมีอยู่เพียงจุดเดียว โดยไม่ต่อเนื่องกับจุดใดในย่านใกล้เคียงเลยอาจจะเป็นขอบภาพที่ไม่สมบูรณ์ หรืออาจจะเป็นสัญญาณรบกวนได้

- ตำแหน่งของจุดที่เป็นขอบภาพควรมีความถูกต้องโดยขอบที่มีความถูกต้องนั้นจะต้องอยู่ในส่วนที่มีค่าความเข้มของจุดภาพที่สูงกว่า

จากภาพประกอบที่ 2-25 ได้แสดงขอบภาพในลักษณะต่างๆ โดยภาพประกอบที่ 2-25 a. แสดงขอบภาพที่มีความหนา และภาพประกอบที่ 2-25 b. เป็นขอบภาพที่ไม่ต่อเนื่อง



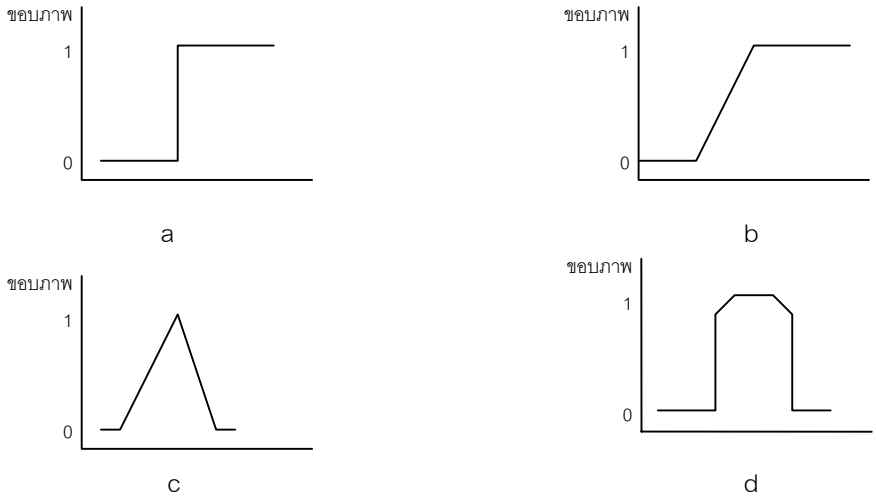
(a) ขอบภาพหนา (กว้าง 2 จุด)



(b) ขอบภาพที่ไม่ต่อเนื่อง

ภาพประกอบที่ 2-25 แสดงลักษณะขอบภาพแบบต่างๆ

รูปแบบของขอบภาพที่ปรากฏในรูปภาพโดยทั่วไปจะประกอบไปด้วยขอบภาพหลายชนิด ดังแสดงในภาพประกอบที่ 2-26 a. เป็นขอบภาพที่เปลี่ยนแปลงทันทีทันใดขอบภาพแบบนี้ส่วนใหญ่เป็นภาพที่มนุษย์จำลองขึ้นมา ส่วนภาพโดยทั่วไปจะเป็นขอบภาพที่มีการเปลี่ยนแปลงความเข้มทีละน้อย ดังภาพประกอบที่ 2-26 b. ภาพประกอบที่ 2-26 c. ขอบภาพที่ค่าความเข้มที่ค่าความเข้มค่อย ๆ เพิ่มขึ้น หรือทดลองทีละน้อยในลักษณะของรูปหน้าจั่ว และภาพประกอบที่ 2-26 d. แสดงขอบภาพที่เป็นสัน



ภาพประกอบที่ 2-26 แสดงขอบภาพชนิดต่างๆ

การเปลี่ยนแปลงค่าความเข้มแสงมีผลต่อการหาขอบภาพ หรือลักษณะทางกายภาพตัววัตถุเอง (เช่น ลักษณะของพื้นผิว และรูปทรง) คุณสมบัติเหล่านี้จะเป็นปัจจัยพื้นฐานที่สำคัญของรูปภาพที่จะบอกขอบเขตทางกายภาพของวัตถุที่อยู่ในภาพ

วิธีการเบื้องต้นที่ใช้สำหรับการหาขอบภาพในลำดับแรก (First order Derivative Edge Detection) มีอยู่ 2 วิธี ด้วยกัน วิธีแรกคือการหาเกรเดียนท์ (gradient) ในทิศทางมุมฉากคือตามแนวนอนและแนวตั้ง และอีกวิธีเราจะทำการหาเกรเดียนท์ในหลาย ๆ ทิศทาง

การสร้างเกรเดียนท์มุมฉาก (Orthogonal Gradient Generation)

ถ้ากำหนดให้ G(x,y) คือเกรเดียนท์ใน 1 มิติ และ θ คือมุมที่พิจารณาโดยเทียบกับแนวระดับ ดังนั้นจะทำการคำนวณเกรเดียนท์สำหรับของภาพจาก

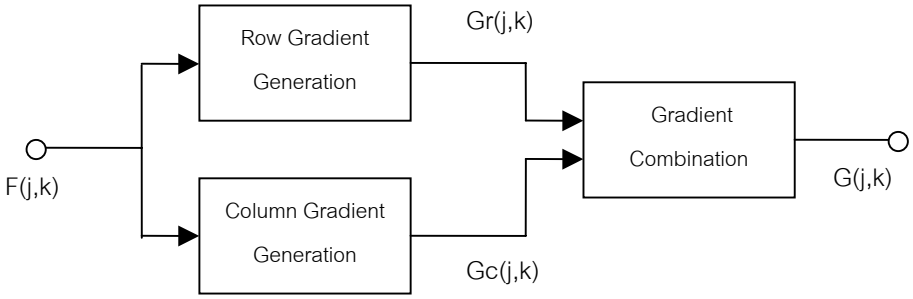
$$G(x, y) = \frac{\partial F(x, y)}{\partial x}\cos\ \theta + \frac{\partial F(x, y)}{\partial y}\sin\theta \qquad (2\text{-}19)$$

ในภาพประกอบที่ 2-27 แสดงการสร้างเกรเดียนท์ของขอบภาพ {G(j,k)} ในขอบเขตที่ไม่ต่อเนื่อง โดยจะแยกเป็นเกรเดียนท์ของแถว (row edge gradient) {Gr(j,k)} และเกรเดียนท์ของหลัก (Column edge gradient) {Gc(j,k)} และค่าแอมปลิจูดของเกรเดียนท์ ซึ่งหมายถึงความเข้มของภาพแสดงได้ดังนี้

$$G(j,k) = \{[Gr(j,k)]+[Gc(j,k)]\}^{1/2} \quad (2\text{-}20)$$

การคำนวณค่าเกรเดียนท์บางครั้งอาจประมาณค่าจาก

$$G(j,k) = |Gr(j,k)| + |Gc(j,k)| \quad (2\text{-}21)$$



ภาพประกอบที่ 2-27 แสดงการหาเกรเดียนท์มุมฉาก

และทิศทางของเกรเดียนท์ตามแนวนอนคือ

$$\theta(j,\ k) = \tan^{-1}\left\{\frac{Gc(j,k)}{Gr(j,k)}\right\} \quad (2\text{-}22)$$

ซึ่งวิธีการที่ง่ายที่สุดสำหรับการสร้างเกรเดียนท์คือ การหาค่าความแตกต่างของความเข้มในแต่ละจุดตามแนวแกนนอน และตามแนวแกนตั้ง ซึ่งเกรเดียนท์ตามแนวแกนนอนหาได้จาก

$$Gr(j,k) = F(j,k) - F(j,k-1) \quad (2\text{-}23)$$

และเกรเดียนท์ในแนวแกนตั้งหาได้จาก

$$Gc(j,k) = F(j,k) - F(j+1,k) \quad (2\text{-}24)$$

โดยเกรเดียนท์นี้จะนำมาใช้หาขอบภาพก็ต่อเมื่อ Gc และ Gr มีค่าเป็นบวกจากซ้ายไปขวาหรือล่างขึ้นบนของภาพอิมเมจ

สำหรับเกรเดียนท์ของขอบในแนวทแยง (diagonal edge gradeint) จะหาได้จากผลต่างของจุดแต่ละคู่ตามแนวทแยง ซึ่งจะใช้วิธีการที่เรียกว่า โรเบิร์ตดิฟเฟอเรนซ์โอเปอร์เตอร์ (Roberts

Difference operator)

2.10.1 โรเบิร์ตดิฟเฟอเรนซ์โอเปอเรเตอร์ (Roberts Difference operator)

สำหรับโอเปอเรเตอร์ชนิดนี้หาได้จาก

$$G(j,k) = |G1(j,k)| + |G2(j,k)| \tag{2-25}$$

หรืออาจหาได้จาก

$$G(j,k) = \{[Gr(j,k)]^2 + [Gc(j,k)]^2\}^{1/2} \tag{2-26}$$

เมื่อ $$G1(j,k) = F(j,k) - F(j+1,k+1) \tag{2-27}$$

$$G2(j,k) = F(j,k+1) - F(j+1,k) \tag{2-28}$$

และทิศทางของขอบภาพคือ

$$\theta(j,\ k) = \frac{\pi}{4} + \tan^{-1}\left\{\frac{Gc(j,\ k)}{Gr(j,\ k)}\right\} \tag{2-29}$$

การหาขอบภาพในทิศทางเดียวเช่นนี้จะไม่สมบูรณ์นักสำหรับภาพอิมเมจที่มีความส่องสว่างเพียงเล็กน้อย ซึ่งการแก้ปัญหานี้ทำได้โดยการหาเกรเดียนท์ใน 2 มิติ คือทำการหาผลต่างใน 1 ทิศทางแต่จะแสดงค่าออกมาเป็นค่าเฉลี่ยทั้งหมดในทิศทางมุมฉาก

2.10.2 พรีวิทท์โอเปอเรเตอร์ (Prewitt operator)

A0	A1	A2
A7	F(j,k)	A3
A6	A5	A4

ภาพประกอบที่ 2-28 ลักษณะของ Template ที่มีขนาด 3x3

โอเปอเรเตอร์ชนิดนี้จะใช้ Template ที่มีขนาด 3x3 ในการหาเกรเดียนท์โอเปอเรเตอร์ ถ้าสมมติให้แต่ละจุดมีการเรียงลำดับตามรูปที่ ดังนั้นเกรเดียนท์ของขอบคือ

$$G(j,k) = \{[Gr(j,k)]^2 + [Gr(j,k)]^2\}^{1/2} \tag{2-30}$$

โดยที่

$$Gr(j,k) = [(A2+KA3+A4)-(A0+KA7+A6)] / (K+2) \quad (2\text{-}31)$$

$$Gc(j,k) = [(A0+KA1+A2)-(A6+KA5+A4)] / (K+2) \quad (2\text{-}32)$$

การหาโอเปอเรเตอร์แบบพรีวิทท์ค่า K ที่นำมาใช้มีค่าเท่ากับ 1 เมื่อแทนค่าลงไปในสูตรข้างต้นของเกรเดียนท์ในแนวนอนแต่ละครั้ง สุดท้ายจะได้ยูนิตเกน (unit gain) ที่มีค่าเป็นบวกและลบของค่าเฉลี่ยความเข้มเพื่อหาตำแหน่งของขอบภาพ

2.10.3 โซเบลโอเปอเรเตอร์ (Sobel operator)

การหาขอบภาพแบบโซเบลนี้จะแตกต่างจากแบบพรีวิทท์ตรงที่ค่าความเข้มของภาพในทิศเหนือ ใต้ ตะวันออก ตะวันตก จะมีค่าเพิ่มขึ้นอีกเท่าตัวในที่นี้คือค่า K ในสมการที่ และ มีค่าเท่ากับ 2 ดังนั้นค่าน้ำหนัก (weignt) ที่เปลี่ยนไปในแต่ละจุดคือตัวที่แสดงการกระจายของเกรเดียนท์

พรีวิทท์โอเปอเรเตอร์จะใช้สำหรับการหาขอบภาพในแนวตั้งและแนวนอนได้ดีกว่าโอเปอเรเตอร์แบบโซเบล แต่โซเบลโอเปอเรเตอร์จะใช้ในการหาขอบภาพในทิศทางแนวทแยงได้ดี การใช้โอเปอเรเตอร์แบบพรีวิทท์ และโซเบลจำทำให้ได้ขอบภาพของวัตถุที่มีความชัดเจนกว่า การใช้โรเบิร์ตโอเปอเรเตอร์ ทั้งนี้เพราะว่ามีขนาดของโอเปอเรเตอร์ที่ใหญ่กว่าทำให้การเฉลี่ยค่าความเข้มในส่วนที่มีความส่องสว่างน้อยได้ดีกว่าเกรเดียนท์ในแนวนอนและแนวตั้งของดีเทคเตอร์ชนิดต่าง ๆ ที่กล่าวมาข้างต้นจะประกอบด้วยการนำความเข้มรอบ ๆ จุดที่ต้องการหาค่ามารวมกัน ดังนั้นเกรเดียนท์ในแนวนอนและแนวตั้งจะคำนวณจากความสัมพันธ์ดังนี้คือ

$$Gr(j,k) = F(j,k) * Hr(j,k) \quad (2\text{-}33)$$

$$Gc(j,k) = F(j,k) * Hc(j,k) \quad (2\text{-}34)$$

โดยที่

$Hr(j,k)$ คืออิมพัลส์เรสปอนส์ในแนวแกนนอน

และ $Hc(j,k)$ คืออิมพัลส์เรสปอนส์ในแนวแกนตั้ง

ในรูปต่อไปนี้จะแสดงเกรเดียนท์ในแนวแกนนอนและแนวตั้งของโอเปอเรเตอร์ชนิดต่าง ๆ

Operator	Row gradient	Column gradient
Pixel difference	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
Roberts	$\begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
Prewitt	$1/3 \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$	$1/3 \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$
Sobel	$1/4 \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$	$1/4 \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$

ภาพประกอบที่ 2-29 แสดงอิมพัลส์เรสปอนส์ของเกรเดียนท์โอเปอเรเตอร์ในแนวมุมฉาก