



## บทที่ 3

# ทฤษฎีที่เกี่ยวข้อง

ในบทนี้จะอธิบายถึงทฤษฎีการประมวลผลภาพที่นำมาใช้ในการสร้างดัชนีภาพ ขั้นตอนวิธีการสร้างและการเปรียบเทียบดัชนีภาพแบบต่าง ๆ ที่เป็นพื้นฐานและเกี่ยวข้องในงานวิจัยนี้และวิธีการทดสอบประสิทธิภาพของแต่ละวิธีที่ใช้ในงานวิจัยนี้

### 3.1 ทฤษฎีการประมวลผลภาพ

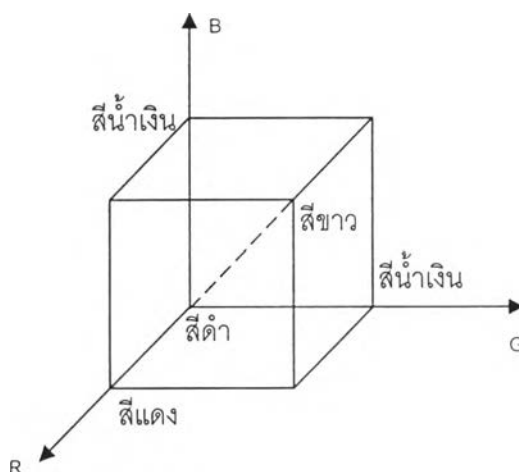
ในการสร้างดัชนีภาพจำเป็นต้องมีการใช้ทฤษฎีการประมวลผลภาพมาช่วยในการจำแนกลักษณะสำคัญของภาพ ซึ่งทฤษฎีการประมวลผลที่นำมาใช้ในงานวิจัยนี้ได้แก่ โมเดลสี และการควอนไทซ์สี โดยมีรายละเอียดดังต่อไปนี้

#### 3.1.1 โมเดลสี [16]

โมเดลสีเป็นการนิยามและแสดงความสัมพันธ์ของสีด้วยคณิตศาสตร์ โดยแต่ละสีจะถูกแทนด้วยจุดหนึ่งจุดใดภายในระบบพิกัดหรือปริภูมิใด ๆ [16] การเลือกใช้โมเดลสีใดนั้นขึ้นอยู่กับอุปกรณ์ที่มีการใช้สี (เช่น จอภาพสีหรือเครื่องพิมพ์สี) หรือขึ้นอยู่กับจุดประสงค์ที่ต้องการประยุกต์ใช้ ตัวอย่างของโมเดลสีที่ใช้กันบ่อยๆ เช่นโมเดลสี RGB (Red, Green Blue) ที่ใช้ในจอภาพสี (Color monitor) หรือกล้องถ่ายภาพดิจิทัล โมเดลสี CMY (Cyan, Magenta, Yellow) และ CMYK (Cyan, Magenta, Yellow, black) ที่ใช้ในเครื่องพิมพ์สี ส่วนโมเดลสี HSV (Hue, Saturation, Value) และโมเดลสี  $L^*a^*b^*$  เป็นโมเดลสีที่ใกล้เคียงกับการรับรู้ของมนุษย์ โดยแบ่งข้อสนเทศของภาพเป็นข้อสนเทศสีและข้อสนเทศความเข้มแสง สำหรับโมเดลสีที่ถูกนำมาใช้ในงานการงานวิจัยนี้ คือ โมเดลสี RGB

#### โมเดลสี RGB

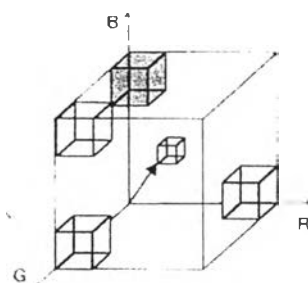
โมเดลสี RGB เป็นโมเดลสีพื้นฐานในงานคอมพิวเตอร์โดยแสดงเป็นลูกบาศก์ของสีบนระบบพิกัดแบบคาร์ทีเซียนที่แบ่งองค์ประกอบหลักออกเป็น สีแดง (Red, R) สีเขียว (Green, G) และสีน้ำเงิน (Blue, B) ดังรูปที่ 3.1 ซึ่งในแนวทแยงมุมหลักของลูกบาศก์ที่มีค่าสีแดง เขียว และน้ำเงินเท่ากันจะแสดงค่าเป็นค่าระดับเทา



รูปที่ 3.1 ลักษณะของโมเดลสี RGB

### 3.1.2 การควอนไทซ์สี (Color quantization)

ในภาพทั่วไปค่าสีจะถูกเก็บโดยใช้โมเดล RGB และใช้จำนวนบิตในการเก็บค่าสีทั้งหมด 24 บิต (สีแดง สีเขียวและสีน้ำเงิน อย่างละ 8 บิต) ทำให้สามารถแสดงสีที่แตกต่างกันได้ถึง 16,777,216 สี การสร้างดัชนีภาพที่มีการใช้สีเป็นดัชนีสำหรับการค้นคืนภาพ หากใช้สีเป็นดัชนีเป็นจำนวนเท่ากับจำนวนสีที่แตกต่างกันในภาพโดยตรงก็จะต้องใช้เนื้อที่เก็บข้อมูลมากและใช้เวลาในการคำนวณสูง ดังนั้นขั้นตอนสำคัญขั้นตอนหนึ่งในการสร้างดัชนีให้กับภาพสีคือการควอนไทซ์สีเพื่อลดจำนวนข้อมูลโดยการตัดข้อมูลบางข้อมูลทิ้งไปหรือแทนค่าของกลุ่มข้อมูลที่มีหลายค่าด้วยค่าเดียว การควอนไทซ์จากค่าของข้อมูลสีเรียกว่า Color-level reduction การควอนไทซ์สีทำได้หลายวิธี วิธีที่ง่ายและนิยมใช้คือการลดจำนวนบิตที่เก็บข้อมูลสีในแต่ละจุดภาพ (bits per pixel) โดยการจัดกลุ่มของค่าสีเข้าด้วยกันเป็นสีเดียว ลักษณะการจัดกลุ่มขึ้นอยู่กับโมเดลสีที่นำมาใช้ สำหรับการควอนไทซ์สีในงานวิจัยนี้ใช้วิธีการแบ่งโมเดลสีเป็นส่วน ๆ ขนาดเท่า ๆ กัน บนโมเดลสี RGB โดยใช้วิธีการจัดกลุ่มสีเป็นลูกบาศก์ย่อยดังรูปที่ 3.2



รูปที่ 3.2 การควอนไทซ์ข้อมูลสีในโมเดลสี RGB

### 3.2 การค้นคืนภาพ

การสร้างดัชนีภาพเป็นขั้นตอนสำคัญขั้นหนึ่งในการค้นคืนภาพโดยการนำลักษณะสำคัญของภาพมาสร้างเป็นดัชนีภาพและนำดัชนีภาพที่ได้ไปเก็บไว้ในฐานข้อมูลดัชนีภาพเพื่อนำไปใช้ในการเปรียบเทียบภาพที่ต้องการ สำหรับงานวิจัยนี้ได้ใช้สีเป็นลักษณะสำคัญของภาพในการสร้างดัชนีภาพซึ่งนำแนวทางมาจากการค้นคืนภาพโดยใช้ฮิสโทแกรมสีและการค้นคืนภาพโดยใช้คอริโลแกรมสีให้สามารถนำมาใช้ค้นคืนภาพจากบางส่วนของภาพได้ ซึ่งวิธีต่าง ๆ ที่นำมาใช้ในงานวิจัยนี้มีวิธีการสร้างดัชนีภาพและเปรียบเทียบดังนี้

เพื่อให้ง่ายและสะดวกในการอธิบายต่อไป จึงขอกำหนดให้  $I$  เป็นภาพในฐานข้อมูลภาพ  $Q$  เป็นภาพที่ใช้ในการสอบถาม และ  $p$  เป็นจุดภาพที่ตำแหน่งพิกัด  $(x, y)$  ใด ๆ ในภาพ

#### 3.2.1 การสร้างดัชนีภาพโดยใช้ฮิสโทแกรมสีและการเปรียบเทียบ

ฮิสโทแกรมสีเป็นลักษณะสำคัญของภาพที่เป็นการวัดการแจกแจงของสีในภาพ ฮิสโทแกรมสีถูกนำมาใช้บ่อย ๆ เนื่องจากคำนวณได้ง่าย สำหรับภาพที่จะถูกนำมาใช้ในการสร้างดัชนีแบบนี้จะต้องมีการควอนไทซ์สีก่อน ถ้ากำหนดให้จำนวนสีหลังจากการควอนไทซ์แล้วมีจำนวน  $m$  สีได้แก่สี  $c_1, \dots, c_m$  ดัชนีของภาพแบบฮิสโทแกรม  $H(I)$  จะมีลักษณะเป็นเวกเตอร์ที่มีขนาดเท่ากับจำนวนสีในภาพซึ่งประกอบด้วย  $H_{c_1}(I), H_{c_2}(I), \dots, H_{c_m}(I)$  โดยสมาชิกในเวกเตอร์แต่ละตัวเป็นจำนวนจุดภาพสีของแต่ละสีในภาพ โดยสามารถคำนวณหาฮิสโทแกรมสีที่สี  $c_i$  จากภาพ  $I$  หรือ  $H_{c_i}(I)$  ได้จากสมการที่ 3.1

$$H_{c_i}(I) = |I_{c_i}| \quad \dots(3.1)$$

โดย  $I_{c_i}$  คือ จุดภาพต่าง ๆ ที่มีสี  $c_i$  ในภาพ  $I$

จากจำนวนจุดภาพสีสามารถนำมาหาค่าความน่าจะเป็นที่จะพบจุดภาพสี  $c_i$  ในภาพ  $I$  ได้ดังสมการที่ 3.2

$$h_{c_i}(I) = \Pr[p \in I_{c_i}] = \frac{H_{c_i}(I)}{n} \quad \dots(3.2)$$

โดย  $n$  คือ จำนวนจุดภาพในภาพ  $I$

ข้อจำกัดสำคัญของฮิสโทแกรมสีคือข้อสนเทศที่แสดงถึงภาพแต่ละภาพไม่ความสัมพันธ์กับตำแหน่งของแต่ละจุดภาพสีในภาพ ทำให้ไม่สามารถบอกรายละเอียดเชิงพื้นที่ในแต่ละภาพได้

การเปรียบเทียบภาพซึ่งทำโดยใช้ฟังก์ชันระยะทางมาวัดความแตกต่างระหว่างฮิสโทแกรมนั้นไม่สามารถบอกได้ว่ามีภาพหนึ่งอยู่ภายในอีกภาพหนึ่งหรือไม่ ดังนั้น Swain และ Ballard [12] จึง

นำเอาทฤษฎีทางคณิตศาสตร์ในเรื่องของการอินเตอร์เซกต์เซตมาใช้ หากภาพในฐานข้อมูลภาพใดที่มีภาพสอบถาม ( $Q$ ) อยู่ภายใน ผลของการอินเตอร์เซกต์ระหว่างดัชนีของภาพนั้นกับดัชนีของภาพสอบถามควรจะได้ผลลัพธ์เป็นดัชนีของภาพสอบถามเช่นเดิม ( $Q \cap I \rightarrow Q$ ) ดังนั้น การเปรียบเทียบจึงทำได้โดยการนำผลลัพธ์จากการอินเตอร์เซกต์ดังกล่าวมาเปรียบเทียบกับภาพสอบถามโดยใช้ฟังก์ชันระยะทาง ผลลัพธ์ที่ได้จะบอกถึงโอกาสที่จะไม่มีภาพสอบถามอยู่ในภาพที่นำมาเปรียบเทียบกับนั้น ๆ

สำหรับวิธีการค้นหาภาพจากภาพสอบถามที่เป็นบางส่วนของภาพที่ใช้การดัชนีภาพแบบฮิสโทแกรมนี้เรียกว่าการอินเตอร์เซกต์ของฮิสโทแกรมสี (*color histogram intersection*) ทำโดยนำฮิสโทแกรมของภาพสอบถาม  $Q$  มาอินเตอร์เซกต์กับแต่ละภาพ  $I$  ในฐานข้อมูล ผลลัพธ์ของการอินเตอร์เซกต์ที่สี  $c$ , ระหว่างภาพ  $Q$  และภาพ  $I$  ( $H_c(Q \cap I)$ ) คือ จำนวนจุดภาพสีที่น้อยกว่าระหว่างจำนวนจุดภาพสีในภาพ  $Q$  และภาพ  $I$  ดังสมการที่ 3.3

$$H_c(Q \cap I) = \min\{H_c(Q), H_c(I)\} \quad \dots(3.3)$$

จากผลลัพธ์ของการอินเตอร์เซกต์ของฮิสโทแกรมสีสามารถนำมาทำให้อยู่บนบรรทัดฐานเดียวกันเพื่อใช้ในการเปรียบเทียบกับภาพสอบถามได้ ดังสมการที่ 3.4

$$h_c(Q \cap I) = \frac{H_c(Q \cap I)}{|Q|} \quad \dots(3.4)$$

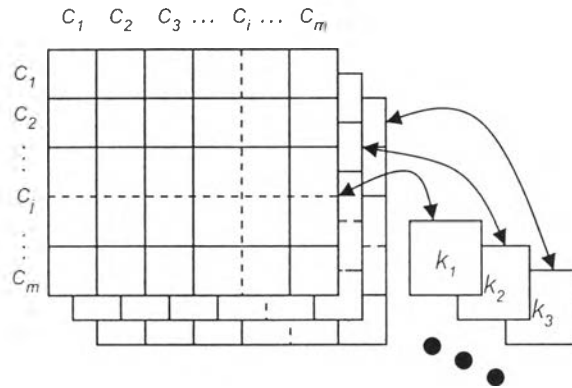
โดย  $|Q|$  คือ จำนวนจุดภาพของภาพสอบถาม

ดังนั้นโอกาสที่ภาพ  $Q$  บรรจอยู่ในภาพ  $I$  วัดได้จากความแตกต่างระหว่างฮิสโทแกรมของภาพที่ใช้สอบถามและฮิสโทแกรมที่ได้จากการอินเตอร์เซกต์โดยใช้ฟังก์ชันระยะทางซึ่งเขียนแทนได้เป็น  $|h(Q) - h(Q \cap I)|$  ฟังก์ชันระยะทางที่นิยมมาใช้ เช่น L1-norm (หรือเรียกได้ว่า City block) ซึ่งหากระยะทางที่ได้มีค่าเป็นศูนย์หมายถึงมีภาพ  $Q$  บรรจอยู่ในภาพ  $I$  และหากระยะทางที่ได้มีค่ามากขึ้นแสดงว่าโอกาสที่มีภาพ  $Q$  บรรจอยู่ในภาพ  $I$  น้อยลง

### 3.2.2 การสร้างดัชนีภาพโดยใช้คอร์ริโลแกรมสีและการเปรียบเทียบ

ในปี ค.ศ. 1997 Jing Huang ได้เสนอวิธีการสร้างดัชนีภาพแบบคอร์ริโลแกรมโดยพิจารณาจากจำนวนคู่สีในแต่ละระยะห่างเชิงพื้นที่แทนฮิสโทแกรมที่เป็นการนับจำนวนจุดภาพสีในภาพโดยไม่คำนึงถึงตำแหน่งเชิงพื้นที่ โดยทั่วไปแล้วคอร์ริโลแกรมคือกราฟ (หรือตาราง) ที่แสดงการเปลี่ยนแปลงความสัมพันธ์ในแต่ละระยะทาง สำหรับคอร์ริโลแกรมสีนั้นเป็นการแสดงการเปลี่ยนแปลงความสัมพันธ์ของสีในแต่ละระยะทางซึ่งพัฒนามาจากข้อจำกัดของฮิสโทแกรมสีที่เก็บการแจกแจงของสีในภาพโดยที่ไม่มีข้อสนเทศสหสัมพันธ์เชิงพื้นที่

การสร้างดัชนีภาพโดยใช้คอร์ริโลแกรมสีสามารถแสดงได้เป็นตารางแสดงจำนวนความสัมพันธ์ของคู่สี  $c_i, c_j$  ในแต่ละระยะห่าง  $k$  ดังตัวอย่างในรูปที่ 3.3



รูปที่ 3.3 ตัวอย่างของตารางแสดงจำนวนความสัมพันธ์แบบคอร์ริโลแกรม

จำนวนความสัมพันธ์ของแต่ละคู่สี  $c_i, c_j$  ในคอร์ริโลแกรมสีคำนวณได้จากการนับจำนวนคู่สี  $c_i, c_j$  ที่ระยะห่าง  $k$  ในภาพดังสมการ 3.5

$$\Gamma_{c_i, c_j}^{(k)}(I) = \left| \{ p_1 \in I_{c_i}, p_2 \in I_{c_j}, |p_1 - p_2| = k \} \right| \quad \dots(3.5)$$

โดย  $\Gamma_{c_i, c_j}^{(k)}(I)$  คือ จำนวนของคู่สี  $c_i, c_j$  ที่ระยะห่าง  $k$  ของภาพ  $I$   
 $k$  คือ ระยะห่างระหว่างคู่จุดภาพที่ต้องการพิจารณา

จากจำนวนคู่จุดภาพสีที่ระยะห่าง  $k$  สามารถนำมาหาเป็นความน่าจะเป็นที่จะพบคู่สี  $c_i, c_j$  ในภาพ  $I$  ได้ ซึ่งเขียนแทนได้ด้วย  $\gamma_{c_i, c_j}^{(k)}$  โดย

$$\gamma_{c_i, c_j}^{(k)} = \Pr[p_2 \in I_{c_j}, |p_1 - p_2| = k | p_1 \in I_{c_i}] = \frac{\Gamma_{c_i, c_j}^{(k)}(I)}{H_{c_i}(I) \cdot 8k} \quad \dots(3.6)$$

โดย  $H_{c_i}(I)$  คือ จำนวนของจุดภาพสี  $c_i$  ที่พบในภาพ  $I$  ดังแสดงในสมการที่ 3.1

วิธีหนึ่งที่ใช้ในการหาระยะห่างระหว่างคู่จุดภาพสี ( $k$ ) คือใช้ฟังก์ชันระยะทางแบบ  $L_\infty$ -norm คือ

$$|p_1 - p_2| = \max\{|x_1 - x_2|, |y_1 - y_2|\} \quad \dots(3.7)$$

โดย  $x$  และ  $y$  คือ พิกัดของจุดภาพสี  $p$  ที่ตำแหน่ง  $x$  และ  $y$  ตามลำดับ

แม้ว่าดัชนีแบบคอรีโลแกรมสามารถใช้แสดงความแตกต่างของภาพได้ละเอียด แต่ขนาดของคอรีโลแกรมค่อนข้างใหญ่เพราะพิจารณาทุก ๆ ความสัมพันธ์ที่ระยะห่างหลายระยะ หากจำนวนสีในภาพมี  $m$  สีและใช้จำนวนระยะห่าง  $d$  ระยะห่าง จะต้องใช้เนื้อที่ในการเก็บข้อมูลคอรีโลแกรมมีขนาดเป็น  $m^2d$  ต่อหนึ่งภาพ ดังนั้น Jing Huang จึงเสนอออโตคอรีโลแกรม (autocorrelogram) เพื่อลดขนาดของดัชนี โดยพิจารณาเฉพาะคู่สีที่มีสีเดียวกันในแต่ละระยะห่างเท่านั้น ทำให้ขนาดของดัชนีภาพแบบออโตคอรีโลแกรมสีใช้เนื้อที่เพียง  $md$  ต่อภาพ โดยดัชนีแบบออโตคอรีโลแกรมคำนวณได้ดังสมการที่ 3.8

$$\Gamma_{c_i, c_i}^{(k)}(I) = |\{p_1, p_2 \in I_{c_i}, |p_1 - p_2| = k\}| \quad \dots(3.8)$$

ความน่าจะเป็นที่จะพบสี  $c_i$  ที่ระยะห่าง  $k$  ซึ่งเขียนแทนได้ด้วย  $\alpha_{c_i}^{(k)}(I)$  สามารถหาได้โดย

$$\alpha_{c_i}^{(k)}(I) = \gamma_{c_i, c_i}^{(k)} = \frac{\Gamma_{c_i, c_i}^{(k)}(I)}{H_{c_i}(I).8k} \quad \dots(3.9)$$

การค้นคืนภาพจากภาพที่ใช้สอบถามที่เป็นบางส่วนของภาพจากดัชนีภาพคอรีโลแกรมได้นำหลักการเดียวกันกับการอินเตอร์เซกต์ของฮิสโทแกรมสีซึ่งเรียกว่าการอินเตอร์เซกต์ของคอรีโลแกรมดัชนีภาพที่ได้จากการอินเตอร์เซกต์ของคอรีโลแกรมหาได้จากสมการที่ 3.10

$$\Gamma_{c_i, c_j}^{(k)}(Q \cap I) = \min\{\Gamma_{c_i, c_j}^{(k)}(Q), \Gamma_{c_i, c_j}^{(k)}(I)\} \quad \dots(3.10)$$

ค่าความน่าจะเป็นที่จะพบคู่สี  $c_i, c_j$  ที่ระยะห่าง  $k$  จากผลของการอินเตอร์เซกต์ของคอรีโลแกรมหรือเขียนแทนได้ว่า  $\gamma_{c_i, c_j}^{(k)}(Q \cap I)$  สามารถหาได้จากสมการที่ 3.11

$$\gamma_{c_i, c_j}^{(k)}(Q \cap I) = \frac{\Gamma_{c_i, c_j}^{(k)}(Q \cap I)}{H_{c_i, c_j}(Q \cap I).8k} \quad \dots(3.11)$$

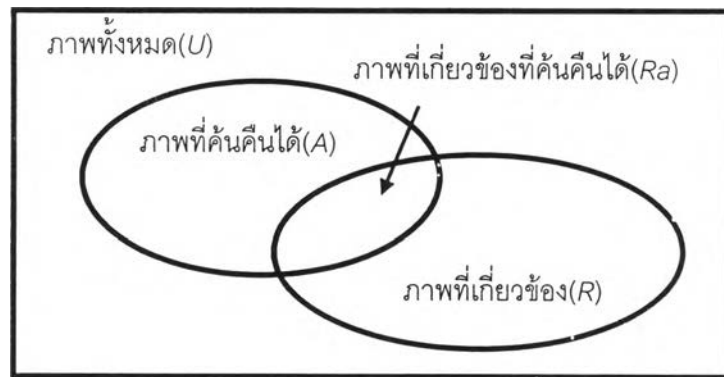
โดย  $|\gamma^{(k)}(Q) - \gamma^{(k)}(Q \cap I)|$  ที่มีค่ามากขึ้นหมายถึงโอกาสที่ภาพ  $Q$  บรรจุอยู่ในภาพ  $I$  น้อยลง และ  $|\gamma^{(k)}(Q) - \gamma^{(k)}(Q \cap I)| = 0$  หมายถึงมีโอกาสสูงที่ภาพ  $Q$  บรรจุอยู่ภายในภาพ  $I$  ในทำนองเดียวกับการเปรียบเทียบของฮิสโทแกรมสี

### 3.3 การวัดและเปรียบเทียบประสิทธิภาพ

ในการเปรียบเทียบความสามารถของขั้นตอนวิธีต่าง ๆ ที่ใช้ในการค้นคืนภาพจำเป็นต้องมีการวัดประสิทธิภาพของแต่ละวิธีที่ใช้ การวัดประสิทธิภาพของวิธีการค้นคืนสามารถทำได้หลายวิธี ซึ่งวิธีการ

วัดประสิทธิภาพที่นิยมใช้มีดังต่อไปนี้ คือ การวัดประสิทธิภาพจาก Precision และ Recall และการวัดประสิทธิภาพจาก Precision เฉลี่ยจากอันดับที่พบภาพที่เกี่ยวข้อง

กำหนดให้  $Q_1, \dots, Q_q$  เป็นกลุ่มของภาพที่ใช้สอบถามทั้งหมดจำนวน  $q$  ภาพ และให้การสอบถามที่ใช้ภาพ  $Q_i$  มีภาพ  $I_1^{(i)}, \dots, I_{a_j}^{(i)}$  เป็นภาพที่เกี่ยวข้อง (Relevant images) ที่ต้องการจากการสืบค้น ( $a_j$  คือจำนวนภาพที่มีภาพ  $Q_i$  บรรจุอยู่) ซึ่งแทนได้ด้วยเซต  $R$  (มีจำนวนเท่ากับ  $|R|$  ภาพ) ดังแสดงในรูปที่ 3.4 โดยในการค้นคืนภาพแต่ละครั้งจะได้ภาพผลลัพธ์ที่ออกมาเป็นภาพในเซต  $A$  (มีจำนวนเท่ากับ  $|A|$  ภาพ) ซึ่งภาพที่เกี่ยวข้องที่ค้นคืนได้เป็นภาพในเซต  $Ra$  (มีจำนวนเท่ากับ  $|Ra|$  ภาพ)



รูปที่ 3.4 เซตของภาพที่ได้จากการค้นคืนภาพ

### 3.3.1 Recall และ Precision [16]

การวัดประสิทธิภาพชนิดนี้เป็นการวัดประสิทธิภาพจากจำนวนของภาพที่เกี่ยวข้องที่ได้จากการสอบถาม โดยแสดงเป็นกราฟแสดงค่า Precision ในแต่ละระดับของค่า Recall ซึ่งสามารถคำนวณได้ดังนี้

Recall ( $r$ ) คือ อัตราส่วนของจำนวนภาพที่เกี่ยวข้องที่ได้จากการค้นคืนภาพต่อจำนวนของภาพที่เกี่ยวข้องที่มีทั้งหมด ดังสมการ 3.12

$$r = \frac{|Ra|}{|R|} \quad \dots(3.12)$$

Precision ( $p$ ) คือ อัตราส่วนของจำนวนภาพที่เกี่ยวข้องที่ได้จากการค้นคืนต่อจำนวนภาพทั้งหมดที่ได้จากการค้นคืนภาพ ดังสมการที่ 3.13

$$p = \frac{|Ra|}{|A|} \quad \dots(3.13)$$

ในการสร้างกราฟแสดงค่า Precision ในแต่ละระดับค่า Recall นั้น ทำได้โดยนำผลที่ได้จากการเปรียบเทียบความแตกต่างระหว่างดัชนีภาพของแต่ละภาพในฐานข้อมูลและภาพสอบถามมาจัดอันดับ

ตามค่าระยะห่างที่ได้จากการเปรียบเทียบดัชนี แล้วพิจารณาภาพจากอันดับที่หนึ่งไปจนถึงอันดับสุดท้าย โดยเมื่อพบภาพที่เกี่ยวข้องเพิ่มขึ้นแสดงถึงการเพิ่มขึ้นของค่า Recall ซึ่งเมื่อได้ค่า Recall ใหม่ให้ทำการ คำนวณค่า Precision ที่ระดับ Recall นั้น ๆ ในบางกรณีอาจพบภาพที่มีอันดับเดียวกันเนื่องจากมีค่า ระยะห่างจากการเปรียบเทียบดัชนีเท่ากัน ให้คำนวณค่า Recall จากจำนวนภาพที่เกี่ยวข้องที่พบใน อันดับก่อนหน้ารวมกับจำนวนภาพที่เกี่ยวข้องที่พบในอันดับที่นี้และหาค่า Precision จากจำนวนภาพ ทั้งหมดในอันดับก่อนหน้ารวมกับจำนวนภาพทั้งหมดในอันดับที่กำลังพิจารณาอยู่นี้

ตัวอย่างการคำนวณ กำหนดให้  $d_1 d_2 d_3 \dots$  แทนภาพแต่ละภาพและให้เซต  $R_q$  เป็นเซตของ ภาพที่เกี่ยวข้องที่ต้องการจากการค้นคืนภาพโดยใช้ภาพสอบถาม  $q$  โดย

$$R_q = \{d_1, d_3, d_5, d_7, d_9, d_{14}\}$$

ผลลัพธ์ที่ได้จากการค้นคืนภาพเรียงตามอันดับ โดยมีภาพที่เกี่ยวข้องที่อยู่ในอันดับของภาพ ผลลัพธ์นี้จะมีจุดสีดำอยู่ด้านหลังและมีภาพที่ได้อันดับที่ 7 จำนวน 3 ภาพ โดยมีผลดังนี้

1. $d_1$ ●	6. $d_6$	11. $d_{11}$
2. $d_2$	7. $d_7$ ●	12. $d_{12}$
3. $d_3$ ●	7. $d_8$	13. $d_{13}$
4. $d_4$	7. $d_9$ ●	14. $d_{14}$ ●
5. $d_5$ ●	10. $d_{10}$	15. $d_{15}$

การคำนวณค่า Precision ที่แต่ละระดับ Recall สามารถคำนวณจากจำนวนภาพที่เกี่ยวข้องที่ พบในแต่ละอันดับที่ ซึ่งจากผลการค้นคืนภาพข้างต้นที่มีจำนวนภาพที่เกี่ยวข้องทั้งหมด 6 ภาพ สามารถ คำนวณค่า Precision ที่แต่ละระดับ Recall ได้ดังนี้

- ภาพที่เกี่ยวข้องภาพแรกพบในอันดับที่ 1 ดังนั้นได้ค่า Recall =  $1/6 = 0.1667$  (1 ภาพที่ เกี่ยวข้องจาก 6 ภาพที่เกี่ยวข้องทั้งหมด) และได้ค่า Precision =  $1/1$  (จากภาพจำนวน 1 ภาพพบภาพที่ เกี่ยวข้อง 1 ภาพ)

- ภาพที่เกี่ยวข้องถัดไปพบในอันดับที่ 3 ดังนั้นได้ค่า Recall =  $2/6 = 0.3334$  (2 ภาพที่ เกี่ยวข้องจาก 6 ภาพที่เกี่ยวข้องทั้งหมด) และได้ค่า Precision =  $2/3 = 0.6667$  (จากภาพจำนวน 3 ภาพ พบภาพที่เกี่ยวข้อง 2 ภาพ)

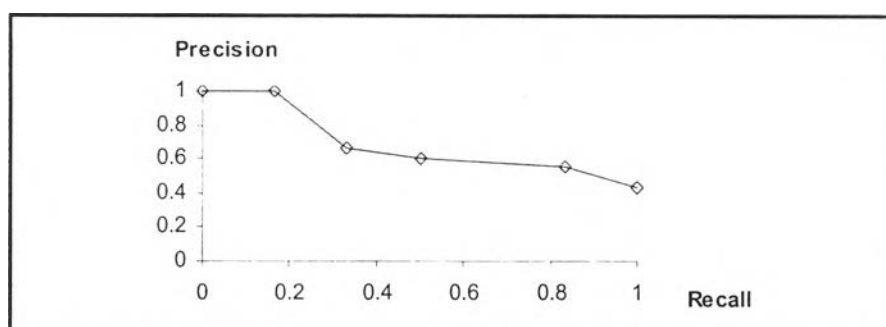
- ภาพที่เกี่ยวข้องถัดไปพบในอันดับที่ 5 ดังนั้นได้ค่า Recall =  $3/6 = 0.5$  (3 ภาพที่เกี่ยวข้อง จาก 6 ภาพที่เกี่ยวข้องทั้งหมด) และได้ค่า Precision =  $3/5 = 0.6$  (จากภาพจำนวน 5 ภาพพบภาพที่ เกี่ยวข้อง 3 ภาพ)

- ในอันดับที่ 7 มีภาพที่อยู่ในอันดับนี้ 3 ภาพซึ่งเป็นภาพที่เกี่ยวข้อง 2 ภาพ สามารถหา ค่า Recall ได้เป็น  $5/6 = 0.8333$  (5 ภาพที่เกี่ยวข้องจาก 6 ภาพที่เกี่ยวข้องทั้งหมด) และได้ค่า Precision =  $5/9 = 0.5556$  (จากภาพจำนวน 9 ภาพในการพบภาพที่เกี่ยวข้อง 5 ภาพ)



- ภาพที่เกี่ยวข้องสุดท้ายพบในอันดับที่ 14 ดังนั้นได้ค่า Recall =  $6/6 = 1$  (6 ภาพที่เกี่ยวข้องจาก 6 ภาพที่เกี่ยวข้องทั้งหมด) และ ได้ค่า Precision =  $6/14 = 0.6$  (จากภาพจำนวน 14 ภาพพบภาพที่เกี่ยวข้อง 6 ภาพ)

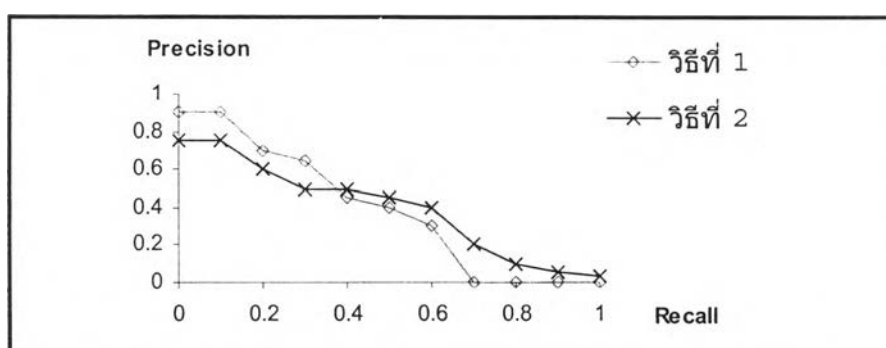
จากผลการคำนวณดังกล่าวสามารถนำไปสร้างเป็นกราฟแสดงค่า Precision ในแต่ละระดับค่า Recall ได้ดังในรูปที่ 3.5 (โดยทั่วไประดับ Recall ที่มีค่า 0 จะกำหนดให้มีค่า Precision เท่ากับค่า Precision ที่ได้จากระดับ Recall ถัดไป)



รูปที่ 3.5 ตัวอย่างการวัดประสิทธิภาพแบบ Recall และ Precision

กราฟ Precision และ Recall ในตัวอย่างข้างต้นนี้เป็นผลจากการค้นคืนภาพด้วยภาพสอบถามเพียงภาพเดียว อย่างไรก็ตาม โดยปกติการวัดประสิทธิภาพของวิธีที่ใช้ในการค้นคืนภาพจะทำการสอบถามจากหลาย ๆ การสอบถาม ดังนั้นในการวัดประสิทธิภาพด้วยวิธี Recall และ Precision จากหลาย ๆ การสอบถาม ค่า Precision ที่ได้จึงใช้เป็นค่าเฉลี่ยจากทุก ๆ การสอบถาม กล่าวคือ ค่า Precision ที่เป็นผลจากหลาย ๆ การสอบถามได้มาจากการเฉลี่ยค่า Precision ที่ระดับค่า Recall เดียวกันจากผลการค้นคืนในทุกการสอบถาม

การเปรียบเทียบประสิทธิภาพของวิธีที่ใช้ในการค้นคืนภาพจะพิจารณาจากค่า Precision ที่ระดับค่า Recall เดียวกันโดยค่า Precision ที่สูงกว่า (มีค่าใกล้ 1 มากกว่า) แสดงถึงประสิทธิภาพของวิธีการค้นคืนที่ดีกว่า ตัวอย่างในรูปที่ 3.6 แสดงให้เห็นว่าวิธีที่ 1 มีประสิทธิภาพดีกว่าวิธีที่ 2 ในช่วงที่ระดับ Recall ต่ำและวิธีที่ 2 มีประสิทธิภาพดีกว่าวิธีที่ 1 ในช่วงของระดับ Recall สูง ๆ



รูปที่ 3.6 ตัวอย่างกราฟเปรียบเทียบประสิทธิภาพแบบ Recall และ Precision

### 3.3.2 ค่า Precision เฉลี่ยจากอันดับที่พบภาพที่เกี่ยวข้อง

จากการวัดประสิทธิภาพแบบ Recall และ Precision นั้น ค่า Precision ที่ได้จะมีหลายค่า ซึ่งทำให้การเปรียบเทียบประสิทธิภาพของแต่ละวิธีทำได้ยาก ดังนั้น Precision เฉลี่ยจึงเป็นวิธีหนึ่งที่ใช้ในการเปรียบเทียบประสิทธิภาพของการค้นคืนภาพ โดยประสิทธิภาพของแต่ละวิธีจะถูกแสดงด้วยค่าเพียงค่าเดียว จากการเฉลี่ยค่า Precision จากทุกตำแหน่งที่พบภาพที่เกี่ยวข้องจากกราฟ Recall และ Precision โดยจากตัวอย่างของผลการค้นคืนภาพในรูปที่ 3.5 สามารถหาค่า Precision เฉลี่ยที่ตำแหน่งที่พบภาพที่เกี่ยวข้องได้เป็น  $(1 + 0.6667 + 0.6 + 0.5556 + 0.4285)/5 = 0.6502$  (ไม่นำค่า Precision ที่ระดับ Recall มีค่าเป็น 0 มาใช้ในการคำนวณ) สำหรับค่าที่ได้จากการวัดประสิทธิภาพวิธีนี้ ค่าที่สูงกว่าแสดงถึงการมีประสิทธิภาพในการค้นคืนดีกว่า สำหรับในการวัดประสิทธิภาพจากหลาย ๆ การสอบถามก็จะนำค่า Precision เฉลี่ยนี้จากทุก ๆ การสอบถามมาทำการเฉลี่ยรวมกันเช่นกัน

### 3.3.3 จำนวนการสอบถามที่ได้ค่า Precision สูงสุด

อีกวิธีหนึ่งในการวัดว่าวิธีใดมีประสิทธิภาพมากกว่ากันจากการสอบถามหลาย ๆ ครั้ง คือ การนับจำนวนการสอบถามที่แต่ละวิธีได้ค่า Precision เฉลี่ยจากแต่ละการสอบถามสูงสุด โดยวิธีการใดที่ได้จำนวนการสอบถามที่มีค่า Precision เฉลี่ยสูงสุดมากกว่าแสดงว่าประสิทธิภาพในการค้นคืนภาพสูงกว่าด้วย

### 3.3.4 ผลรวมของค่าน้ำหนัก

ในบางครั้งแต่ละการสอบถามนั้นมีภาพที่เกี่ยวข้องไม่เท่ากัน การวัดประสิทธิภาพจากค่าเฉลี่ย Precision จากทุกการสอบถามจึงไม่เหมาะสม วิธีการให้ผลรวมของค่าน้ำหนักแสดงความถูกต้องของแต่ละวิธีจึงเป็นอีกวิธีหนึ่งที่ใช้ในการเปรียบเทียบ ซึ่งคำนวณได้จากการให้น้ำหนักจากผลการค้นคืนภาพ โดยพิจารณาที่จะการสอบถามว่าผลการค้นคืนภาพวิธีใดที่มีค่า Precision เฉลี่ยสูงสุดจะมีค่าน้ำหนักมากที่สุดค่าหนึ่ง (ในการทดลองนี้ให้ค่าเท่ากับจำนวนวิธีที่ทำการเปรียบเทียบ) และวิธีที่มีค่า Precision เฉลี่ยรองลงมาจะมีค่าน้ำหนักน้อยลงไปตามลำดับ แล้วนำค่าน้ำหนักจากวิธีการค้นคืนภาพเดียวกันในแต่ละการสอบถามมารวมกัน ซึ่งวิธีที่ได้ค่าผลรวมที่มากกว่าแสดงว่ามีประสิทธิภาพสูงกว่า

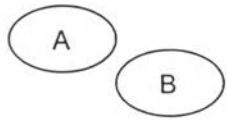
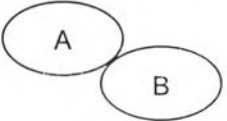
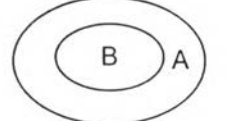
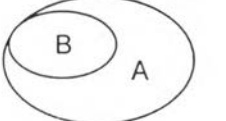
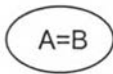
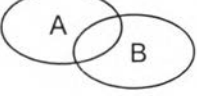

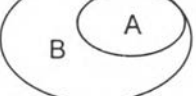
## 3.4 ความสัมพันธ์ทอพอโลยี [11]

ความสัมพันธ์ทอพอโลยีเป็นความสัมพันธ์ในเชิงพื้นที่ที่ความสัมพันธ์ไม่เปลี่ยนแปลงเมื่อมีการเปลี่ยนแปลงแบบสัมพรรค (Affine transformations) เช่นการหมุน ย่อขยาย หรือการยืดหดของภาพ วิธีการหนึ่งที่ใช้ในการแสดงความสัมพันธ์แบบทอพอโลยีคือการแสดงความสัมพันธ์ทอพอโลยีฐานสอง (Binary topological relations) [11] กล่าวคือพิจารณาความสัมพันธ์ของสิ่งที่พิจารณาที่ละคู่ โดย Egenhofer [11] ได้กำหนดความสัมพันธ์โดยใช้การอินเตอร์เซกต์ของภายใน (Interiors) ของบริเวณและขอบ (Boundaries) ของบริเวณโดยวิธีที่เรียกว่า *4-intersection*

กำหนดให้  $A$  เป็นบริเวณที่ไม่มีหลุมภายในบริเวณ โดยภายในของบริเวณแทนด้วย  $A^\circ$  ขอบของบริเวณแทนด้วย  $\partial A$  และภายนอกบริเวณแทนด้วย  $A^-$  ความสัมพันธ์ทอพอโลยีฐานสองระหว่างสองบริเวณ  $A$  และ  $B$  สามารถหาได้โดยวิธี 4-intersection ของขอบและภายในของบริเวณ  $A$  กับขอบและภายในของบริเวณ  $B$  โดยใช้เมตริกซ์ขนาด  $2 \times 2$  ดังแสดงได้ดังนี้

$$\begin{pmatrix} \partial A \cap \partial B & \partial A \cap B^\circ \\ A^\circ \cap \partial B & A^\circ \cap B^\circ \end{pmatrix}$$

จากผลของการอินเตอร์เซกต์ที่เป็นเซตว่าง ( $\emptyset$ ) และเซตที่ไม่เป็นเซตว่าง ( $\neg\emptyset$ ) สามารถแบ่งความสัมพันธ์เชิงทอพอโลยีฐานสองได้ 16 แบบ ซึ่งมีเพียง 8 แบบเท่านั้นที่เกิดขึ้นจริง [11] ซึ่งได้แก่ ความสัมพันธ์แบบ ไม่มีส่วนร่วม (disjoint) ติดกัน (meets) บรรจุ (contains) คลุม (covers) เท่ากัน (equals) ล้ำกัน (overlaps) ภายใน (inside) และ คลุมโดย (covered-by) ดังรูปที่ 3.7 ซึ่งแสดงความสัมพันธ์แบบทอพอโลยี 8 แบบจาก 4-intersection ระหว่างบริเวณ  $A$  และ  $B$

 <p><math>\partial B \ B^\circ</math>  <math>\partial A \ \begin{pmatrix} \emptyset &amp; \emptyset \end{pmatrix}</math>  <math>A^\circ \ \begin{pmatrix} \emptyset &amp; \emptyset \end{pmatrix}</math>                      ไม่มีส่วนร่วม</p>	 <p><math>\partial B \ B^\circ</math>  <math>\partial A \ \begin{pmatrix} \neg\emptyset &amp; \emptyset \end{pmatrix}</math>  <math>A^\circ \ \begin{pmatrix} \emptyset &amp; \emptyset \end{pmatrix}</math>                      ติดกัน</p>	 <p><math>\partial B \ B^\circ</math>  <math>\partial A \ \begin{pmatrix} \emptyset &amp; \emptyset \end{pmatrix}</math>  <math>A^\circ \ \begin{pmatrix} \neg\emptyset &amp; \neg\emptyset \end{pmatrix}</math>                      บรรจุ</p>	 <p><math>\partial B \ B^\circ</math>  <math>\partial A \ \begin{pmatrix} \neg\emptyset &amp; \emptyset \end{pmatrix}</math>  <math>A^\circ \ \begin{pmatrix} \neg\emptyset &amp; \neg\emptyset \end{pmatrix}</math>                      คลุม</p>
 <p><math>\partial B \ B^\circ</math>  <math>\partial A \ \begin{pmatrix} \neg\emptyset &amp; \emptyset \end{pmatrix}</math>  <math>A^\circ \ \begin{pmatrix} \emptyset &amp; \neg\emptyset \end{pmatrix}</math>                      เท่ากัน</p>	 <p><math>\partial B \ B^\circ</math>  <math>\partial A \ \begin{pmatrix} \neg\emptyset &amp; \neg\emptyset \end{pmatrix}</math>  <math>A^\circ \ \begin{pmatrix} \neg\emptyset &amp; \neg\emptyset \end{pmatrix}</math>                      ล้ำกัน</p>	 <p><math>\partial B \ B^\circ</math>  <math>\partial A \ \begin{pmatrix} \emptyset &amp; \neg\emptyset \end{pmatrix}</math>  <math>A^\circ \ \begin{pmatrix} \emptyset &amp; \neg\emptyset \end{pmatrix}</math>                      ภายใน</p>	 <p><math>\partial B \ B^\circ</math>  <math>\partial A \ \begin{pmatrix} \neg\emptyset &amp; \neg\emptyset \end{pmatrix}</math>  <math>A^\circ \ \begin{pmatrix} \emptyset &amp; \neg\emptyset \end{pmatrix}</math>                      คลุมโดย</p>

รูปที่ 3.7 ความสัมพันธ์แบบทอพอโลยี 8 แบบจาก 4-intersection ระหว่างบริเวณ  $A$  และ  $B$

การกำหนดความสัมพันธ์ 8 แบบข้างต้นนี้เป็นการหาความสัมพันธ์จากวิธี 4-intersection ที่ใช้ขอบและภายในของบริเวณ แต่โดยปกติแล้วบริเวณที่ถูกนำมาพิจารณานั้นจะไม่มีขอบของบริเวณมาก่อน ทำให้ผู้หาความสัมพันธ์ระหว่าง 2 บริเวณนี้จะต้องกำหนดขอบของแต่ละบริเวณขึ้นมาก่อน ซึ่งการหาขอบนั้นทำได้ยุ่งยากและอาจเกิดปัญหาเช่น ไม่แน่ใจว่าจะกำหนดให้ขอบอยู่ภายในหรือภายนอกของแต่ละบริเวณ เป็นต้น ดังนั้นเพื่อหลีกเลี่ยงการพิจารณาความสัมพันธ์ที่มีการใช้ขอบเข้ามาเกี่ยวข้อง

จึงมีการพิจารณาความสัมพันธ์ทอพอโลยีโดยใช้พื้นที่ภายในและภายนอกของแต่ละบริเวณแทน [19] การหาความสัมพันธ์ทอพอโลยีจากวิธี 4-intersection จากภายในและภายนอกของบริเวณ A และ B หาได้ดังนี้

$$\begin{pmatrix} A^{\circ} \cap B^{\circ} & A^{\circ} \cap B^{-} \\ A^{-} \cap B^{\circ} & A^{-} \cap B^{-} \end{pmatrix}$$

ผลลัพธ์ที่ได้จากการหาความสัมพันธ์ทอพอโลยีโดยใช้พื้นที่ภายในและภายนอกบริเวณทำให้สามารถจำแนกความสัมพันธ์ทอพอโลยี 8 แบบที่แสดงในรูปที่ 3.7 ได้เป็น 5 กลุ่ม เช่น ความสัมพันธ์ทอพอโลยีแบบไม่มีส่วนร่วมและติดกันจะถูกนำมาจัดเป็นกลุ่มเดียวกัน ซึ่งผลลัพธ์ที่การอินเตอร์เซกต์พื้นที่ภายในและภายนอกบริเวณและความสัมพันธ์ของแต่ละกลุ่มแสดงในรูปที่ 3.8

$B^{\circ} \quad B^{-}$	$B^{\circ} \quad B^{-}$	$B^{\circ} \quad B^{-}$	$B^{\circ} \quad B^{-}$	$B^{\circ} \quad B^{-}$
$A^{\circ} \quad \begin{pmatrix} \neg\phi & \phi \\ \phi & \phi \end{pmatrix}$	$A^{\circ} \quad \begin{pmatrix} \phi & \phi \\ \phi & \phi \end{pmatrix}$	$A^{\circ} \quad \begin{pmatrix} \phi & \neg\phi \\ \neg\phi & \phi \end{pmatrix}$	$A^{\circ} \quad \begin{pmatrix} \phi & \phi \\ \neg\phi & \phi \end{pmatrix}$	$A^{\circ} \quad \begin{pmatrix} \phi & \neg\phi \\ \phi & \phi \end{pmatrix}$
$A^{-} \quad \begin{pmatrix} \phi & \phi \\ \phi & \phi \end{pmatrix}$	$A^{-} \quad \begin{pmatrix} \phi & \phi \\ \phi & \phi \end{pmatrix}$	$A^{-} \quad \begin{pmatrix} \phi & \neg\phi \\ \neg\phi & \phi \end{pmatrix}$	$A^{-} \quad \begin{pmatrix} \phi & \phi \\ \neg\phi & \phi \end{pmatrix}$	$A^{-} \quad \begin{pmatrix} \phi & \neg\phi \\ \phi & \phi \end{pmatrix}$
ไม่มีส่วนร่วม / ติดกัน	ล้ำกัน	เท่ากัน	บรรจุ / คลุม	ภายใน / คลุมโดย

รูปที่ 3.8 ความสัมพันธ์ทอพอโลยี 5 แบบจากความสัมพันธ์ 8 แบบ

ในงานวิจัยนี้ได้นำความสัมพันธ์ระหว่างสองบริเวณในรูปความสัมพันธ์ทอพอโลยีมาใช้แทนความสัมพันธ์อิงระยะทางเพื่อใช้เป็นดัชนีภาพสำหรับการค้นคืนภาพเพื่อให้ทนทานต่อการย่อและขยายขนาด ซึ่งรายละเอียดการพัฒนาจะได้กล่าวในบทถัดไป