หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

พื้นฐานและงานวิจัยที่เกี่ยวข้อง

ปัจจุบัน เทคนิคในการค้นคืนภาพได้ถูกนำเสนอขึ้นมามากมายหลายวิธี เทคนิคเหล่านี้สามารถแบ่งตามลักษณะของการค้นคืนได้ 3 แบบ [2] ได้แก่ การค้นคืนภาพจากข้อความ (Text-based retrieval) การค้นคืนภาพจากเนื้อหา (Content-based retrieval) และการค้นคืนภาพจากความหมาย (Semantic-based retrieval) โดยแต่ละวิธีจะมีข้อดีและข้อจำกัดที่ต่างกัน ในการค้นคืนภาพจากข้อความ แต่ละภาพจะถูกอธิบายด้วยคำสำคัญและทำการค้นหาภาพโดยเปรียบเทียบหาคำสำคัญตรงกับข้อความที่สอบถาม สำหรับการค้นคืนภาพจากเนื้อหา แต่ละภาพจะถูกวิเคราะห์หาลักษณะที่สำคัญเพื่อใช้แสดงถึงภาพ ลักษณะสำคัญที่นิยมใช้ได้แก่ สี ลักษณะพื้นผิว หรือรูปร่าง เป็นต้น และในการค้นคืนภาพจากความหมาย จะต้องมีการกำหนดความหมายตามลักษณะในฐานข้อมูลเช่น การกำหนดให้กลุ่มภาพที่มีจุดภาพสีส้มกระจายอยู่หมายถึงภาพพระอาทิตย์ตกดิน เป็นต้น ระบบของการค้นคืนภาพลักษณะนี้บางระบบจำต้องมีการสร้างฐานข้อมูลความรู้ (Knowledge base) ประกอบด้วยเพื่อให้สามารถเรียนรู้และจำแนกความหมายได้เองเมื่อมีภาพเพิ่มเข้ามา จากการค้นคืนภาพทั้ง 3 แบบนี้สามารถสรุปข้อดีและข้อเสียของแต่ละแบบได้ดังตารางที่ 2.1

ตารางที่ 2.1 ตารางเปรียบเทียบวิธีการค้นคืนภาพในแต่ละแบบ [2]

วิธีการค้นคืน	ข้อดี	ข้อเสีย
การค้นคืนภาพจากข้อความ	ใช้คำถามเป็นข้อความซึ่งสามารถทำได้ง่าย	ประสิทธิภาพดีหรือไม่ขึ้นอยู่กับการกำหนดคำอธิบายภาพ
การค้นคืนภาพจากเนื้อหา	จำแนกลักษณะสำคัญได้โดยอัตโนมัติ	ยากในการกำหนดคำถามเป็นลักษณะข้อความ
การค้นคืนภาพจากความหมาย	สามารถกำหนดคำถามตามความเข้าใจของมนุษย์	ขอบเขตความหมายค่อนข้างชี้เฉพาะไปในเรื่องใดเรื่องหนึ่งเท่านั้น

จากข้อดีและข้อเสียในตารางที่ 2.1 ในการค้นคืนภาพจากข้อความนั้น ต้องมีการกำหนดคำสำคัญให้แต่ละภาพซึ่งต้องใช้เวลามากและยังไม่สามารถอธิบายภาพได้อย่างถูกต้องและครบถ้วน และในการค้นคืนจากความหมายถูกจำกัดความหมายจากความรู้ของผู้สร้างระบบ ดังนั้น จากข้อจำกัด

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ดังกล่าวทำให้การค้นค้นจากเนื้อหาเป็นที่นิยมมากกว่าเนื่องจากข้อมูลสามารถจำแนกได้โดยอัตโนมัติ และไม่ถูกจำกัดขอบเขตจากผู้สร้างระบบ

2.1 การสอบถามในการค้นคืนภาพ

โดยทั่วไป ในระบบการค้นคืนภาพจากเนื้อหาจะป้อนตัวอย่างของภาพเข้ามาเป็นคำถามให้กับระบบ หรือเรียกว่า *การสอบถามโดยใช้ตัวอย่าง* (*Query-by-example*) [3] ซึ่งแต่เดิมนั้นการสอบถามแบบนี้ได้มีการนำมาใช้กับฐานข้อมูลแบบข้อความ โดยใช้การป้อนค่าและข้อกำหนดเป็นตัวอย่างของการสอบถาม เพื่อให้ระบบทำการแปลงรูปแบบจากค่าตัวอย่างและข้อกำหนดไปเป็นของภาษาสอบถามเชิงโครงสร้าง (SQL) ก่อน แล้วนำไปสอบถามกับระบบอีกทีหนึ่ง ซึ่งบริษัท IBM เป็นบริษัทแรกที่สร้างระบบการสอบถามโดยใช้ตัวอย่างแบบนี้ ในปัจจุบัน ตัวอย่างหนึ่งของระบบการสอบถามโดยใช้ตัวอย่าง คือ โปรแกรม Microsoft Access ซึ่งเป็นการค้นหาในฐานข้อมูลเชิงสัมพันธ์ (Relational database) ที่ค่าของข้อมูลเป็นข้อความหรือตัวเลข ระบบการสอบถามโดยใช้ตัวอย่างของฐานข้อมูลชนิดนี้เป็นเพียงการสร้างส่วนติดต่อให้ผู้ใช้สามารถใช้งานได้สะดวกสบายขึ้นเท่านั้น แต่ไม่ได้เพิ่มความสามารถให้กับระบบแต่อย่างใด

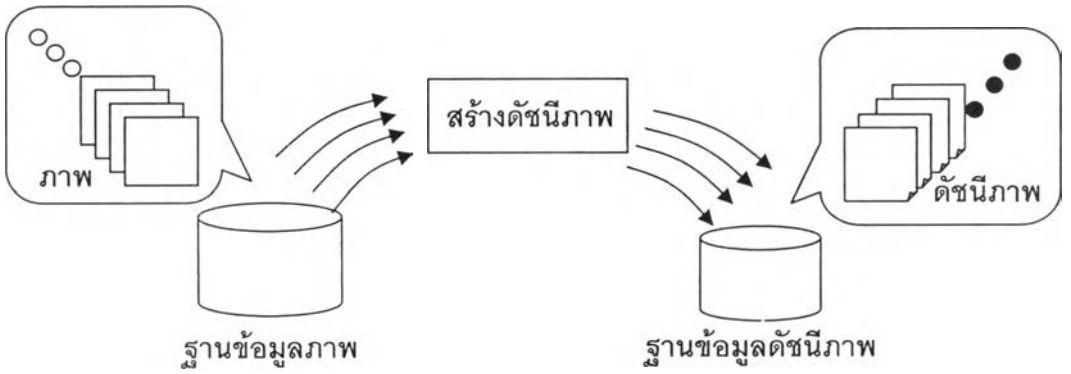
สำหรับฐานข้อมูลภาพ การสอบถามโดยใช้ตัวอย่างที่เป็นนิยมใช้กันมาก โดยสามารถให้คำถามเป็นภาพตัวอย่าง หรือภาพจากการระบายสีบนหน้าจอ หรือภาพร่างของวัตถุแทนที่จะทำการสอบถามโดยทำการพิมพ์คำถามให้กับระบบ ซึ่งระบบสามารถให้ผลลัพธ์เป็นภาพที่คล้ายคลึงหรือภาพที่มีวัตถุคล้ายคลึงกับภาพที่สอบถามได้

ในการค้นคืนภาพนอกจากต้องการค้นหาภาพที่มีลักษณะคล้ายกันทั้งภาพแล้ว การค้นหาว่ามีภาพใดบ้างที่มีวัตถุดังตัวอย่างที่สอบถามก็เป็นปัญหาที่ถูกนำมาถามบ่อย การค้นคืนภาพเพื่อได้ภาพที่มีภาพสอบถามอยู่ภายในนี้ เรียกว่าการค้นคืนภาพโดยใช้บางส่วนของภาพ (Image subregion querying) [1] ซึ่งเป็นปัญหาเบื้องต้นของปัญหาการรู้จำวัตถุ (Object recognition) คำถามลักษณะนี้เกิดขึ้นเมื่อผู้สอบถามสนใจวัตถุที่ต้องการมากกว่าความคล้ายคลึงของพื้นหลังหรือลักษณะโดยรวมของทั้งภาพ โดยผู้สอบถามจะป้อนคำถามด้วยภาพตัวอย่างที่เป็นบางส่วนของภาพที่ต้องการ แล้วทำการค้นคืนภาพในกลุ่มภาพหรือในฐานข้อมูลเพื่อให้ได้ภาพที่ในภาพมีบางส่วนของภาพนี้ปรากฏอยู่ โดยกลุ่มของภาพอาจจะประกอบด้วยภาพนิ่งหรือภาพวีดิทัศน์หรือทั้งสองอย่างผสมกัน การค้นคืนภาพในลักษณะนี้ค่อนข้างยากกว่าการค้นคืนภาพโดยทั่วไปเนื่องจากภาพวัตถุเดียวกันที่อยู่ในภาพคนละภาพอาจมีความแตกต่างกันซึ่งอาจเกิดมาจากมุมมองที่ต่างกัน ตำแหน่งการวางตัวของวัตถุที่ต่างกัน การมีสัญญาณรบกวนในภาพถ่าย และการที่วัตถุถูกปิดบังบางส่วน เป็นต้น

2.2 ขั้นตอนของระบบการค้นคืนภาพ

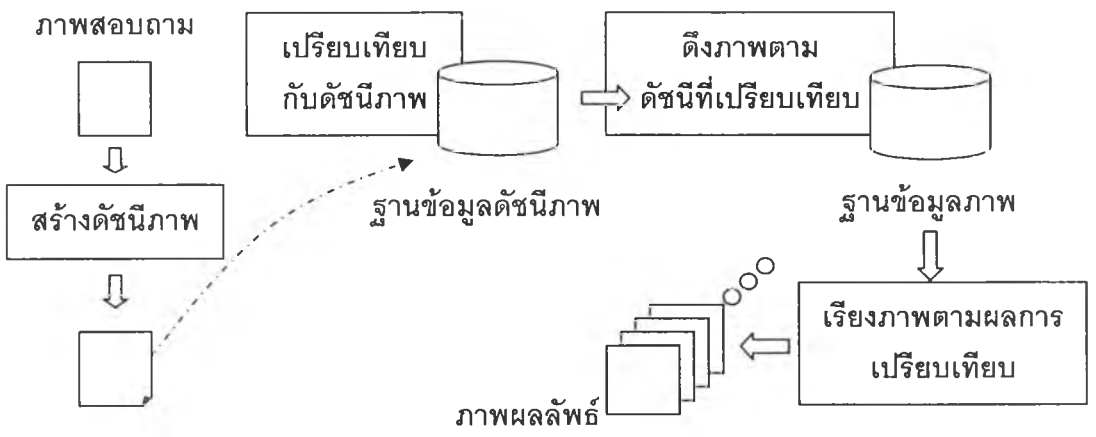
ขั้นตอนของระบบการค้นคืนภาพในแต่ละระบบจะขึ้นอยู่กับการออกแบบระบบของผู้สร้างระบบ แต่โดยทั่วไปแล้วระบบการค้นคืนภาพจะประกอบด้วย 2 ขั้นตอนหลัก ๆ [1] คือ

1) *สร้างดัชนีภาพ* เป็นการสร้างดัชนีเพื่อใช้ในการแสดงถึงแต่ละภาพในฐานข้อมูล โดยทั่วไปดัชนีของภาพจะได้มาจากการจำแนก*ลักษณะสำคัญของภาพ* (*image feature*) ดัชนีของภาพเหล่านี้จะถูกเก็บไว้ในฐานข้อมูลที่เก็บดัชนีภาพ ดังรูปที่ 2.1



รูปที่ 2.1 ขั้นตอนการสร้างดัชนีภาพ

2) *ค้นหาภาพ* เป็นขั้นตอนที่ถูกใช้เมื่อมีการสอบถามเกิดขึ้นโดยภาพที่นำมาสอบถามจะถูกนำมาสร้างดัชนีภาพ แล้วนำเอาดัชนีภาพที่ได้จากภาพสอบถามไปเปรียบเทียบกับดัชนีภาพของแต่ละภาพในฐานข้อมูลที่ได้ถูกเก็บไว้ ดังรูปที่ 2.2



รูปที่ 2.2 ขั้นตอนในการค้นหาภาพ

การสร้างดัชนีภาพนั้น นอกจากเป็นการจำแนกลักษณะที่สำคัญในภาพออกมาแล้วยังช่วยให้สามารถค้นคืนภาพได้รวดเร็วยิ่งขึ้นอีกด้วย เพราะไม่ต้องเสียเวลาจำแนกลักษณะสำคัญในขณะที่ทำการค้นหาภาพ

2.3 การจำแนกลักษณะสำคัญ

ในขั้นตอนการค้นหาภาพของระบบการค้นคืนภาพข้างต้นจะต้องทำการเปรียบเทียบดัชนีภาพ โดยที่ส่วนใหญ่แล้วดัชนีภาพได้มาจากลักษณะสำคัญของภาพ ดังนั้นผลของการค้นคืนภาพจะดีหรือไม่นั้นลักษณะสำคัญจึงเป็นสิ่งสำคัญที่ต้องคำนึงถึงในการค้นคืนภาพ ในทางคณิตศาสตร์ ลักษณะสำคัญหมายถึง เวกเตอร์ที่มีขนาด n โดยที่สมาชิกในเวกเตอร์มาจากข้อมูลที่วิเคราะห์และจำแนกได้จากภาพ ลักษณะสำคัญที่นิยมใช้ในการค้นคืนภาพจากเนื้อหาคือลักษณะสำคัญจากสิ่งที่เห็น (Visual feature) [1] ซึ่งได้แก่ ข้อมูลสี ข้อมูลพื้นผิว หรือข้อมูลรูปร่าง เป็นต้น ตัวอย่างหนึ่งของการนำข้อมูลสีในภาพมาแสดงในรูปของลักษณะสำคัญ เช่น ฮิสโทแกรมของสีในภาพ เป็นต้น บางครั้งสมาชิกในเวกเตอร์ลักษณะสำคัญสามารถประกอบด้วยข้อมูลจากสิ่งที่เห็นหลายชนิดก็ได้ เช่น ลักษณะสำคัญที่ประกอบด้วยข้อมูลสีและพื้นผิวของภาพ เป็นต้น

ลักษณะสำคัญจากสิ่งที่เห็นของภาพสามารถแบ่งเป็น 3 ระดับ [1] ได้ดังนี้

1) *ลักษณะสำคัญในระดับต่ำของภาพ* (*Low-level image features*) เป็นลักษณะสำคัญที่หาได้จากภาพโดยตรง เช่น ข้อมูลสี รูปร่าง พื้นผิว เป็นต้น

2) *ลักษณะสำคัญในระดับกลางของภาพ* (*Middle-level image features*) เช่น บริเวณหรือขอบเขต (region) เป็นต้น ลักษณะสำคัญชนิดนี้ต้องมีการนำลักษณะสำคัญในระดับต่ำมาผ่านกระบวนการบางอย่างก่อน เช่น การแบ่งส่วนภาพแล้วนำลักษณะสำคัญของทั้งบริเวณหรือขอบเขตไปใช้งาน เป็นต้น

3) *ลักษณะสำคัญในระดับสูงของภาพ* (*High-level image features*) ลักษณะสำคัญชนิดนี้ต้องมีมนุษย์เป็นผู้ช่วยกำหนด เช่น การแบ่งกลุ่มสีโดยการกำหนดสีเป็นกลุ่มสีร้อนและกลุ่มสีเย็น เป็นต้น

ระบบการค้นคืนภาพส่วนใหญ่นิยมนำลักษณะสำคัญระดับต่ำมาใช้เนื่องจากสามารถจำแนกได้โดยอัตโนมัติ ไม่จำเป็นต้องใช้คนช่วยกำหนดเหมือนลักษณะสำคัญระดับสูง ส่วนลักษณะสำคัญในระดับกลางนั้นมักถูกใช้ในการแบ่งส่วนเพื่อหาวัตถุที่ต้องการ

ในการเลือกใช้หรือออกแบบลักษณะสำคัญสำหรับระบบการค้นคืนภาพ ควรคำนึงถึงลักษณะสำคัญของภาพที่นำมาใช้ [1] ดังนี้

1) *สำเหนียกถึงความคล้ายคลึงได้* (*Perceptual similarity*): ภาพที่ต่างกันมากควรจะมีค่าความแตกต่างของลักษณะสำคัญมากด้วย

2) *มีประสิทธิภาพ* (*Efficiency*): สามารถคำนวณหาลักษณะสำคัญได้อย่างรวดเร็ว

3) *ประหยัด* (*Economy*): ลักษณะสำคัญที่เลือกใช้ควรจะมีขนาดเล็ก ขนาดของลักษณะสำคัญไม่ได้มีผลกระทบต่อประสิทธิภาพของการค้นคืนเท่านั้น แต่มีผลต่อการออกแบบโครงสร้างข้อมูลในการเก็บดัชนีด้วย

2.4 การเปรียบเทียบลักษณะสำคัญของภาพ (Image Feature Comparison)

โดยปกติแล้ว ความคล้ายคลึงของภาพสองภาพเปรียบเทียบมาจากความแตกต่างของลักษณะสำคัญของภาพ เนื่องจากดัชนีของภาพเป็นเวกเตอร์ลักษณะสำคัญของภาพ ดังนั้นจึงสามารถใช้ฟังก์ชันระยะทาง (Distance function) ในการวัดความแตกต่างของเวกเตอร์ลักษณะสำคัญของทั้งสองภาพโดยที่ค่าระยะทางเป็นศูนย์หมายถึงภาพทั้งสองที่นำมาเปรียบเทียบกันเหมือนกัน และค่าระยะทางยิ่งมากขึ้นแสดงให้เห็นถึงระดับความแตกต่างของทั้งสองภาพที่มากขึ้น ฟังก์ชันระยะทางที่รู้จักกันดีและนิยมใช้ คือ Euclidean norm (หรือ L_2 norm) และ city-block metric (หรือ L_1 norm) [4] นอกจากนี้ในบางครั้งเวกเตอร์ลักษณะสำคัญของภาพเป็นการแสดงถึงการแจกแจงต่าง ๆ ในภาพ เช่น ฮิสโทแกรมสี เป็นการแสดงการกระจายของสีในภาพ ดังนั้นฟังก์ชันระยะทางในการวัดความแตกต่างของการแจกแจงจึงควรเป็นฟังก์ชันที่สามารถวัดความแตกต่างของการแจกแจงของทั้งสองภาพด้วย เช่น ฟังก์ชัน Mahalanobis distance [5] เป็นต้น

2.5 ลักษณะสำคัญสำหรับการค้นคืนภาพจากเนื้อหา

ลักษณะสำคัญในระดับต่ำเช่น สี พื้นผิว รูปร่าง เป็นลักษณะสำคัญที่มักจะถูกนำมาใช้ในระบบการค้นคืนภาพจากเนื้อหาเนื่องจากลักษณะสำคัญเหล่านี้สามารถคำนวณได้โดยอัตโนมัติ ในหัวข้อนี้จึงเป็นการกล่าวถึงลักษณะของแต่ละลักษณะสำคัญต่าง ๆ ที่นำมาใช้ในการค้นคืนภาพอย่างคร่าว ๆ

2.5.1 สี

สีเป็นลักษณะสำคัญจากสิ่งที่เห็นที่มีการนำมาใช้กันอย่างมากในการค้นคืนภาพ ฮิสโทแกรมสีเป็นลักษณะสำคัญของสีอย่างหนึ่งที่ถูกนำมาใช้บ่อย ๆ โดยมีงานวิจัยหลายงานที่นำลักษณะสำคัญแบบฮิสโทแกรมสีมาใช้ อย่างไรก็ตาม แม้ว่ามีการใช้ฮิสโทแกรมสีกันมาก แต่โดยปกติแล้วฮิสโทแกรมจะเป็นการแจกแจงของสีของทั้งภาพ ทำให้ไม่มีการนำเอาข้อมูลเชิงพื้นที่มาใช้ ดังนั้นงานวิจัยจำนวนมากจึงได้พัฒนาต่อมาจากฮิสโทแกรม ซึ่งส่วนใหญ่เป็นการนำข้อมูลเชิงพื้นที่เข้ามาใช้ประกอบด้วย

2.5.2 พื้นผิว (Texture)

พื้นผิวเป็นอีกสิ่งหนึ่งที่นำมาใช้เป็นลักษณะสำคัญของภาพ พื้นผิวของภาพหมายถึงรูปแบบของสิ่งที่เห็นซึ่งไม่สามารถแทนได้ด้วยสีเพียงสีเดียวหรือความสว่างเพียงค่าเดียว การใช้พื้นผิวในการค้นคืนภาพจะต้องคำนึงถึง 2 สิ่ง [3] คือ (1) โมเดลของพื้นผิว (Texture model) และ (2) ฟังก์ชันระยะทางที่ใช้วัดความแตกต่างระหว่างรูปแบบของพื้นผิว โดยโมเดลของพื้นผิวมักจะถูกแสดงด้วยเวกเตอร์อธิบายพื้นผิว (Texture description vector) การค้นคืนภาพที่ใช้พื้นผิวเป็นลักษณะสำคัญของภาพส่วนใหญ่จะถูกนำไปใช้ในการค้นหาภาพพื้นผิวจากกลุ่มภาพพื้นผิวของ Brodatz ซึ่งเป็นชุดภาพพื้นผิวแบบต่าง ๆ เช่น ชุดภาพพื้นผิวของหิน ชุดภาพลายผ้า เป็นต้น ทำให้ไม่เหมาะกับในการนำไปใช้ค้นคืนภาพทั่วไป

2.5.3 รูปร่าง

จากลักษณะสำคัญทั้งสองลักษณะที่กล่าวมา ทั้งสีและพื้นผิวแสดงลักษณะสำคัญของภาพทั้งภาพ ดังนั้นเพื่อให้สามารถอ้างถึงวัตถุหรือบริเวณใด ๆ ในภาพได้ รูปร่างของวัตถุหรือรูปร่างของบริเวณจึงถูกนำมาเป็นลักษณะสำคัญของภาพ ซึ่งในการจำแนกลักษณะสำคัญนี้ออกมา มักจะใช้วิธีการหาขอบภาพ (Edge detection) หรือการแบ่งส่วนภาพ (Image segmentation) สำหรับวิธีการวิเคราะห์ความแตกต่างของรูปร่างนั้นมีอยู่หลายวิธี ไม่ว่าจะเป็นวิธีทางสถิติ วิธีที่ใช้พื้นฐานการวัดระยะแบบเฮาส์-ดอร์ฟ (Hausdorff-distance based approach) [6] วิธีการประมาณจากรูปหลายเหลี่ยม (Polygonal approximations) [7] และการแสดงเป็นเส้นแสดงรูปร่าง (Contour-oriented representations) [8] อย่างไรก็ตามปัญหาอย่างหนึ่งของการใช้รูปร่างเป็นลักษณะสำคัญคือการวัดความคล้ายคลึงของรูปร่างว่าเหมือนหรือแตกต่างกันอย่างชัดเจนนั้นทำได้ค่อนข้างยาก

2.5.4 **ความสัมพันธ์เชิงพื้นที่** (Spatial Relationship)

ความสัมพันธ์เชิงพื้นที่เป็นการเก็บความสัมพันธ์ของวัตถุที่แสดงอยู่บนภาพในเชิงพื้นที่หรือตำแหน่ง ในการค้นคืนภาพ ความสัมพันธ์เชิงพื้นที่เป็นส่วนประกอบที่ใช้ช่วยอธิบายลักษณะสำคัญของภาพซึ่งสามารถบอกความสัมพันธ์ของวัตถุหรือคุณสมบัติเช่น สี รูปร่าง และพื้นผิวของสิ่งที่เห็นในภาพ ในระบบสารสนเทศทางภูมิศาสตร์ (Geographic information systems : GISs) ได้แบ่งความสัมพันธ์เชิงพื้นที่ไว้ 3 ชนิด [9] คือ ความสัมพันธ์ระยะทาง (Distance relation) ความสัมพันธ์ทิศทาง (Direction relation) และความสัมพันธ์ทอพอโลยี (Topology relation)

1) *ความสัมพันธ์ระยะทาง* เป็นการเก็บความสัมพันธ์จากระยะทางของวัตถุในการที่จะย้ายวัตถุจากตำแหน่งหนึ่งไปอีกตำแหน่งหนึ่ง โดยทั่วไประยะทางที่เก็บไว้จะเป็นระยะทางเชิงปริมาณ เช่น ระยะทางยุคลิเดียน (Euclidean distance) ต่อมาในปี ค.ศ.1994 Hong [10] ได้มีการแปลงระยะทางเชิงปริมาณเป็นระยะทางเชิงคุณภาพ โดยใช้ระยะใกล้ (Near) ปานกลาง (Medium) ไกล (Far) และไกลมาก (Very far) เป็นต้น

2) *ความสัมพันธ์ทิศทาง* จะเก็บความสัมพันธ์จากตำแหน่งของวัตถุหนึ่งกับอีกวัตถุหนึ่งบนปริภูมิ (Space) เช่น ทางเหนือ ทางใต้ ทางซ้าย ทางขวา เป็นต้น ตัวอย่างการแสดงความสัมพันธ์เชิงทิศทางเช่น วัตถุ A อยู่ทางเหนือของวัตถุ B เป็นต้น

3) *ความสัมพันธ์ทอพอโลยี* เป็นการเก็บความสัมพันธ์เชิงภูมิศาสตร์ระหว่างวัตถุซึ่งทนทานต่อการเปลี่ยนแปลงแบบสัมพรรค (Affine transformations) เช่น การเลื่อนที่ (Translation) การหมุน (Rotation) และการย่อขยาย (Scaling) ตัวอย่างของการแสดงความสัมพันธ์ทอพอโลยีเช่น ถ้าวัตถุ 2 ชิ้นไม่มีส่วนร่วมกัน (Disjoint) ถึงแม้ว่าปริภูมิที่วัตถุวางอยู่มีการหมุน วัตถุทั้งสองก็ยังคงไม่มีส่วนร่วมกันเช่นเดิม เป็นต้น ในปี ค.ศ.1991 Egenhofer และ Franzosa [11] ได้จำแนกความสัมพันธ์ทอพอโลยีระหว่างวัตถุ 2 ชิ้นได้ 8 แบบที่แตกต่างกัน คือความสัมพันธ์ระหว่างวัตถุแบบ ไม่มีส่วนร่วม (Disjoint) ล้ำ

กัน (Overlaps) ติดกัน (Meets) เท่ากัน (Equals) ภายใน (Inside) บรรจุ (Contains) คลุม (Covers) และ คลุมโดย (Covered-by) ซึ่งรายละเอียดของความสัมพันธ์เหล่านี้จะแสดงอยู่ในบทที่ 3

2.6 งานวิจัยที่เกี่ยวข้อง

สีเป็นลักษณะสำคัญที่ใช้กันอย่างมากในการค้นคืนภาพ งานวิจัยนี้ก็ได้นำสีมาใช้เป็นลักษณะสำคัญของภาพด้วย ซึ่งงานวิจัยที่เกี่ยวข้องมีดังนี้

ใน ค.ศ.1991 Swain และ Ballard [12] ได้นำฮิสโทแกรมสี (Color histogram) มาใช้เป็นดัชนีของภาพและเสนอการอินเตอร์เซกต์ของฮิสโทแกรมสี (Color histogram intersection) เป็นตัววัดความคล้ายของภาพ วิธีนี้เป็นวิธีที่ง่ายในการคำนวณ ไม่ต้องทำการแบ่งส่วนของพื้นหน้าพื้นหลัง แต่เนื่องจากฮิสโทแกรมไม่ได้เก็บข้อสนเทศที่เกี่ยวข้องกับตำแหน่ง ดังนั้นข้อสนเทศที่แสดงออกจึงเป็นความสัมพันธ์ของทั้งภาพ ทำให้ประสิทธิภาพในการค้นคืนภาพไม่สูงนัก

ค.ศ.1995 Matas และคณะ [13] ได้นำเสนอ Color adjacency graph (CAG) เพื่อค้นหาภาพที่มีวัตถุที่ต้องการ โดยการสร้างกราฟความสัมพันธ์ของสีในภาพ แต่ละโหนดของกราฟเป็นสีที่ได้มาจากการแบ่งส่วนของภาพโดยใช้ฮิสโทแกรม วิธีการค้นคืนภาพทำโดยหาว่ามีบางส่วนของกราฟในภาพใดที่เหมือนหรือคล้ายกราฟของวัตถุที่ค้นหา วิธีนี้ได้ผลค่อนข้างดีแต่ต้องใช้เวลานานในการคำนวณหากราฟที่เหมือนกัน

ค.ศ.1996 Pass และ Zabih [14] ได้นำฮิสโทแกรมมาปรับปรุงโดยแบ่งแต่ละแท่งฮิสโทแกรมออกเป็น 2 ส่วน คือ ส่วนของจุดภาพที่ Coherent กันและส่วนของจุดภาพที่ Incoherent กัน โดยจุดภาพที่ Coherent กันหมายถึงจุดภาพที่มีจุดภาพสีเดียวกันอยู่ติดกัน และจุดภาพที่ Incoherent หมายถึงจุดภาพที่ไม่มีจุดภาพสีเดียวกันติดกัน ฮิสโทแกรมที่ปรับปรุงใหม่นี้เรียกว่า Color Coherent Vector (CCV) ซึ่งสามารถคำนวณได้รวดเร็วและมีประสิทธิภาพในการค้นคืนภาพได้ดีกว่าการใช้ฮิสโทแกรมแบบเดิม

ค.ศ.1997 Huang และคณะ [1] ได้เสนอการแสดงความสัมพันธ์ของสีแบบใหม่มาเป็นดัชนีแทนฮิสโทแกรมเรียกว่า คอร์ริโลแกรม (Correlogram) ซึ่งคอร์ริโลแกรมนี้เก็บข้อมูลความสัมพันธ์ของทั้งสีและตำแหน่งโดยการนับจำนวนคู่สีที่ระยะห่างต่าง ๆ นอกจากนี้ยังได้เสนอดัชนีแบบออโตคอร์ริโลแกรม (Autocorrelogram) ซึ่งเก็บความสัมพันธ์เฉพาะคู่สีเดียวกันที่ระยะห่างต่าง ๆ นอกจากนี้ยังได้นำดัชนีแบบคอร์ริโลแกรมมาใช้ในการค้นคืนภาพจากบางส่วนของภาพแทนวิธีการอินเตอร์เซกต์ของฮิสโทแกรมสีเรียกว่าวิธีการอินเตอร์เซกต์ของคอร์ริโลแกรมสี (Color correlogram intersection) แต่ข้อจำกัดของการใช้คอร์ริโลแกรมคือต้องใช้เนื้อที่ในการเก็บดัชนีมากและการที่ดัชนีมีขนาดใหญ่ทำให้ต้องใช้เวลาในการคำนวณมาก งานวิจัยของ Huang และคณะจึงได้นำเอาออโตคอร์ริโลแกรมมาใช้ในการค้นคืนภาพ ซึ่งได้ผลการค้นคืนภาพดีกว่าการใช้ฮิสโทแกรมในการค้นคืนภาพ

ค.ศ.1999 Nicu Sebe และคณะ [15] ได้เสนอวิธีการค้นคืนภาพโดยใช้บางส่วนของภาพเพื่อให้สามารถค้นคืนภาพจากภาพสอบถามหลาย ๆ ขนาดได้ โดยวิธีการนี้จะต้องมีการเลือกบริเวณต่าง ๆ ที่สนใจในภาพก่อน ซึ่งบริเวณเหล่านี้สามารถคาบเกี่ยวกันได้ และสร้างดัชนีภาพจากผลต่างหรือระยะทางระหว่างค่าทางสถิติต่าง ๆ จากสีของทั้งภาพกับค่าทางสถิติต่าง ๆ จากสีของแต่ละบริเวณที่เลือกไว้ โดยค่าสถิติต่าง ๆ เหล่านี้ได้แก่ ค่าเฉลี่ยของสี (Average color)และค่าเมทริกซ์ความแปรปรวนร่วมของสี (Covariance matrix) และระยะทางดังกล่าวนี้เรียกว่า Inter Hierarchical Distance (IHD) ในการค้นคืนภาพก็จะนำภาพสอบถามมาหาระยะทางระหว่างค่าทางสถิติต่าง ๆ ของทั้งภาพกับค่าทางสถิติต่าง ๆ ของภาพสอบถามหรือกล่าวได้ว่าหาระยะทาง IHD กับภาพสอบถาม แล้วนำค่าผลต่างนี้ไปเปรียบเทียบค่าต่าง ๆ ของดัชนีของภาพ ซึ่งค่าผลต่างที่น้อยที่สุดระหว่าง IHD ของภาพสอบถามกับ IHD ของบริเวณที่เลือกไว้จะถือว่าเป็นค่าระยะทางของการค้นคืนภาพ นอกจากการค้นคืนภาพดังกล่าวแล้วดัชนีชนิดนี้ยังสามารถนำไปใช้ในการหาตำแหน่งของภาพสอบถามในภาพได้ (Localization)