

# การค้นคืนข้อมูลภาพโดยใช้เนื้อหาของภาพ

## Content-Based Image Retrieval

สาวลักษณ์ วรรณภา

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

### 1. บทนำ

ความสนใจในศักยภาพของข้อมูลภาพดิจิตอลมีเพิ่มมากขึ้นเป็นอันมากในช่วงไม่กี่ปีมานี้ ทั้งนี้ส่วนหนึ่งมาจากการเติบโตอย่างรวดเร็วของการใช้ข้อมูลภาพในอินเทอร์เน็ตและเว็บไซต์เว็บ (World Wide Web) หรือ WWW ซึ่งยังผลให้ในปัจจุบันมีภาพเก็บอยู่จำนวนมาก ผู้ใช้ในหลายสาขาอาชีพต่างใช้ประโยชน์จากความสามารถในการเข้าถึงและจัดการข้อมูลภาพที่เก็บอยู่ใน WWW จากระยะไกล อย่างไรก็ตาม ผู้ใช้ส่วนมากพบว่ากระบวนการในการค้นหาเพื่อให้ได้มาซึ่งภาพที่ต้องการต้องการของตนจาก WWW ซึ่งเป็นหัวใจของการทำงานทางมาศึกษาและมีความหลากหลายมากนั้นเป็นไปได้ยากลำบากยิ่ง [1] จากสาเหตุดังกล่าว จึงก่อให้เกิดความสนใจในเรื่องเกี่ยวกับการสืบค้นข้อมูลภาพ (Image Retrieval) ขึ้น และเป็นผลให้เกิดงานวิจัยและพัฒนาความรู้และเทคโนโลยีในด้านดังกล่าวเพิ่มขึ้นอย่างมากภายในปัจจุบันวิธีการดังเดิมที่ใช้ในการทำดัชนีภาพนั้นไม่มีประสิทธิภาพเพียงพอ [2] จึงก่อให้เกิดความสนใจในการค้นหาเทคโนโลยีใหม่ๆ เพื่อที่จะสืบค้นข้อมูลภาพที่ต้องการให้ได้โดยอัตโนมัติ โดยอาศัยคุณลักษณะของภาพที่มีอยู่ในภาพ อาทิ ลักษณะพื้นที่สี แสง รูปทรงของวัตถุในภาพ เกิดเมื่อเทคโนโลยีใหม่ที่เป็นที่รู้จักกันในชื่อ การค้นคืนข้อมูลภาพโดยใช้เนื้อหาของภาพ (Content-Based Image Retrieval) หรือ CBIR หลังจากผ่านการวิจัยมาเป็นเวลาอย่างยาวนาน การใช้เทคโนโลยี CBIR ก็ถูกขยายจากห้องปฏิบัติการวิจัย ออกมายังการธุรกิจ ในรูปของซอฟต์แวร์สำเร็จรูปอย่าง QBIC ของบริษัท IBM อย่างไรก็ตามเทคโนโลยีนี้ก็ยังมีได้ถูกใช้อย่างแพร่หลายเท่าไหรก็แล้วแต่ยังต้องการการพัฒนาเพิ่มเติมอีกมาก

### 2. การใช้งานข้อมูลภาพดิจิตอล

ใช้งานข้อมูลภาพในลักษณะนี้มีมานานมากแล้ว นับย้อนหลังไปตั้งแต่ยุคสมัยที่มีการวางแผนบนแผ่นดิน นอกจากนั้นยังมีหลักฐานการใช้งานแผนที่และผังอาคารมาตั้งแต่ก่อนสมัยโบราณ อย่างไรก็ตาม ในยุคศตวรรษที่ 20 นี้ได้มีการเติบโตอย่างมากมากทั้งในแง่จำนวนการใช้งานและความสำคัญของข้อมูลภาพที่มีต่อการดำเนินชีวิตของผู้คนในปัจจุบัน ข้อมูลภาพมีบทบาทสำคัญในหลาย ๆ ด้าน ตัวอย่างเช่น วงการแพทย์ วงการสื่อสารมวลชน วงการโฆษณา การศึกษา และการบันทึก เป็นต้น เทคโนโลยีที่อยู่ในรูปของสิ่งประดิษฐ์อย่างการถ่ายรูปและเครื่องรับโทรทัศน์ ได้เข้ามามีบทบาทสำคัญในการจัดเก็บและเลือกรายชื่อ ข้อมูลภาพ แต่สิ่งประดิษฐ์ล้ำๆ ที่สุดในการปฏิวัติข้อมูลภาพเห็นจะได้แก่ เครื่องคอมพิวเตอร์ซึ่งช่วยอำนวยความสะดวกในการจัดเก็บ ประมวลผล และส่งผ่านข้อมูลภาพในรูปแบบดิจิตอล ยุคของ การใช้งานคอมพิวเตอร์เพื่อช่วยจัดการข้อมูลภาพเริ่มต้นตั้งแต่ในรูปแบบดิจิตอล ต่อมาเมื่อเทคโนโลยีพัฒนาขึ้น อุปกรณ์คอมพิวเตอร์ต่างๆ ไม่ว่าจะเป็นหน่วยประมวลผล หน่วยความจำ แม่ปะลิกิจภาพดิจิตอล และราชากุลง ทำให้ความนิยมในการจัดเก็บข้อมูลภาพในรูปแบบดิจิตอลมากขึ้น และเพิ่มขึ้น แม้แต่ในหลายสาขาวิชา เช่น วิศวกรรมศาสตร์ สถาปัตยกรรมศาสตร์ การแพทย์ ห้องแสดงภาพ และพิพิธภัณฑ์ และเมื่อมีการพัฒนาระบบ WWW ขึ้นในช่วงประมาณปีค. 1991-1992 ก็ยิ่งอำนวยความสะดวกให้ผู้ใช้คอมพิวเตอร์สามารถเข้าถึงข้อมูลต่างๆ ที่ต้นเองต้องการจากทั่วทุกมุมโลก และยิ่งกระตุ้นให้มีการใช้งานข้อมูลภาพดิจิตอลเพิ่มขึ้นเป็นอันมาก ปัจจุบัน จำนวนภาพที่เก็บอยู่ใน WWW มีจำนวนมากกว่าล้านภาพและยังคงทวีจำนวนเพิ่มขึ้นเรื่อยๆ

### 3. การจัดการข้อมูลภาพ

ในทัวร์อีก 2. ได้แก่ส่วนไว้ว่าซึ่งมูลภาพพิจิตอลได้รับความนิยมเพิ่มมากขึ้นทุกขณะอย่างรวดเร็ว อย่างไรก็ตามกระบวนการการทำให้เข้ามูลภาพกล้ายเป็นหัวข้อในรูปแบบดิจิตอลเพื่อจัดเก็บในระบบคอมพิวเตอร์ได้นั้น มีด้วยวิธีที่การจัดการกลุ่มของหัวข้อมูลภาพจำนวนมากกันอย่างรวดเร็วและถูกต้องได้ ดังนั้นจึงยังคงมีความต้องการที่จะให้มีการทำการจัดหมวดหมู่และทำด้วยหัวข้อมูลภาพดังกล่าวอยู่ ซึ่งในยุคแรกๆ ระบบการจัดการหัวข้อมูลภาพจะใช้การจัดหมวดหมู่และทำด้วยโดยอาศัยคำสำคัญ (keyword) ซึ่งเป็นคำหรือข้อความที่มีนัยยะเป็นสัญลักษณ์เข้าไป และการค้นคืนก็จะใช้เทคนิคของการค้นคืนข้อความ (Text Retrieval) ซึ่งเป็นเทคโนโลยีที่พัฒนาเป็นเวลาหนาจนมีประสิทธิภาพพิเศษเป็นที่ยอมรับได้อยู่แล้ว ทว่าการทำด้วยโดยอาศัยมนุษย์เป็นผู้ตัดความหมายจากภาพแล้วป้อนคำสำคัญนั้นเมื่อเสียคือต้องใช้เวลาและแรงงานมนุษย์เป็นอย่างมาก ทั้งนี้ทำด้วยมือแล้วก็ต้องเสียเวลากันจำนวนมาก แต่ความต้องการที่จะลดเวลาลงและลดความซ้ำซ้อนของผู้ป้อนหัวข้อมูลก็มีอย่างมาก ซึ่งความเห็นดังกล่าวอาจแตกต่าง หรือไม่ครอบคลุมความต้องการของผู้ใช้ที่จะมาค้นคืนหัวข้อมูลภาพของระบบในภายหลัง ยังผลให้ผู้ใช้อาจไม่ได้ผลลัพธ์ที่ตรงตามความต้องการที่แท้จริงจากระบบนั้นๆ จากข้อบกพร่องดังกล่าว จึงมีความพยายามที่จะสร้างระบบการจัดการหัวข้อมูลภาพที่มีประสิทธิภาพดีขึ้นกว่าระบบดั้งเดิม โดยเป็นระบบที่สามารถดึงเอาสารสนเทศบางอย่างที่มีอยู่ในภาพออกมายได้โดยอัตโนมัติ แทนที่จะให้มนุษย์เป็นผู้ตัดความลับสารสนเทศเหล่านั้นแล้วแปลงเป็นรูปข้อความในปัจจุบันที่คิดทางของ การค้นคืนวิจัยเกี่ยวกับการจัดเก็บและเรียกค้นคืนหัวข้อมูลภาพ อาจถูกแบ่งได้เป็นสี่ด้านใหญ่ๆ ได้แก่ การแทนค่าข้อมูล (Data Representation) การดึงคุณลักษณะสำคัญของภาพออกมายังการทำการทำด้วย (Feature Extractions and Indexing) การเรียบเรียงภาพ (Image Organization) และ การเชื่อมต่อกับผู้ใช้ (User Interfacing) ปัญหาใหญ่ ประการหนึ่งในหลายประการที่พบในปัจจุบัน ได้แก่ความยากลำบากในการค้นหาหัวข้อมูลภาพที่ต้องการของมาจากที่เก็บภาพขนาดใหญ่ ซึ่งมีภาพอยู่เป็นจำนวนมาก และโดยเฉพาะที่เก็บภาพที่มีภาพหลากหลายประเภทลักษณะที่แตกต่างกัน ทั้งนี้แม้ว่าการค้นหาหัวข้อมูลภาพที่ต้องการจากที่เก็บภาพซึ่งมีภาพจำนวนไม่มากนักนั้นจะทำได้อย่างรวดเร็ว โดยอาศัยวิธีเลือกภาพที่เก็บอยู่ที่ลักษณะ

ไปจนพบภาพที่ต้องการ (browsing) แต่จะเห็นได้ว่าวิธีการดังกล่าว้นั้น นำมาให้กับที่เก็บภาพขนาดใหญ่ที่มีภาพหน้าพันภาพ (หรือกว่านั้น) ไม่ได้ จำเป็นที่จะต้องหาวิธีที่มีประสิทธิภาพมาก กว่าที่นำมาใช้แทน

#### 4. ลักษณะของคำสั่งในการค้นคืนข้อมูลภาพ

โดยทั่วไปชื่อมูลภาพมักถูกใช้ในหมายความอุปประสงค์ อาทิ  
ให้ในการแสดงภาพประกอบนหมวดความต่างๆเพื่อสื่อความหมาย  
ของสารสนเทศหรืออารมณ์ความรู้สึกต่างๆที่ยากแก่การอธิบาย  
ด้วยถ้อยคำ หรืออาจใช้ในการแสดงข้อมูลรายละเอียดเพื่อการ  
วิเคราะห์บางประการ เช่นภาพถ่ายทางรังสีแพทย์ซึ่งช่วยให้แพทย์  
ท่าการวินิจฉัยโรคของผู้ป่วยได้ นอกจากนี้ชื่อมูลภาพยังอาจถูก  
ใช้ในการบันทึกชื่อมูลภาพรอคายแบบเพื่อใช้ในหมายหน้า เช่นภาพ  
แบบแปลนทางสถาปัตยกรรม เป็นต้น ส่วนการค้นหาภาพที่  
ต้องการจากที่เก็บมาพนั้นเมื่อได้หลากรูปแบบขึ้นกับความต้องการ  
ของผู้ใช้ ตัวอย่างเช่นผู้ใช้อาจต้องการภาพที่แสดงฉลากหรือสถาน  
ที่เฉพาะหรือชนิดของวัสดุเฉพาะอย่าง หรืออาจต้องการภาพที่มี  
เห็นแล้วให้ความรู้สึกบางประการ หรืออาจต้องการภาพที่มี  
ลักษณะเฉพาะ เช่นพื้นผิวน้ำหรือสีฟ้าทึบงามอย่าง โดยทั่วไปแล้ว  
ภาพใดๆจะมีคุณลักษณะต่างๆที่สามารถนำมาใช้ในการค้นคืนอยู่  
หลายประการด้วยกัน ซึ่งก่อให้เกิดรูปแบบของคำสั่งที่ใช้ในการ  
ค้นคืน (queries) ได้หลากหลายรูปแบบ ทั้งนี้คำสั่งการค้นคืน  
อาจต้องการให้ค้นคืนชื่อมูลภาพโดยอาศัย

-ลักษณะพื้นฐานบางอย่างของภาพ เช่นสี รูปทรงของวัตถุในภาพ หรือลักษณะพื้นผิวของภาพ (เช่นภาพดาวสีเหลือง ภาพวงกลมสีฟ้า)

-ลักษณะการจัดวางตุ๊กในภาพ (เช่นภาพเก้าอี้วางรอบโต๊ะ ภาพคนอยู่หน้าบ้าน)

-ชื่อของบุคคล วัฒน หรือสถานที่สำคัญในภาพ ( เช่นภาพ  
สถานีรถไฟหัวลำโพง ภาพอัลเบรต ไอน์สไตน์)

-ลักษณะของเหตุการณ์เฉพาะอย่าง (เช่น ภาพการแข่งขันฟุตบอล ภาพการแสดงรำไทย)

- 亚洲的佛教文化与基督教（基督教福音派）  
- 亚洲的佛教文化与基督教（基督教福音派）

-ข้อมูลอื่นๆเกี่ยวกับภาพ เช่นชื่อผู้เขียนหรือบันทึกภาพ วันเวลาที่เขียนหรือบันทึกภาพนั้นๆ (เช่นภาพที่วาดโดยปีคัลลิฟภาพที่ถ่ายโดยแอนน์ เกตส์ในช่วงปีค. 1995-1996)

เมื่อพิจารณาดูรูปแบบของคำสั่งการค้นคืนที่กล่าวมาแล้ว ทั้งหมด จะเห็นว่าทุกรูปแบบยกเว้นรูปแบบสุดท้ายนั้นเป็นเนื้อหา (contents) หรือคุณลักษณะ (features) ที่มีอยู่ในภาพ และแต่ละรูปแบบจะมีความซับซ้อนหรือเป็นนามธรรมมากขึ้นกว่ารูปแบบก่อนหน้านี้เป็นลำดับไป ในการตอบสนองคำสั่งการค้นคืนที่มีความซับซ้อนมากก็ขึ้นนึ้นก็ต้องการความรู้จากภายนอกมา ประกอบมากขึ้นด้วย ในงานเขียนของ Eakins [3] ได้ทำการจำแนกประเภทของคำสั่งการค้นคืนออกเป็นสามระดับ ตามความซับซ้อนที่เพิ่มขึ้นดังต่อไปนี้

ระดับที่ 1 จะเป็นคำสั่งการค้นคืนที่สามารถตอบสนองได้โดยการเรียกคืนข้อมูลที่อาศัยคุณลักษณะพื้นฐานที่ดึงออกมาจากภาพได้โดยตรง คุณลักษณะพื้นฐานนี้หมายรวมถึงสี รูปทรง ลักษณะ พื้นผิว และตำแหน่งของวัตถุภายในภาพ ด้วยอ่างของคำสั่งค้นคืนที่จัดว่าอยู่ในระดับนี้อาจเป็น "จงหาวุปภาพที่สีรูปสีเหลืองสีแดงอยู่มุมล่างซ้ายมือของภาพ" หรือ "ภาพที่ประกอบด้วยวงกลมสีแดงอยู่ในสีเหลืองสีขาว" เป็นต้น การค้นคืนในระดับนี้ใช้คุณลักษณะที่ได้จากการโดยตรง ไม่จำเป็นต้องอาศัยความรู้จากภายนอกใดๆมาประกอบเลย แต่การใช้งานของคำสั่งค้นคืนในระดับนี้ไม่มากนัก และมักจะจำกัดอยู่ในบางลักษณะเฉพาะเท่านั้น เช่นการตรวจสอบว่ามีการใช้เครื่องหมายการค้าช้ากับเครื่องหมายที่เคยขึ้นทะเบียนมาก่อนแล้วหรือไม่ ในงานการรับข้อเท็จจริงนี้ กระบวนการจำแนกประเภทของคำสั่งค้นคืนที่มีความซับซ้อนนี้จะมีความซับซ้อนมากขึ้น

ระดับที่ 2 เป็นการค้นคืนที่อาศัยคุณลักษณะที่สามารถตีความ (derive) ได้จากภาพ ซึ่งต้องใช้การวินิจฉัยทางตรรกะ (logic) การค้นคืนในระดับนี้ยังสามารถถูกแยกย่อยออกเป็นสองระดับ ย่อย ได้แก่

ก. การค้นคืนโดยอาศัยชนิดของวัตถุในภาพ (เช่น ภาพที่มีเด็กผู้ชายกับสุนัข)

ข. การค้นคืนโดยอาศัยเชื่อมโยงของคนหรือวัตถุ หรือสถานที่ในภาพ (เช่น ภาพของน้ำตกในและการ)

คำสั่งการค้นคืนโดยทั่วไป มีแนวโน้มที่จะอยู่ในระดับนี้มากกว่าระดับที่ 1 อย่างไรก็ตาม การจะตอบสนองคำสั่งค้นคืนที่อยู่ในระดับนี้ โดยเฉพาะอย่างยิ่งในระดับ 2 ข. จะต้องมีการอ้างถึงความรู้ภายนอกมาประกอบด้วย ระดับที่ 3 เป็นการค้นคืนที่อาศัยคุณลักษณะที่เป็นนามธรรม (abstract) ซึ่งต้องใช้การตีความระดับสูง จึงจะสามารถทราบความหมายและจุดประสงค์ของวัตถุหรือภาพที่ปรากฏในภาพ และการตีความนั้นมักมีลักษณะที่ขึ้นกับความคิดของแต่ละคน คำสั่งค้นคืนในระดับนี้อาจถูกแบ่งเป็นสองระดับย่อย ดังนี้

ก. การค้นคืนโดยอาศัยลักษณะของเหตุการณ์เฉพาะอย่าง (เช่น ภาพของอุบัติเหตุรถชน)

ข. การค้นคืนโดยอาศัยอารมณ์หรือความรู้สึกที่ภาพเลือก (เช่น ภาพที่ให้ความรู้สึกเศร้า ภาพที่แสดงความรัก) องค์มา (เช่น ภาพที่ให้ความรู้สึกเศร้า ภาพที่แสดงความรัก)

ซึ่งในระดับที่ 3 นี้ การจะตอบสนองคำสั่งค้นคืนได้ต้องใช้การตีความขั้นสูงที่ลับซับซ้อน และมักมีความคิดเห็นส่วนบุคคลเข้ามาร่วมด้วย แม้ว่าคำสั่งสืบคันที่อยู่ในระดับนี้จะพบได้ไม่บ่อยเท่ากับคำสั่งสืบคันในระดับที่ 2 แต่ก็ยังจัดว่าพบได้บ่อยในบางสาขาเช่น เช่นในห้องสมุดภาพ หรือห้องแสดงภาพศิลปะ หรือในงานหนังสือพิมพ์

มีงานวิจัยบางชิ้น เช่น[4] ได้จำแนกประเภทของการค้นคืนข้อมูลภาพออกเป็นสองระดับ ได้แก่ระดับที่ใช้คุณสมบัติพื้นฐาน (primitive features) และ ระดับที่ใช้คุณสมบัติเชิงความหมาย (semantic features) โดยระดับเชิงความหมายนี้จะรวมເວລະດັບທີ 2 ແລະ ອະດັບທີ 3 ຕາມการจำแนกของ Eakins ที่กล่าวถึงข้างต้นเข้าไว้ด้วยกัน

สำหรับการค้นคืนข้อมูลภาพโดยใช้ข้อมูลประกอบอื่นๆ เช่นชื่อผู้เขียนหรือบันทึกภาพนั้น หรือวันเวลาที่บันทึกหรือเกิดเหตุการณ์ในภาพนั้น มีได้ถูกจัดอยู่ในการจำแนกประเภทที่ได้กล่าวมาแล้วข้างต้น เมื่อจากข้อมูลเหล่านี้มีลักษณะเป็นข้อความ (text) ที่ผู้ใช้ต้องป้อนเข้าเก็บในระบบของ ซึ่งข้อมูลดังกล่าวสามารถถูกจัดการได้โดยใช้ระบบค้นคืนข้อมูล (text retrieval systems) ทั่วไป

## 5. เทคโนโลยีการค้นคืนข้อมูลภาพโดยอาศัยเนื้อหาของภาพ

CBIR เป็นคำย่อของ Content-Based Image Retrieval หรือการค้นคืนข้อมูลภาพโดยอาศัยเนื้อหาของภาพ ค่าว่า CBIR ถูกใช้เพื่อธุรกิจกระบวนการในการค้นคืนข้อมูลภาพที่ต้องการจากที่เก็บภาพจำนวนมากโดยใช้คุณลักษณะต่างๆ ของภาพ (เช่น สี รูปทรงของวัตถุในภาพ และลักษณะพื้นผิว) ซึ่งสามารถถูกดึงหัวใจตรวจสอบจากภาพได้โดยอัตโนมัติ ซึ่งคุณลักษณะต่างๆ ของภาพนั้นอาจเป็นคุณลักษณะขั้นพื้นฐาน (primitive features) หรืออาจเป็นคุณลักษณะในเชิงความหมาย (semantic features) ก็ได้ อย่างไรก็ตาม กระบวนการในการค้นคืนคุณลักษณะนี้จะพยายามที่ต้องทำให้โดยอัตโนมัติ ไม่ต้องอาศัยความช่วยเหลือจากมนุษย์ ดังนั้นระบบการค้นคืนข้อมูลภาพแบบดังเดิมที่ใช้ค่าสำคัญที่มีนูนย์เป็นผู้ป้อนเข้าไปล่วงหน้าก่อนหนึ่งไม่จำต้องเป็น CBIR เมื่อว่าที่จริงแล้วค่าสำคัญที่สามารถอธิบายเนื้อหาของภาพได้ก็ตาม

CBIR ต่างจากการบันการค้นคืนสารสนเทศแบบดั้งเดิมซึ่งเป็นการค้นคืนข้อมูลความลับ (text-based systems) อย่างลึกซึ้ง เนื่องมาจากฐานข้อมูลภาพนั้นมีลักษณะที่ไม่มีโครงสร้าง (unstructured) ในขณะที่ฐานข้อมูลที่เก็บข้อมูลความนัยการจัดเก็บอย่างมีโครงสร้าง (structured) ทั้งนี้เป็นเพราะภาพดิจิตอลนั้นประกอบด้วยชุดของจุดภาพ (pixels) นับล้านจุดที่มีความเข้มข้นแห่งสีต่างกัน ทว่าแต่ละจุดภาพนั้นไม่ได้มีความหมายใดๆ โดยตัวของมันเอง จุดประสงค์หลักก็คือหนึ่งของการทำการประมวลผลข้อมูลภาพ คือการที่จะสามารถดึงเอกสารสนเทศที่มีประโยชน์ออกมาจากตัวข้อมูลเดิมนั้นได้ (เช่นสามารถบอกได้ว่าในภาพนั้นมีรูปทรง หรือลักษณะพื้นผิวที่ต้องการอยู่หรือไม่) ก่อนที่จะสามารถนำสารสนเทศนั้นไปตีความเป็นเนื้อหาของภาพในระดับสูงขึ้นไปได้ ฐานข้อมูลรูปภาพเจ้มที่ลักษณะการจัดการที่ต่างไปจากฐานข้อมูลข้อความซึ่งตัวข้อมูลเดิมอยู่ในรูปของคำ (ซึ่งประกอบขึ้นจากชุดตัวอักษรที่แทนด้วยรหัส ASCII) ที่ถูกจัดเก็บอย่างมีโครงสร้างทางตรรกะโดยผู้เชี่ยวชาญตั้งแต่แรกแล้ว [Santini and Jain, 1997] ในฐานข้อมูลข้อความเจ้มไม่มีการค้นคืนในระดับพื้นฐาน (หรือระดับที่ 1 ถ้าเทียบกับฐานข้อมูลภาพ)

วิธีการที่ใช้ในเทคโนโลยี CBIR ส่วนมากได้มาจากสาขาวิชาการประมวลผลข้อมูลภาพ (Image Processing) และคอมพิวเตอร์วิชัน (Computer Vision) จึงเป็นเหตุให้เกิดวิชาการจำนวนหนึ่งจัดว่า CBIR นั้นเป็นแขนงหนึ่งของสาขาวิชาดังกล่าว

อย่างไรก็ตาม CBIR นั้นจะมุ่งเน้นเฉพาะเรื่องของการค้นคืนข้อมูลภาพที่มีคุณลักษณะตรงตามที่ต้องการจากที่เก็บภาพจำนวนมาก ในขณะเดียวกันการประมวลผลข้อมูลนั้นจะสนใจในหลากหลายหัวข้อ เช่นการปรับรูปภาพ (Image Enhancement) การย่อขนาดไฟล์ข้อมูลภาพ (Image File Compression) และการแปลงรูปของข้อมูลภาพ (Image Transmission) เป็นต้น

หัวข้อที่มีการที่วิจัยและพัฒนาในด้าน CBIR มีอยู่มาก หลายหัวข้อมีความสัมพันธ์ใกล้ชิดกับสาขาวิชาการประมวลผลข้อมูลภาพ และสาขาวิชาการค้นคืนสารสนเทศด้วย ด้วยอย่างของหัวข้อวิจัยที่สำคัญได้แก่

- การเข้าใจความต้องการ และพฤติกรรมการค้นหาของผู้ใช้ข้อมูลภาพ

- การหาวิธีการที่เหมาะสมในการอธิบายเนื้อหา (contents) ของภาพ

- การดึงคุณลักษณะที่อาจใช้ในการค้นคืนข้อมูลภาพออกมากจากภาพ

- การหาวิธีจัดเก็บฐานข้อมูลภาพขนาดใหญ่โดยใช้เนื้อที่เก็บอย่างมีประสิทธิภาพ

- การตอบสนองค่าสั่งค้นคืนให้ได้มากที่สุดโดยการค้นหาของมนุษย์มากที่สุด

- การออกแบบระบบเชื่อมต่อผู้ใช้ (user interface) ที่มีประสิทธิภาพ

ตัวอย่างของการนำเทคโนโลยี CBIR ไปประยุกต์ใช้ในชีวิตจริงนั้นมีมากมาย ดังต่อไปนี้

- การป้องกันและปราบปรามอาชญากรรมของกรมตำรวจ เช่น ระบบเบรย์บลายด์มือ ระบบเบรย์บลายด์ในหน้าผู้ต้องสงสัย ฯลฯ

- การป้องกันการละเมิดทรัพย์สินทางปัญญา เช่นการขึ้นทะเบียนเครื่องหมายการค้า

- การแพทย์ อาจใช้เทคโนโลยี CBIR ช่วยในการวินิจฉัยโรค โดยค้นหาข้อมูลภาพของกรณีศึกษาในอดีตที่มีลักษณะอาการของโรคใกล้เคียงกัน

- การค้นหาข้อมูลภาพที่ต้องการจากที่เก็บภาพจำนวนมหาศาล เช่นใน WWW

## 6. เทคนิค CBIR ที่ใช้เพื่อค้นคืนข้อมูลภาพในปัจจุบัน

## 6.1 คุณลักษณะพื้นฐาน (Primitive Level)

จากที่ได้กล่าวมาในหัวข้อก่อนนี้แล้วว่าการทำงานของระบบที่ใช้เทคนิค CBIR นั้นแตกต่างไปจากเทคนิคการค้นคืนข้อมูลโดยใช้ข้อความ (text-based systems) ในระบบดังเดิมอย่างสั้นเริ่ง ทั้งนี้เพราะระบบที่ใช้ CBIR จะค้นคืนข้อมูลภาพที่ต้องการอุดมมาจากที่เก็บภาพโดยการเปลี่ยนเทียบคุณลักษณะที่ถูกดึงออกมาจากภาพได้โดยอัตโนมัติ ซึ่งคุณลักษณะของภาพที่ใช้กันมากที่สุด ได้แก่ค่าในเชิงคณิตศาสตร์ของลักษณะพื้นผิว และรูปทรงของตัวถุที่ปรากฏในภาพ นอกจากนี้ยังมีคุณลักษณะพื้นฐานอื่นๆ อีก เช่น ตำแหน่งและระยะห่างของวัตถุในภาพ (spatial location) หรือการเปลี่ยนรูปของจุดภาพต่างๆ ในภาพ (complex transformations of pixel intensities) โดยระบบ CBIR ในปัจจุบันเกือบทั้งหมดจะค้นคืนข้อมูลภาพโดยใช้คุณลักษณะพื้นฐานเหล่านี้ ไม่ว่าจะเป็นระบบ CBIR ที่มีงานวิจัยอย่าง Photobook ของ MIT [6] หรือระบบเชิงพาณิชย์อย่าง QBIC ของบริษัท IBM [7],[8] หรือ Visual RetrievalWare ของบริษัท Excalibur Technologies [9] ซึ่งระบบ CBIR โดยทั่วไป มากใช้คุณลักษณะพื้นฐานหลายประการ ที่กล่าวมาแล้วนี้ร่วมกัน ใน การค้นคืนข้อมูลภาพกว่าที่จะเลือกใช้คุณลักษณะพื้นฐานอย่างใดอย่างหนึ่งเพียงประการเดียว นอกจากนี้ระบบ CBIR สามารถจัดอันดับให้ผู้ใช้สังเคราะห์ในการค้นคืน (queries) ในรูปของตัวอย่างภาพที่ผู้ใช้ต้องการ ซึ่งอาจเป็นภาพเมื่อวันจริง หรือภาพที่เกิดจากการวัดโดยเส้นกีดี จากนั้นระบบจะทำการค้นหาในที่เก็บข้อมูลภาพดูว่ามีภาพใดที่มีคุณลักษณะพื้นฐานใกล้เคียงกับภาพตัวอย่างมากที่สุด ก็จะแสดงภาพขนาดเล็ก (thumbnails) ของภาพหรือกลุ่มภาพผลลัพธ์ (ระบบอาจให้คำตอบมากกว่าหนึ่งภาพ เช่น อาจเลือกภาพ 10 หรือ 20 ภาพที่มีลักษณะใกล้เคียงกับภาพตัวอย่างมากที่สุด) ออกมาที่หน้าจอเพื่อให้ผู้ใช้เลือกภาพที่ตรงกับความต้องการของตนมากที่สุดไปใช้งานต่อไป

### 6.1.1 การค้นคืนโดยอาศัยความคล้ายคลึงของสี (Colour Retrieval)

มีงานวิจัยหลายชิ้นที่ทำเสนอวิธีการทำ การค้นคืนภาพโดยอาศัยความคล้ายคลึงของสี [10], [11],[12] ซึ่งเน้นรายละเอียดของวิธีการที่ใช้จะแตกต่างกัน แต่คงอยู่ภายใต้แนวคิดพื้นฐานเดียวกัน นั่นคือ ก่อนหนีข้อมูลภาพแต่ละภาพจะถูกจัด

เก็บลงในที่เก็บข้อมูลภาพเหล่านั้นจะถูกวิเคราะห์และคำนวณทาง histogram สี (Colour Histogram) ซึ่งเป็นชุดของตัวเลขที่แสดงสัดส่วนของจุดภาพแต่ละสีที่มีอยู่ในภาพนั้นๆ จากนั้นอัลกอริทึมสีของแต่ละภาพจะถูกจัดเก็บลงในฐานข้อมูลเพื่อใช้สำหรับมีนัดชั้นในการค้นหา เมื่อต้องการค้นหาภาพที่ต้องการ ผู้ใช้อาจสังเคราะห์ในการค้นคืนข้าสู่ระบบได้สองวิธี วิธีแรกคือจะส่งตัวเลขค่าสัดส่วนของแต่ละสีในภาพที่ต้องการ (เช่น ให้ค้นหาภาพที่มีสีแดง 30 % และสีเหลือง 70%) ส่วนอีกวิธีหนึ่งคือส่งภาพตัวอย่างเข้ามาเพื่อให้ระบบค้นหาภาพที่มีโภนสีคล้ายคลึงกับภาพตัวอย่าง ซึ่งสำหรับวิธีที่สองนี้ ระบบจะต้องทำการคำนวณเพื่อหาอัลกอริทึมสีของภาพตัวอย่างก่อน จากนั้นໄ่าว่าจะใช้ค่าสัดส่วนคืนในรูปแบบใด ระบบก็จะต้องมีการเปลี่ยนเทียบค่าตัวเลขที่ได้จากการสังเคราะห์คืน กับค่าตัวเลขในอัลกอริทึมสีของแต่ละภาพที่เก็บอยู่แล้วจึงทำการเรียงค้นภาพที่มีความคล้ายคลึงของกลุ่มสีกับสั่งค้นคืนมากที่สุดอีกมาส่องเป็นผลลัพธ์ ซึ่งผลลัพธ์ที่ได้จะจากระบบค้นคืนข้อมูลภาพที่ใช้เทคนิคนี้ส่วนมากจะให้ผลที่ตรงความต้องการของผู้ใช้ค่อนข้างมาก

### 6.1.2 การค้นคืนโดยอาศัยความคล้ายคลึงของลักษณะพื้นผิว (Texture Retrieval)

ความสามารถในการค้นคืนข้อมูลภาพโดยอาศัยความคล้ายคลึงของลักษณะพื้นผิวของภาพแต่เพียงอย่างเดียวันอาจให้ผลลัพธ์ที่ไม่น่าพอใจนัก แต่มีอีกหนึ่งวิธีที่ใช้ร่วมกับการค้นคืนโดยอาศัยความคล้ายคลึงของสีแล้ว ซึ่งช่วยให้ผลลัพธ์จากการค้นคืนได้ผลดีขึ้นเป็นอันมาก เมื่อจากลักษณะพื้นผิวจะช่วยให้ระบบแยกภาพของสิ่งที่แตกต่างกันแต่สีใกล้เคียงกันออกจากกันได้มีน้อยดี (ตัวอย่างเช่นภาพของห้องพักกับภาพห้องนอนมีสีฟ้าคล้ายกัน แต่ลักษณะพื้นผิวต่างกัน หรือภาพของใบไม้กับภาพของสนมหาดูเหมือนกัน แต่ลักษณะพื้นผิวต่างกัน) ทั้งนี้การคำนวณลักษณะพื้นผิวของภาพ อาจจะออกมากในรูปค่าตัวเลขของความสว่าง (brightness) ความหยาบ (coarseness) กิจทาง (directionality) ของแต่ละจุดที่ประกอบขึ้นเป็นต้น

นอกจากนี้ ในงานวิจัยบางชิ้นยังมีการนำเอาเทคนิคของแฟร์คอล (fractals) [13] หรือการเบอร์ฟิลเตอร์ (Gabor Filters) [14] มาใช้ในการวิเคราะห์ลักษณะพื้นผิวอีกด้วย

คำสั่งค้นคืนข้อมูลภาพในระบบคันคืนโดยอาศัยความคล้ายคลึงของลักษณะพื้นผิวนี้ จะมีลักษณะใกล้เคียงกับคำสั่งคันคืนโดยอาศัยความคล้ายคลึงของลีส์ นั่นคืออาจอนุญาตให้ผู้ใช้ใส่ภาพตัวอย่างที่ต้องการมาในคำสั่งคันคืน หรืออาจมีตัวอย่างลักษณะพื้นผิวหลากหลายแบบมาให้ผู้ใช้เลือกแบบที่ต้องการก็ได้

### 6.1.3 การค้นคืนโดยอาศัยความคล้ายคลึงของรูปทรง (Shape Retrieval)

ในงานวิจัยของ Biederman [15] สรุปไว้ว่าวัตถุในธรรมชาติโดยทั่วไปแล้วมักจะถูกแยกแยะหรือรู้จ้าได้โดยรูปทรงของมัน ในกระบวนการของการค้นคืนข้อมูลภาพโดยอาศัยความคล้ายคลึงของรูปทรงของวัตถุในภาพนั้นจะทำได้ก็โดยต้องมีการคำนวนหาค่าของคุณลักษณะบางประการของรูปทรง (โดยไม่เขียนกับมุมมองหรือขนาด) ของวัตถุแต่ละชิ้นในแต่ละภาพที่เก็บอยู่ อย่างเช่น aspect ratio สำหรับคำสั่งการค้นคืนจะถูกตอบสนองได้โดยการคำนวนค่าของคุณลักษณะชุดเดียวกันนั้นจากภาพตัวอย่างในคำสั่งคันคืน เพื่อนำมาเปรียบเทียบกับข้อมูลภาพทั้งหมดที่เก็บอยู่ แล้วจึงดึงกลุ่มของภาพที่มีค่าคุณลักษณะใกล้เคียงกับภาพตัวอย่างมากที่สุดขึ้นมาเป็นผลลัพธ์ [16] นอกจากนี้ยังอาจใช้วิธีอื่นๆ อีก การเปรียบเทียบอิสโตแกรมพิกัดทาง (directional histogram) ของขอบของรูปทรงที่ดึงออกมายังไงจากภาพ [17] หรือการใช้การแทนค่ารูปทรงหนึ่งด้วยแกนของรูปทรงนั้นๆ (skeletal representations) ซึ่งข้อมูลแกนนี้สามารถนำมาระบบเปรียบเทียบกันได้โดยใช้เทคนิคการเปรียบเทียบกราฟ [18], [19]

สำหรับคำสั่งคันคืนข้อมูลภาพในระบบคันคืนโดยอาศัยความคล้ายคลึงของรูปทรงนี้ อาจอยู่ในรูปของภาพตัวอย่างที่เป็นภาพจริง หรืออาจอยู่ในรูปของภาพวาดลายเส้น (sketch) จากผู้ใช้ก็ได้ [20]

### 6.2 คุณลักษณะในเชิงความหมาย (Semantic Level)

ระบบคันคืนข้อมูลภาพที่ใช้เทคนิค CBIR ในปัจจุบัน ส่วนใหญ่จะถูกออกแบบมาให้ใช้คันคืนข้อมูลภาพโดยอาศัยลักษณะพื้นฐานเท่านั้น อย่างไรก็ตามนักวิจัยจำนวนหนึ่งก็ได้พยายามที่จะนำคุณลักษณะเชิงความหมายมาใช้ในการค้นคืนข้อมูล เพื่อให้ได้ระบบคันคืนข้อมูลที่ตอบสนองความต้องการของผู้ใช้มากขึ้น ตัวอย่างเช่นระบบ GRIM\_DBMS [21] ซึ่งถูกออกแบบมาให้สามารถแปลงและเรียกคันภาพลายเส้นของวัตถุบาง

อย่างที่อยู่ในขอบเขตคำจำกัดที่ได้กำหนดไว้ล่วงหน้า อย่างเช่นผังการจัดห้องสำหรับอาคารที่อยู่อาศัย เป็นต้น ระบบจีวิเคราะห์ภาพคาดคะเนวัตถุในภาพ แล้วทำการให้เชื่อมต่อแต่ละชิ้นเท่าที่ระบบจะสามารถตีความໄດ้ ซึ่งชื่อที่ได้มานั้นจะช่วยให้ระบบสามารถตีความได้ว่าภาพที่มีวัตถุดังกล่าวอยู่นี่เป็นภาพอะไร

สำหรับในช่วงปัจจุบันนี้ งานวิจัยในระดับนี้มักมุ่งเน้นที่จะแก้ปัญหาสองประการด้วยกัน ประการแรกคือการรู้จำภาพหรือสถานที่ที่ปรากฏในภาพ (scene recognition) ตัวอย่างเช่นระบบ IRIS [22] ซึ่งใช้ลักษณะพื้นฐานอย่างข้อมูลลีส์ พื้นผิวและตำแหน่งของพื้นที่ในภาพ เพื่อช่วยในการแปลความหมายของฉากที่ปรากฏในภาพ แล้วทำการสรุปเป็นข้อความบรรยายภาพนั้น (เช่นห้องพัก ทะเล ป่า เป็นต้น)

ส่วนปัญหาอีกประการหนึ่งที่มีงานวิจัยอุปกรณามีงบอยู่ได้แก่การรู้จั vwatsttu ในภาพ (object recognition) ซึ่งอาจนำความรู้ด้านคอมพิวเตอร์วิชั่นมาประยุกต์ใช้ร่วมด้วย ในปัจจุบันมีการคิดหาเทคนิคมาหลายเพื่อการรู้จ้าและแยกแยะวัตถุ เช่นงานของ Forsyth และคณะ [23] ซึ่งพัฒนาเทคนิคในการรู้จ้าและจำแนกภาพร่างเป็นลีอยของมนุษย์ และเทคนิคดังกล่าวถูกนำมาประยุกต์ใช้กับการรู้จั vwatsttu ของมนุษย์ แล้วม้า และต้นไม้ เทคนิคดังๆ เหล่านี้มีกระบวนการพื้นฐานได้แก่การสร้างรูปแบบ (model) ของประเภท (class) ต่างๆของวัตถุ จากนั้นจีวิเคราะห์หาส่วนของภาพที่น่าจะมีวัตถุอยู่ และพิสูจน์ว่าในส่วนนั้นของภาพมีวัตถุอยู่จริงหรือไม่ และตัดสินใจว่าจะ

อย่างไรก็ตามงานวิจัยต่างๆที่มีในปัจจุบันเกี่ยวกับการค้นคืนโดยใช้คุณลักษณะเชิงความหมายนั้นยังคงจำกัดอยู่ในขอบเขตเฉพาะบางด้านเท่านั้น ยังคงมีความต้องการที่จะทำกระบวนการอัตโนมัติได้ก้าวไปสู่ความสามารถที่จะสามารถทำได้ทันทีและค้นคืนข้อมูลภาพโดยอาศัยชีวิดของวัตถุหรือหากที่ปรากฏในภาพ

## 7. บทสรุป

ในปัจจุบันนี้ CBIR ยังคงเป็นหัวข้อที่นักวิจัยให้ความสนใจเป็นอันมาก เทคโนโลยีนี้ยังคงต้องการการศึกษาพัฒนาเพิ่มเติมอีกมาก ทั้งยังไม่ถูกนำไปประยุกต์ใช้ในวงกว้างเท่าใดนัก สาเหตุประการหนึ่งอาจเนื่องมาจากในระบบคันคืนข้อมูลภาพที่ใช้เทคโนโลยี CBIR ในปัจจุบันส่วนใหญ่ยังใช้เทคนิคที่อาศัยเฉพาะคุณลักษณะพื้นฐานของภาพเท่านั้น ทำให้ไม่เหมาะสมกับลักษณะ

งานและความต้องการของผู้ใช้โดยทั่วไป จึงยังคงมีความต้องการที่จะพัฒนาเทคโนโลยี CBIR ที่สามารถอ่านคุณลักษณะเชิงความหมายของภาพให้มีประสิทธิภาพและใช้ได้ในหลากหลายลักษณะงานต่อไป ทั้งนี้เพื่อยังผลให้สามารถปั้นรูปแบบค้นหานิรุ่มภาพให้มีประสิทธิภาพดีและเหมาะสมกับผู้ใช้งานหลากหลายประเภทและสาขาวิชาซึ่งมากยิ่งขึ้น

## 8. บรรณานุกรม

- [1] Jain, R., "World-wide maze" IEEE Multimedia 3), p.3,1995.
- [2] Enser, P. G. B., Pictorial Information Retrieval, Journal of Documentation , 51(2), pp.126-170, 1995.
- [3] Eakins, J.P., Automatic Image Content Retrieval - Are We Getting Anywhere?, Proceedings of Third International Conference on Electronic Library and Visual Information Research (ELVIRA3) , De Montfort University, Milton Keynes, pp.123-135,1996.
- [4] Gudivada V. N. and Raghavan V. V., Content-Based Image Retrieval Systems, IEEE Computer 8(9), pp.18-22, 1995.
- [5] Santini, S. and Jain, R. C., The Graphical Specification of Similarity Queries, Journal of Visual Languages and Computing 7, pp. 403-421, 1997.
- [6] Pentland A. et. al., Photobook: Tools for Content-Based Manipulation of Image Databases, International Journal of Computer Vision 18(3), pp.233-254, 1996.
- [7] Flickner, M. et. al., Query by Image and Video Content: the QBIC System, IEEE Computer 28 (9), pp.23-32, 1995.
- [8] Niblack, W. et. al., Updates to the QBIC System, in Storage and Retrieval for Image and Video Databases VI (Sethi, I K and Jain, R C, eds), Proc SPIE 3312, pp.150-161, 1998.
- [9] Feder, J., Towards Image Content-Based Retrieval For The World-Wide Web, Advanced Imaging 11(1), pp.26-29, 1996.
- [10] Swain, M J and Ballard, D H., Color Indexing, International Journal of Computer Vision 7(1), pp.11-32,1991.
- [11] Carson, C.S. et. al., Region-Based Image Querying, Proceedings of IEEE Works Content-Based Access of Image and Video Libraries, San Juan, Puerto Rico, pp.42-49, 1997.
- [12] Stricker, M. and Dimai, A., Color Indexing With Weak Spatial Constraints, in Storage and Retrieval for Image and Video Databases IV, (Sethi, I K and Jain, R. C. eds), Proc SPIE 2670, pp.29-40,1996.
- [13] Kaplan, L. M. et. al., Fast texture database retrieval using extended fractal features, in Storage and Retrieval for Image and Video Databases VI (Sethi, I. K. and Jain, R. C., eds), Proc SPIE 3312, p p.162-173, 1998.
- [14] Manjunath, B. S. and Ma, W. Y., Texture Features for Browsing and Retrieval of Large Image Data, IEEE Transactions on Pattern Analysis and Machine Intelligence, 18, pp.837-842, 1996.
- [15] Biederman, I., Recognition-by-Components: A Theory of Human Image Understanding, Psychological Review 94(2), pp.115-147, 1987.
- [16] Mehrotra, R. and Gary, J. E., Similar-Shape Retrieval in Shape Data Management, IEEE Computer 28(9), pp.57-62, 1995 .
- [17] Androutsas, D. et. al., Image Retrieval Using Directional Detail Histograms, in Storage and Retrieval for Image and Video Databases VI, Proc SPIE 3312, pp.129-137, 1998 .

- [18] Kimia, B. B. et. al., A Shock-Based Approach for Indexing of Image Databases Using Shape, in *Multimedia Storage and Archiving Systems II* (Kuo, C C J et al, eds), Proc SPIE 3229, p.288-302, 1997.
- [19] Tirthapura, S. et. al., Indexing Based on Edit-Distance Matching of Shape Graphs, in *Multimedia Storage and Archiving Systems III* (Kuo, C C J et al, eds), Proc SPIE 3527, pp.25-36, 1998.
- [20] Chan, Y. and Kung, S.Y., A Hierarchical Algorithm for Image Retrieval by Sketch, First IEEE Workshop on Multimedia Signal Processing, pp.564-569, 1997 .
- [21] Rabbitti, F. and Stanchev, P., GRIM\_DBMS: A Graphical Image Database Management System, in *Visual Database Systems* (Kunii, T. ed), Elsevier, Amsterdam, pp.415-430, 1989.
- [22] Hermes, T. et. al., Image Retrieval for Information Systems, in *Storage and Retrieval for Image and Video Databases III* (Niblack, W R and Jain, R. C. eds), Proc SPIE 2420, pp.394-405,1995.
- [23] Forsyth, D. A. et. al., Finding Pictures of Objects in Large Collections of Images, in *Digital Image Access and Retrieval: 1996 Clinic on Library Applications of Data Processing* (Heidorn, P B and Sandore, B, eds), pp.118-139, Graduate School of Library and Information Science, University of Illinois at Urbana- Champaign, 1997.