การสกัดคุณลักษณะข้อมูลพื้นฐานสำหรับระบบค้นคืนข้อมูลภาพ

A Study of Low Level Feature Extraction Techniques Used for Content-Based Image Retrieval System

นัศพ์ชาณัณ ชินปัญช์ธนะ*

บทคัดย่อ

ปัจจุบันภาพดิจิทัลมีจำนวนเพิ่มมากขึ้นทำให้เกิดการพัฒนาเป็นระบบค้นคืนภาพ ขึ้นมานานกว่าสิบปีในสาขาคอมพิวเตอร์วิชันปัญหาส่วนใหญ่ของระบบค้นคืนภาพคือวิธีการสกัดข้อมูลคุณลักษณะจากภาพอย่างไร เพื่อให้สามารถแทนเนื้อหาในภาพได้อย่างมีประสิทธิภาพ ซึ่งการสกัดข้อมูลภาพนั้นจำเป็นจะต้องได้คุณลักษณะของภาพที่สมบูรณ์ เพื่อนำมาใช้ในการค้นคืนภาพต่อไป บทความได้สรุปและทบทวนเนื้อหาพื้นฐานเทคนิคต่างๆ ของระบบค้นคืนภาพ โดยแบ่งออกเป็น 4 ส่วนดังนี้ ส่วนที่ 1 เป็นการแนะนำระบบการค้นคืนภาพโดยรวม ส่วนที่ 2 เป็นการทบทวนคุณลักษณะข้อมูลภาพ เช่น สี, รูปทรง, ลวดลาย เป็นต้น ส่วนที่ 3 อธิบายถึง การวัดความเหมือนของการค้นคืนภาพ และส่วนที่ 4 ทิศทางของระบบค้นคืนภาพ จะเห็นว่าปัญหายังไม่สามารถแก้ไขได้ด้วยการใช้คุณลักษณะข้อมูลพื้นฐานทั้งหมดทำให้นักวิจัยต้องพยายามหาวิธีการที่เหมาะสมเพื่อพัฒนาระบบการค้นคืนภาพต่อไป

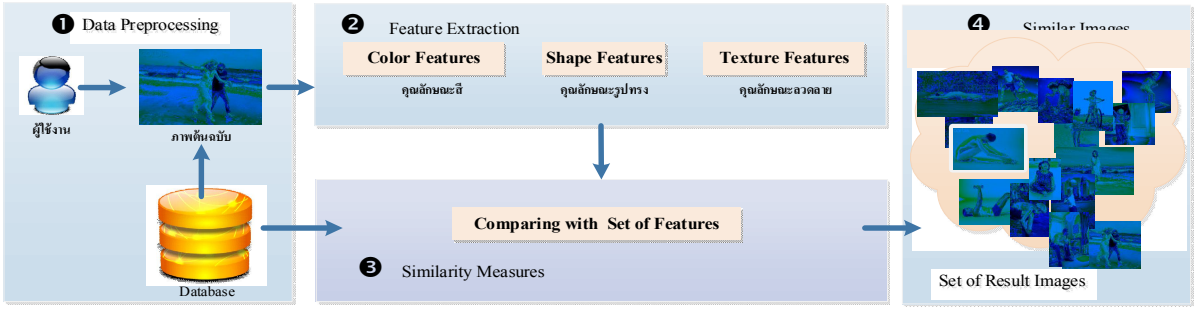
Abstract

Due to the enormous increase indigital image sizes, many Content Based Image Retrieval (CBIR) systems have been developedover the last 10 yearsin the field of computer vision. Most current issue in CBIR is to extract the set of features that effectively way to represent the image contents. Such afeature extraction requires a detailed evaluation of retrieval performance of image features.The articleis composed of 4 sections: Section 1 gives an introduction into the system of CBIR. Section 2 gives a review of available literature infundamental aspects of content based image retrieval. CBIR method uses various visual features of images such as color, shape, textureetc. Section 3 explainsthe various approachesof similarity measures for the matching of image retrieval. Last section is concluded the possible research directions in CBIR. The problem still remains unsolved, many research attempting to find a suitable method for developingthe system of image retrieval. Since this problem remains unsolved, many research in image retrieval attempt to find the new framework to characterize the image content with higher level semantics, closer to that familiar to the user in mind.

*ผู้ช่วยศาสตราจารย์ สาขาเทคโนโลยีสารสนเทศ วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์

บทนำ

วิวัฒนาการของเทคโนโลยีสมัยใหม่ถูกพัฒนาไปอย่างรวดเร็ว ทำให้แอพพลิเคชันต่างๆ ถูกพัฒนาขึ้นเพื่อตอบสนองความต้องการของผู้ใช้งานอย่างหลากหลาย แต่ละแอพพลิชันมีวัตถุประสงค์เพื่ออำนวยความสะดวกต่อผู้ใช้งานทั้งสิ้น รวมทั้งแอพพลิเคชันสำหรับการค้นคืนข้อมูลภาพดิจิทัลเป็นอีกหนึ่งแอพพลิเคชันที่ถูกสร้างขึ้นมา เพื่อตอบสนองความต้องการของปริมาณภาพดิจิทัลที่เพิ่มมากขึ้นอย่างรวดเร็ว ปัญหาที่ตามมาก็ คือ การจัดเก็บข้อมูลภาพที่เพิ่มมากขึ้นอย่างไร้ขีดจำกัดนี้และสามารถสืบค้นข้อมูลภาพ และจำแนกข้อมูลภาพให้ตรงตามความหมายของภาพที่ต้องการของผู้ใช้มากที่สุดได้อย่างไร ทำให้งานวิจัยในปัจจุบันที่เกี่ยวข้องกับการค้นคืนรวมทั้งการจัดกลุ่มภาพให้ตรงกับความต้องการที่เพิ่มมากขึ้น รวมไปถึงระยะเวลาในการสืบค้นที่น้อยลงกับปริมาณของภาพที่เพิ่มทวีคูณ ทำให้ผู้ใช้ประสบปัญหาในการค้นคืนภาพที่ตรงกับความต้องการดังนั้นจึงเกิดความต้องการระบบที่ จะสามารถทำหน้าที่ในการค้นคืนภาพให้ตรงตามความต้องการของผู้ใช้ได้ที่เรียกว่า ระบบค้นคืนภาพด้วยลักษณะเฉพาะ ของภาพ (Content Based Image Retrieval: CBIR) ซึ่งได้รับความสนใจจากนักวิจัยหลายกลุ่มในสาขาการประมวลผลภาพ (Image processing) ด้านการค้นคืนภาพ (Image retrieval) เพื่อคัดเลือกภาพให้ตรงตามความต้องการของผู้ใช้งาน บทความนี้ได้กล่าวถึงทฤษฎีที่เกี่ยวข้องกับระบบค้นคืนภาพ (CBIR) ด้วยคุณลักษณะเฉพาะของภาพ (Image features) เช่น รูปร่าง, ขนาด, สี หรือ พื้นผิว เป็นต้น (Adnan A, 2010) (B. S. Manjunath, 1996) ในการจัดเก็บและค้นคืนภาพโดยระบบจะทำการแยกลักษณะ เฉพาะของภาพแต่ละภาพในฐานข้อมูลออกมา ซึ่งจะอยู่ในรูปของค่าตัวเลขที่เรียกว่าเวกเตอร์คุณลักษณะเฉพาะ (Feature vector) เพื่อใช้เป็นดัชนีภาพสำหรับการจัดเก็บภาพในฐานข้อมูล เมื่อต้องการค้นคืนภาพ ผู้ใช้จะทำการเลือกภาพสอบถาม จากนั้นระบบจะทำการดึงเวกเตอร์ลักษณะเฉพาะของภาพสอบถามมาทำการเปรียบเทียบกับเวกเตอร์คุณลักษณะเฉพาะของภาพที่อยู่ในฐานข้อมูล เพื่อค้นหาภาพที่มีลักษณะเฉพาะคล้ายกับภาพสอบถาม เพื่อแสดงผลลัพธ์ของการค้นคืนภาพที่ใกล้เคียงมาก ที่สุดเป็นชุดภาพผลลัพธ์



ภาพที่ 1 ระบบค้นคืนภาพด้วยคุณลักษณะเฉพาะ

จากภาพที่ 1 แสดงวิธีการค้นคืนภาพเริ่มต้นจากการกำหนดภาพต้นแบบที่ต้องการจากผู้ใช้งาน เข้าสู่ระบบการค้นคืนภาพเพื่อผ่านกระบวนการสกัดข้อมูลภาพตามแต่ละอัลกอริทึมเพื่อให้ได้คุณลักษณะของภาพต้นแบบ และใช้คุณลักษณะเพื่อวัดความเหมือนกันของภาพจากฐานข้อมูลภาพ ผลลัพธ์จะแสดงให้ผู้ใช้งานเป็นกลุ่มภาพที่ถูกเปรียบเทียบด้วยค่าความคล้ายกันของภาพต้นแบบตามคุณลักษณะเฉพาะที่ได้มา และทำการส่งภาพคืนเป็นชุดตามค่าความเหมือนกันของภาพ ระบบค้นคืนภาพประกอบด้วย 2 ส่วนที่สำคัญ [Chen, C, 2015) (Divya Srivastava, 2015) (1) การสกัดข้อมูลภาพ (Feature extraction) การใช้ข้อมูลที่ได้จากคุณลักษณะของพิกเซลภายในภาพ (Visual information) คือ นำภาพมาสกัดข้อมูลคุณลักษณะภายในพิกเซลต่างๆ (Image features) โดยคำนึงถึงการแทนค่าข้อมูลภาพด้วยตัวเลข ซึ่งวิธีการนี้จะเป็นการนำคุณลักษณะของสี ลวดลาย และคุณลักษณะของรูปทรง ที่มีความสัมพันธ์ภายในภาพมาใช้ และ (2) การวัดความคล้าย (Similarity measure) เป็นการวัดความ

คล้ายกันของภาพสองภาพด้วยการวิเคราะห์ทางตัวเลขจากคุณลักษณะภายในภาพที่สกัดมาซึ่งการวัดความคล้ายนี้ไม่ได้ขึ้นอยู่กับมนุษย์แต่เกิดจากการเปรียบจากตัวเลขเท่านั้น

การสกัดข้อมูลภาพ (Feature extraction)

สำหรับงานวิจัยทางการประมวลผลภาพในการค้นคืนสารสนเทศกลุ่มแรกๆ จะมีการค้นคืนตามคุณลักษณะพื้นฐานของภาพที่ถูกสกัดคุณลักษณะภาพด้วยอัลกอริทึมต่างๆ เช่น สี (Color) ลวดลาย (Texture) รูปทรง (Shape) เป็นต้น (B. S. Manjunath, 1996) (Chen, C, 2015) ได้กล่าวโดยย่อไว้ ดังนี้

1. คุณลักษณะสี

คุณลักษณะสี (Color features) ถือเป็นคุณลักษณะของภาพที่มนุษย์มองเห็นได้เด่นชัดที่สุด แต่มนุษย์มีความสามารถในการแยกแยะระดับของสีได้ เช่น สีของผลแอปเปิ้ล แทนด้วยสีแดง แต่ในความเป็นจริงแล้วสีของแอปเปิ้ลที่มองด้วยมนุษย์ อาจจะเห็นเป็นสีแดงสด ก็เป็นได้ เช่นเดียวกันกับการสกัดข้อมูลภาพด้วยเครื่องคอมพิวเตอร์ อาจจะทำให้ได้ข้อมูลไม่ครบถ้วน ซึ่งในความเป็นจริงแล้วสีที่ปรากฏในโลกของความเป็นจริงนั้นจะประกอบด้วย ลวดลายบนพื้นผิว ความสว่าง และเฉดสี เป็นต้น จึงจะแสดงข้อมูลภาพที่มองเห็นด้วยมนุษย์ได้ครบถ้วน วิธีการวิเคราะห์ด้วยการสกัดข้อมูลสีบนภาพมีหลายวิธีดังนี้

1.1 แบบจำลองสี (John R. Smith, 1994) เป็นโครงสร้างที่กำหนดขึ้นเพื่อจัดระบบสีให้เป็นระเบียบให้เป็นไปตามมาตรฐานและควบคุมการเปลี่ยนแปลงสีให้ง่ายขึ้นซึ่งแบบจำลองสีที่ใช้อยู่ในปัจจุบันได้สร้างขึ้นมาเพื่อรองรับอุปกรณ์หรือโปรแกรมประยุกต์ต่างๆ โดยแบบจำลองสี (Color model) คือระบบการแทนค่าทฤษฎีเป็นมาตรฐาน เพื่อใช้อ้างอิงรูปแบบและค่าสีระหว่างเครื่องคอมพิวเตอร์ หรือระหว่างโปรแกรมหรืออุปกรณ์เพื่อให้เข้าใจตรงกัน แบบจำลองสีที่นิยมทั่วไปมีดังนี้

1.1.1 แบบจำลอง RGB เป็นแบบจำลองที่นิยมใช้ ในโปรแกรมคอมพิวเตอร์ทั่วไป รวมทั้งอุปกรณ์ต่างๆ ประกอบด้วยแม่สี 3 สีคือ แดง (red), เขียว (Green) และน้ำเงิน (Blue) เป็นแบบจำลองที่มาจากหลักการทำงานของจอภาพโทรทัศน์หรือจอคอมพิวเตอร์ ดังนั้นค่าของสีในแบบจำลองนี้จึงมีความหมายในเชิงเทคนิคที่ใกล้เคียงกับสายมนุษย์มองเห็น

1.1.2 แบบจำลอง CMY / CMYK (Cyan Magenta Yellow Key) เป็นแบบจำลองสีที่พัฒนาขึ้นมาใช้สำหรับการพิมพ์ภาพสีโดยมีสีหลัก คือ สีน้ำเงินเขียว (Cyan) สีม่วงแดง (Magenta) และสีเหลือง (Yellow) ระบบสีแบบ CMY สามารถหาได้โดยการนำสีในระบบ RGB ลบกับสีขาว

1.1.3 แบบจำลอง HSI (Hue Saturation Intensity Model) เป็นแบบจำลองที่อยู่บนพื้นฐานการมองเห็นสีด้วยตามนุษย์ และมีการแทนค่าที่คล้ายคลึงกับการแปลความหมายของสีในภาพที่มนุษย์มองด้วยตามากที่สุด ค่าสี (Hue) ค่าสีของสีหลัก (แดง, เขียว, น้ำเงิน) เป็นสีที่มีการสะท้อนมาจากสีของวัตถุค่าสีจะเปลี่ยนแปลงไปตามสเปกตรัมของสีจนถึง 255

1.2 ฮิสโทแกรมสี (Color histogram) (Jong-Seung Park, 2004) (M. Singha, 2012) เป็นวิธีการนับจำนวนพิกเซลของแต่ละสีบนภาพ ผลที่ได้เป็นจำนวนพิกเซลของเม็ดสีบนภาพทั้งหมด ฮิสโตแกรมสีเป็นกราฟแสดงความถี่หรือจำนวนจุดภาพที่ระดับค่าสีแต่ละค่าโดยแกนนอนแทนระดับค่าสีต่างๆ ส่วนแกนตั้งแทนจำนวนจุดภาพที่ระดับค่าสีนั้นๆ ในการคำนวณค่าฮิสโทแกรมสีภาพ แต่ละภาพจะถูกควอนไทซ์สีภายในภาพ เพื่อลดมิติของเวกเตอร์และลดความซับซ้อนในการคำนวณลงโดยการแบ่งกลุ่มสีออกเป็นถังสี (Bin) ซึ่ง ส่วนใหญ่นิยมใช้ 32, 64 หรือ 256 ถังสีเนื่องจากการแยกแยะความแตกต่างของระดับค่าสีด้วยสายตามนุษย์มีความละเอียดไม่มากนักสามารถ เขียนสมการสำหรับการคำนวณฮิสโตแกรมสีได้ดังนี้ $h_{C_i} = \frac{H_{C_i}}{n_1 \times n_2}$ โดยที่ $h_{C_i}(I)$ คือฮิสโทแกรมของสี C_i เมื่อให้ C_i แทนสีที่ i ของภาพ I, M คือจำนวนสีภายในภาพหลังการควอนไทซ์สี n_1 คือความกว้างของภาพและ n_2 คือความยาวของภาพ ดังนั้นภาพ I มีขนาด $n_1 \times n_2$ จุดภาพและ H_{C_i} แทนจำนวนจุดภาพที่มีสี C_i ของภาพ I สามารถเขียนเป็นเวกเตอร์แทนฮิสโตแกรมสีของ I ภาพ $H(I) = (h_{C_{i_1}}, h_{C_{i_2}}, \dots, h_{c_m})$ ดังสมการ ฮิสโทแกรมสีเป็นลักษณะเฉพาะทางสี ของภาพที่นิยมนำมาใช้ในระบบการค้นคืนภาพแต่อย่างไรก็ตามฮิสโทแกรมของภาพบางภาพมีค่าฮิสโทแกรมของสีเหมือนกันทั้งที่ภาพ

ทั้งสองมีลักษณะที่แตกต่างกัน จึงยังทำให้ การใช้ค่าฮิสโทแกรมเพียงอย่างเดียวสำหรับการค้นคืนข้อมูลนั้นไม่เพียงพอ

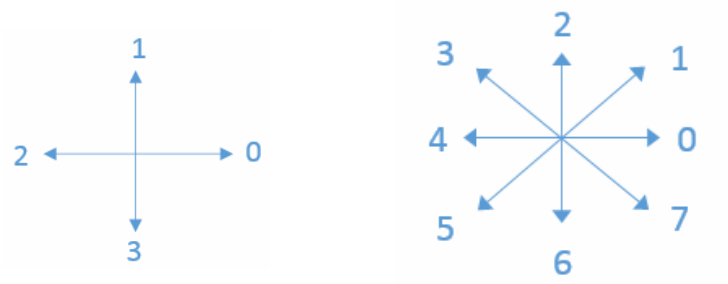
1.3 การเชื่อมติดกันของสี (Color coherence) (Reddy G., 2015) (S. Mangijao Singh, 2012) ฮิสโตแกรมสีที่ถูกนำมาใช้เนื่องจากสามารถคำนวณได้ง่ายและรวดเร็วแต่มีข้อเสียคือมีเฉพาะข้อมูลเกี่ยวกับการกระจายของสีภายในภาพเท่านั้นไม่มีข้อมูลเชิงตำแหน่ง (Spatial information) ทำให้ภาพบางภาพมีฮิสโทแกรมที่เหมือนกันแต่ลักษณะของภาพต่างกัน จึงได้มีการปรับปรุงโดยการเพิ่มเติมข้อมูลเชิงตำแหน่งเข้ามาหลักการ ของการเชื่อมติดกันของสีก็คือการแยกกลุ่มจุดภาพออกเป็น 2 กลุ่ม คือจุดภาพที่เชื่อมติดกัน (Coherent pixel) และจุดภาพที่ไม่เชื่อมติดกัน (Incoherent pixel) โดยจุดภาพใดๆ จะเป็นจุดภาพที่เชื่อมติดกันก็ต่อเมื่อองค์ประกอบที่เชื่อมติดกัน (Connected component) ของจุดภาพนั้นมีค่ามากกว่าค่าคงที่ค่าหนึ่งในลักษณะที่เป็นระดับซึ่งจุดภาพของสีนั้นๆ เป็นสมาชิกของบริเวณขนาดใหญ่ที่มีสีเหมือนกันเรียกบริเวณสำคัญนี้ว่าบริเวณที่เชื่อมติดกัน (Coherent regions) มีความสำคัญในการแสดงลักษณะเฉพาะของภาพ กำหนดให้ $H_{ccv}(I)$ คือ เวกเตอร์ขนาด m มิติแทน การกำหนดการเชื่อมติดของภาพ I ซึ่งคำนวณได้จา$H_{ccv}(I) = \langle(\alpha_1,\beta_1),(\alpha_2,\beta_2),\dots,(\alpha_m,\beta_m)\rangle$ ข้อดี คือเมื่อเปรียบเทียบกับฮิสโทแกรมสี คือมีประสิทธิภาพในการค้นคืนภาพ มากกว่าเนื่องจากมีการกำจัดจุดภาพที่ไม่ได้เป็นองค์ประกอบสำคัญของภาพทิ้งไปแต่จะใช้เวลาการประมวลผลมากกว่าฮิสโทแกรมสีจาก (Suhasini P,2009) (M. Singha, 2012) ได้ทำการเปรียบเทียบวิธีการสกัดคุณลักษณะสีทั้ง 3 วิธีการคือ Conventional Color Histogram, Invariant Color Histogram และ Fuzzy Color Histogram จะเห็นว่าความสัมพันธ์ของความถี่ของสีนั้นสามารถแสดงได้ด้วยวิธีการ Conventional Color Histogram แต่อย่างไรก็ตามการเก็บข้อมูลสีด้วยวิธีการนี้มีจำนวนมิติที่สูงทำให้การวัดความคล้ายกันของภาพค่อนข้างยุ่งยากกว่า

2. คุณลักษณะรูปทรง

คุณลักษณะรูปทรง (Shape feature) (Sanjay Singh, 2014) การอธิบายรูปทรงของวัตถุเป็นการนำสารสนเทศที่แสดงคุณลักษณะของวัตถุนั้นออกมาเพื่อลดปริมาณข้อมูลที่ไม่จำเป็นในระดับพิกเซลให้เหลือเพียงข้อมูลที่ใช้ในการแทนเพียงรูปทรงข้อมูลภาพหรือรูปร่างของวัตถุที่สนใจเท่านั้น วิธีการที่ใช้ในการอธิบายรูปทรงของภาพมีหลายวิธีการในที่นี้จะกล่าวพอสังเขปดังนี้

2.1 แบบขอบเขต (Boundary-based) เส้นขอบจะพิจารณาเฉพาะเส้นล้อมรอบวัตถุ เป็นการอธิบายคุณลักษณะของรูปร่าง จะแสดงคำอธิบายรูปร่างเพียงพื้นที่ทั้งหมดที่อยู่ในเส้นล้อมรอบวัตถุคุณลักษณะของรูปร่างที่ต้องการจะต้องไม่ขึ้นกับการเปลี่ยนตำแหน่งการหมุนและการเปลี่ยนขนาดวิธีหนึ่งที่มีคุณสมบัตินี้ คือรหัสลูกโซ่ (Chain code) เป็นรหัสแทนคุณลักษณะขอบวัตถุหรือรูปร่างวัตถุบนภาพด้วยตำแหน่งของการเรียงลำดับตำแหน่งและทิศทางของพิกเซลข้างเคียงตามรูปทรงของวัตถุบนภาพทั่วไปการกำหนดทิศทางมาตรฐานของรหัสลูกโซ่มีทั้งแบบ 4 ทิศทาง เรียกว่า รหัสลูกโซ่ 4 ทิศทาง (4 Direction Chain Code) และทิศทางรหัสลูกโซ่แบบ 8 ทิศทาง (8 Direction Chain Code) มีอาณาเขตติดต่อกันของตำแหน่งรวมทั้งหมด 8 พิกเซล

ตัวเลขรูปร่าง (Shape number) เป็นการแสดงรูปร่างด้วยรหัสลูกโซ่จะเป็นการแสดงในส่วนของรูปร่างด้วยลำดับของทิศทางของขอบวัตถุด้วยรหัสลูกโซ่แบบ 4 ทิศทาง เป็นวิธีการที่ใช้ข้อมูลน้อยและสามารถอธิบายได้อย่างดี แต่อย่างไรก็ตามจำเป็นต้องมีการกำหนดจุดเริ่มต้นของการอธิบายบนวัตถุให้ตรงกันเสมอ เพราะถ้ามีการกำหนดจุดเริ่มต้นที่ต่างกันบนวัตถุรูปแบบเดียวกันจะทำให้ผลของรหัสลูกโซ่มีความแตกต่างกันทันทีทั้งที่เป็นวัตถุเดียวกัน



(ก) รหัสลูกโซ่ 4 ทิศทาง (ข) รหัสลูกโซ่ 8 ทิศทาง

ภาพที่ 2 ทิศทางการอธิบายตำแหน่งของรหัสลูกโซ่

2.2 แบบบริเวณ (Region-based) เป็นการหาคุณลักษณะของรูปทรงจากพื้นที่ภายใน เช่น การใช้วิธีโปรเจคชัน (Projection) เวอร์ติเคิลโปรไฟล์ (Vertical profile) การนับจำนวนพิกเซลที่อยู่ในแนวคอลัมม์หรือฮอริสันเทิลโปรไฟล (Horizontal profile) เป็นการนับจำนวนพิกเซลในแนวแถว ซึ่งวิธีการนี้เคยนำมาใช้ในแอพพลิเคชันการรู้จำตัวอักษร เมื่อตัวอักษร "L" และ "T" ที่มีความแตกต่างกันของจำนวนพิกเซลในแนวแถวแต่แนวคอลัมม์มีความเหมือนกัน หรือตัวอักษร "A" และ "H" มีความเท่ากันของจำนวนพิกเซลในแนวแถวแตกต่างกันในแนวคอลัมม์ จะเห็นว่าวิธีการแบบโปรไฟล์จะสามารถใช้ได้ดีสำหรับรูปทรงที่เป็นสัญลักษณ์หรือบริเวณที่แยกออกจากกัน วิธีการนี้จะถูกเรียกว่า "Signature parsing" คือวิธีการวนซ้ำ เพื่อแยกบริเวณของเวอร์ติเคิลโปรไฟล์และฮอริสันเทิลโปรไฟล์จากกัน

2.3 การหาขอบภาพ (Edge detection) เป็นวิธีการที่แสดงโครงร่างของวัตถุบนภาพ เป็นวิธีการหนึ่งที่สำคัญที่สามารถนำไปประยุกต์ใช้กับการแบ่งส่วนภาพได้ เป็นการหาจุดเริ่มเปลี่ยน เพื่อแบ่งส่วนวัตถุบนภาพด้วยการหาขอบของวัตถุ โดยทั่วไปขอบของวัตถุจะมีการเปลี่ยนแปลงค่าระดับความเข้มของพิกเซลข้างเคียง หรือพิกเซลเพื่อนบ้าน (Neighborhood pixels) จากค่าระดับความเข้มต่ำเป็นค่าความเข้มสูงอย่างรวดเร็ว ทำให้เกิดความไม่ต่อเนื่องของระดับความเข้ม ความต่างของค่าระดับความเข้มนั้นกลายเป็นขอบของภาพได้อย่างชัดเจน โดยทั่วไปวิธีการพื้นฐานที่ใช้สำหรับการหารูปทรงคือการหาขอบภาพด้วยอนุพันธ์อันดับหนึ่งและอนุพันธ์อันดับสองในระยะใกล้และผลลัพธ์ที่ได้จากการทำอนุพันธ์อันดับหนึ่งและอนุพันธ์อันดับสอง จากค่าระดับเทา สำหรับการหาขอบภาพด้วยอนุพันธ์อันดับหนึ่งอีกวิธีการหนึ่งหรือการใช้ตัวกรอง (Filter) เช่น ตัวกรองโซเบล (Sobel filter) และ ตัวกรองพรีวิท (Prewitt filter) เป็นต้นเพื่อแสดงตัวกรองอนุพันธ์ (Derivative filter)

3. คุณลักษณะพื้นผิว

คุณลักษณะพื้นผิว (Texture feature) (Smeulders A, 2000) หรือลวดลาย ลักษณะเด่นของการเป็นพื้นผิวคือการมีรูปแบบที่ซ้ำกันในบริเวณหนึ่งๆ เป็นรูปแบบการเรียงตัวของสีที่ต่างกันออกไป สามารถกล่าวได้ว่าพื้นผิว (Texture) เป็นลักษณะเฉพาะที่ใช้อธิบายความหยาบ ความละเอียด หรือความซับซ้อนของวัตถุภายในภาพ สามารถสกัดคุณสมบัติของพื้นผิวด้วยวิธีการดังต่อไปนี้

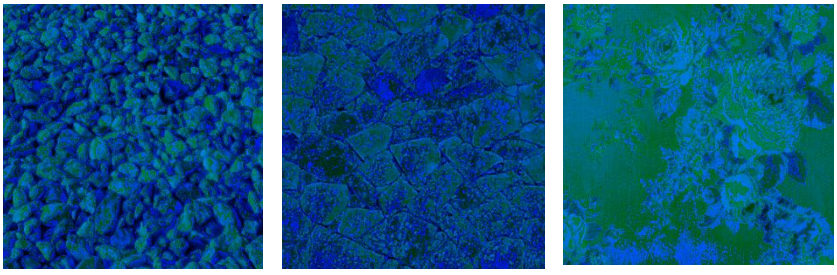
3.1 Discrete Wavelet Transform (DWT) (Smeulders A,2000) และ Gabor (Suhasini P, 2009) เป็นวิธีที่สามารถอธิบายรูปแบบของลวดลายได้ ด้วยการใช้ Band Pass Filters เป็นการอธิบายถึงความเหมือนและความซับซ้อนกันที่มีลักษณะเหมือนกับการมองด้วยตามนุษย์ ซึ่งสามารถแทนได้ด้วยค่าที่เป็นตัวเลข แต่อย่างไรก็ตามขึ้นอยู่กับบริเวรของการค้นคืนด้วยว่าเป็นแบบมีรูปแบบซ้ำ (Repetitive patterns) หรือไม่มีรูปแบบ (Non-structured) ดังแสดงในภาพที่ 3 การแปลงเวฟเล็ตแบบดีสครีท (Discrete Wavelet Transform: DWT) การแปลงเวฟเล็ตของข้อมูลภาพ เป็นการแปลงเวฟเล็ตแบบดีสครีทแบบ 2 มิติ (2-D Discrete Wavelet Transform: 2D-DWT) ใช้หลักการแยกองค์ประกอบ

เป็นแบนด์ย่อย (Subband Decomposition) สามารถแยกภาพออกเป็นส่วนย่อยตามความถี่ได้ ค่าสัมประสิทธิ์ของค่ากลาง (Mean) และความแปรปรวน (Variance of the coefficients) ในแต่ละส่วนย่อยสามารถคำนวณได้เป็นคุณลักษณะของพื้นผิวหรือลวดลายสามารถเขียนเป็นสมการได้ดังนี้

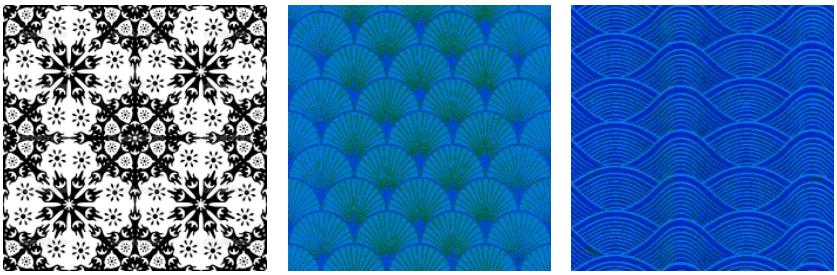
$$g(x,y) = \sum_{(k,l)\in K_v} k_{kl}\varphi_{kl}(x,y), (x,y) \in L$$

เมื่อกำหนดให้บริเวณ $f(x,y), (x,y) \in A$ เก็บข้อมูล Mพิกเซล, Aเป็นบริเวณและขอบ ใช้วิธี 2D-DWTและ L มีขนาดN(> M)มีบริเวณปิดเมื่อให้ $g(x,y)$และ k_{kl} เป็นสัมประสิทธิ์ที่มีความสัมพันธ์กับพื้นฐานฟังก์ชัน φ_{kl} ซึ่งภาพแต่ละภาพอาจจะประกอบไปด้วยวัตถุที่มีลักษณะพื้นผิวที่แตกต่างกันซึ่งภาพแต่ละภาพอาจจะประกอบไปด้วยวัตถุที่มีลักษณะพื้นผิวที่แตกต่างกัน การวิเคราะห์พื้นผิวจะช่วยให้สามารถแยกแยะความแตกต่างของวัตถุได้ดียิ่งขึ้น เช่น การแยกแยะความแตกต่างระหว่างภาพเสือดาวซึ่งมีลายเป็นจุดกับเสือชนิดอื่นซึ่งมีลายเป็นริ้วหรือแถบยาวออกจากกันได้ เป็นต้น

3.2 Tamura และคณะ (Smeulders A, 2000) (Suhasini P,2009) ได้ค้นพบความสัมพันธ์ระหว่างการรับรู้ทางการมองเห็นของมนุษย์และคำนวณหาความแตกต่างระหว่างพื้นผิว 6 คุณลักษณะ ได้แก่ ความหยาบ (Coarseness) ความตรงกันข้าม (Contrast) ทิศทาง (Directionality) ความเหมือนกันของเส้น (Line-likeness) ความสม่ำเสมอ (Regularity) และความขรุขระ (Roughness) คุณลักษณะพื้นผิวที่สำคัญที่สุดคือ ความหยาบเป็นการให้ข้อมูลส่วนประกอบของพื้นผิว (Texture elements) ดังนั้นในการคำนวณค่าความหยาบจะถูกใช้เพื่อแปรเปลี่ยน ตามขนาด สำหรับความตรงกันข้ามจะแปรเปลี่ยนได้ 4 ปัจจัย คือ ระดับเทา ค่าความกระจายของสีดำ และขาว บนฮิสโทแกรม ความคมชัดของขอบ และระดับความซ้ำกันของข้อมูลรูปแบบ



(ก) ข้อมูลลวดลายไม่มีรูปแบบ



(ข) ข้อมูลลวดลายแบบมีรูปแบบซ้ำ

ภาพที่ 3 คุณลักษณะพื้นผิว

4. การผสมคุณลักษณะภาพ

นักวิจัยพยายามปรับปรุงวิธีการสกัดข้อมูลพื้นฐานของภาพ เพื่อให้ได้ผลลัพธ์ของการค้นคืนที่ใกล้เคียงกับภาพที่ต้องการมากที่สุด ด้วยการนำวิธีการมาผสมผสานกันระหว่างคุณลักษณะภาพ (Combined features) เพื่อให้สามารถวิเคราะห์ในรูปแบบที่ซับซ้อนได้มากขึ้น เช่น การรวมเทคนิคด้วยคุณลักษณะสีและรูปทรงของภาพเพื่อทำการค้นคืนภาพ หรือมีการใช้อัลกอริทึมเพื่อทำการสกัดข้อมูลภาพเป็นคุณลักษณะเพื่อนำมาใช้ในการ ค้นคืนภาพในรูปแบบที่แตกต่างกันออกได้ (Reddy et al, 2005) (Wenjing Jia, 2006) ได้พยายามใช้คุณลักษณะของรูปทรง สีและขนาดในการค้นคืนภาพจากการทดลองแล้วผลลัพธ์ที่ได้ส่วนใหญ่ตรงตามความต้องการแต่อย่างไรก็ตามยังมีบางกลุ่มภาพที่ได้ผลไม่ถูกต้อง Adnan et al. (Wu J, 2010) ใช้วิธีการค้นคืนภาพด้วยรูปทรงเช่นเดียวกันรวมทั้งสีและลวดลายเพื่อระบุและแสดงถึงความสัมพันธ์ของวัตถุในภาพ (Wu et al,2010 (Reddy G,2005) ได้นำเสนอการรวมกันของคุณลักษณะสีและลวดลายเพื่อใช้ในการค้นคืนโดยลวดลายถูกสกัดด้วยวิธีการ Dual Tree Complex Wavelet Transform (DT-CWT) และ Rotated Wavelet Filter และคุณลักษณะของสีใช้ฮีสโทแกรม (Color histograms) ของรูปแบบ RGB และ HSV

การวัดความคล้าย (Similarity measure)

การวัดความคล้ายกันของภาพสองภาพด้วยการวิเคราะห์ทางตัวเลขจากคุณลักษณะภายในภาพที่ได้สกัดมา ซึ่งการวัดความคล้ายนี้ไม่ได้ขึ้นกับความคล้ายในดุลยพินิจของมนุษย์แต่จะเป็นการเปรียบจากตัวเลขของคุณลักษณะด้วยเครื่องคอมพิวเตอร์เท่านั้นดังนั้นจะพบว่าสี รูปทรง ลวดลายเป็นคุณลักษณะหนึ่งที่ใช้ในการวัดค่าความคล้ายกันของภาพได้ โดยค่าความคล้ายเป็นปริมาณที่สะท้อนถึงของความสัมพันธ์ระหว่างวัตถุ 2 วัตถุ หรือ 2 คุณลักษณะ เมื่อกำหนดให้วัดค่าความคล้ายของ 2 เวกเตอร์ที่ประกอบด้วย $x, y \in R^d$ (Zainab Ibrahim Abood, 2013) ในบทความนี้จะแสดงวิธีการวัดค่าความคล้าย 3 วิธี ดังนี้

1. Euclidean distance เป็นการวัดระยะห่างของ 2 เวกเตอร์ที่ประกอบด้วย $x, y \in R^d$ คำนวณได้จากสมการดังนี้

$$\delta_1(x,y) = \|x-y\|_2 = \sqrt{\sum_{j=1}^{d}(x_j - y_j)^2}$$

วิธีการวัดความคล้ายนี้สามารถเขียนได้อีกรูปแบบได้ดังนี้

$$\tau_1(x,y) = \|x-y\|_1 = \sum_{j=1}^{d}|x_j - y_j|$$

$$\tau_2(x,y) = \max_{1\le j\le d}|x_j - y_j|$$

2. Mahalanobis distance เป็นการวัดระยะความห่างของ 2 เวกเตอร์ที่ประกอบด้วย โดยจะใช้การเรียนรู้รูปแบบของ สามารถเขียนสมการได้ดังนี้

$$\delta_2(x,y) = \sqrt{(x-y)^t S^{-1}(x-y)}$$

เมื่อกำหนดให้ S เป็นเมตริกซ์ความแปรปรวนร่วม (Covariance Matrix) จากเวกเตอร์ $\{x_i | 1 \le i \le n\}$ เมื่อกำหนดขนาดเป็น n แล้วสามารถคำนวณได้ $S = \frac{1}{n}\sum_{i=1}^{n}(x_i - u)(x_i - u)^t$ เมื่อ $u = \frac{1}{n}\sum_{i=1}^{n} x_i$ ดังนั้นจะเห็นว่าการคำนวณโดยอาศัยค่าจากสถิติ $\delta_2(x, y)$ ก็คือช่วงห่างของค่านั้นกับการแจกแจง ซึ่งสามารถทดสอบนัยสำคัญได้ถ้ามีตัวแปรอิสระเพียงตัวเดียว Mahalanobis distance ก็คือค่ากำลังสองของค่ามาตรฐานของตัวแปร

3. Chord distance เป็นการวัดระยะห่าง ระหว่างสองเวกเตอร์คือ เพื่อวัดระยะห่าง ระหว่างเวกเตอร์ สามารถคำนวณได้จากสมการ

$$\delta_2(x, y) = \left\| \frac{x}{r} - \frac{y}{s} \right\|_2$$

เมื่อ $r = \|x\|_2, s = \|y\|_2$ โดยปกติการคำนวณจะนำไปสู่ $\delta_2(x, y) = 2\sin\left(\frac{\alpha}{2}\right)$ เมื่อค่ามุมของเวกเตอร์ x และ y คือ α มีค่าน้อยมาก ดังนั้นค่าความคล้ายที่มีพื้นฐานจากมุมเวกเตอร์ x และ y สามารถเขียนเป็นสมการได้ดังนี้ $\tau_3(x, y) = 1 - |\cos(\alpha)|$, เมื่อ $\cos(\alpha) = \frac{x.y}{\|x\|_2\|y\|_2}$

สำหรับการวัดค่าความคล้ายกันของข้อมูลภาพจะเป็นส่วนที่มีความสำคัญไม่น้อยไปกว่าการสกัดคุณลักษณะข้อมูลภาพ การวัดค่าความคล้ายกันสามารถวัดได้หลากหลายวิธีขึ้นกับคุณลักษณะของข้อมูลที่ได้มา แต่ในที่นี้จะกล่าวถึงข้อมูลที่ได้มามีลักษณะเป็นรูปแบบของเวกเตอร์ สามารถเลือกใช้วิธีการได้ ดังที่กล่าวมาข้างต้น ความแตกต่างของรูปแบบหรือผลลัพธ์ที่ได้มานั้น เมื่อนำมาเปรียบเทียบด้วยตามนุษย์แล้วในบางครั้ง แล้วอาจมีความต่างกัน แต่เนื่องจากคุณลักษณะต่างของข้อมูลภาพถูกแปลงออกเป็นตัวเลข ดังนั้นการวัดความคล้ายกันจึงต้องถูกเปรียบออกมาในลักษณะของตัวเลขเช่นกัน ทำให้ความเหมือนกันจากดุลยพินิจมนุษย์และรูปแบบการวัดความคล้ายจึงได้ผลอาจต่างกัน ในความเป็นจริงแล้วการวัดค่าความเหมือนของทั้งภาพนั้นวัดได้ยากอย่างไรก็ตามได้มีงานวิจัยที่พยายามจะปรับเปลี่ยนรูปแบบของการดึงข้อมูลภาพและวิธีการวัดความคล้ายให้มีความใกล้เคียงกับมนุษย์มากที่สุด จึงยังคงเป็นงานวิจัยที่ท้าทายสำหรับการค้นคืนข้อมูลภาพ

สรุป

ระบบค้นคืนภาพเป็นแอพพลิเคชันหนึ่งที่ถูกพัฒนาขึ้นเพื่อตอบสนองความต้องการของผู้ใช้งานสำหรับการค้นคืนข้อมูลภาพดิจิทัล เพื่อรองรับปริมาณภาพดิจิทัลที่เพิ่มมากขึ้นอย่างรวดเร็ว จึงเป็นสิ่งท้าทายสำหรับกลุ่มนักวิจัยทางด้านการประมวลผลภาพ ที่จะต้องสกัดข้อมูลจากภาพดิจิทัลอย่างไรและนำข้อมูลเหล่านั้นมาใช้ในการค้นคืนอย่างไรถึงจะได้ผลลัพธ์ที่ถูกต้องตรงตามความต้องการมากที่สุด สำหรับในบทความนี้เป็นการรวบรวมเนื้อหาของการสกัดข้อมูลภาพและทฤษฎีที่เกี่ยวข้องกับระบบค้นคืนภาพด้วยคุณลักษณะเฉพาะของภาพรวมทั้งการวัดความคล้ายของภาพเพื่อทำการแสดงผลลัพธ์ที่ได้ แต่อย่างไรก็ตามความท้าทายของการทำให้ได้มาของภาพผลลัพธ์ที่ต้องการนั้นยังคงเป็นสิ่งที่สำคัญที่สุด มีกลุ่มนักวิจัยพยายามที่จะทำการแทนค่าของเนื้อหา ภาพในรูปแบบต่างๆ รวมไปถึงการใช้คำศัพท์ (Keyword) เข้ามาช่วยเพิ่มความสัมพันธ์ของภาพ รวมไปถึงการป้อน ข้อมูลที่มีความเกี่ยวพันย้อนกลับ (Relevance feedback) เพื่อให้ได้ผลลัพธ์ดียิ่งขึ้น

บรรณานุกรม

Adnan A., Nawaz M., Anwar S., Ali T., & Ali M. (2010). " Object identification with color, Texture, and Object-Correlation in CBIR system. *World Academy of Science, Engineering and Technology*. 64(40) : 117-122.

B. S. Manjunath, & W. Y. Ma. (1996). "Texture features for browsing and retrieval of large image data". *IEEE Trans. On Pattern Anal.and Mach. Intell*. 18(8) : 837-842.

Chen, C., & Chu, H. (2015). "Similarity measurement between images". *Proceedings of 29thAnnual International Computer Software and Application Conference*. 2(1): 41-42.

DivyaSrivastava, Rajesh Wadhvani & Manasi Gyanchandani. (2015). "A review : Color feature extraction methods for content based image retrieval". *IJCEM Interna tional Journal of Computational Engineering & Management*. 18(3) : 2230-7893.

John R. Smith, & Shih-Fu Chang. (1994). "Transform features for texture classification and discrimination InLarge image databases". *Proc. IEEE Inter. Conf. on Image Processing*. (3) : 407-411.

Jong-Seung Park & TaeYong Kim. (2004). "Shape-based image retrieval using invariant features". *Advances in Multimedia Information Processing-PCM. Lecture Notes in Computer Science*. 33(32) : 146-153.

M. Singha & K. Hemachandran. (2012). "Content based image retrieval using color and texture". *Signal and Image processing. An international journal (SIPIJ)*. 3(1) : 299-309.

Reddy G., Babu G., & Somasekhar P. (2005). "Image retrieval by semantic indexing". *Journal of Theoreticaland Applied Information Technology*. 5(6) : 745-750.

S. Mangijao Singh & K. Hemachandran. (2012). "Content based image retrieval using color moment and gabor based image retrieval using color moment and gabor". *IJCSI International Journal of Computer Science*. 9(5) :13-22.

Sanjay Singh & Trimbak Ramchandra Sontakke. (2014). "An effective mechanism to neutralize the semantic gap in content based image retrieval (CBIR)". *The International Arab Journal of Information Technology*. 11(2) : 124-133.

Smeulders A., Worring M., Santini S., Gupta A., & Jain R. (2000). "Content-based image retrieval at the end of the early years". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 22(12) : 1349-1380.

Suhasini P., Krishna S., & Krishna M. (2009). "CBIR using color histogram processing". *Journal of Theoretical and Applied Information Technology*. 6(1) : 116-122.

WenjingJia, Huaifeng Zhang, Xiangjian He, & Qiang Wu. (2006). A comparison on histogram based image matching methods. *Proceedings of the IEEE International Conference on Video and Signal Based Surveillance*.

Wu J., Wei Z., & Youli C. (2010). "Color and texture feature for content based image retrieval. *The International Journal of Digital Content Technology and its Applications*". 4(3) : 43-49.

Zainab Ibrahim Abood, sraaJameel & Nabeel Jameel Tawfiq Muhsin. (2013). "Content-based image retrieval (CBIR) using hybrid technique". *International Journal of Computer Applications*. 83(12) : 17-24.