บทที่ 1
บทนำ

1.1 ความเป็นมาของปัญหา

สมัยก่อนการประมวลผลภาพ จะอยู่ในรูปของการประมวลผลสัญญาณแอนะล็อก (analog) โดยใช้อุปกรณ์จับภาพแต่งภาพที่ถ่ายออกมายังคงเป็นภาพขาวดำและไม่ชัดเจนเท่าที่ควร ทำให้ส่วนใหญ่ภาพถ่ายมักจะถูกถ่ายเฉพาะภาพเหตุการณ์ที่สำคัญหรือบุคคลสำคัญเท่านั้น เนื่องจาก อุปกรณ์ที่นำมาใช้ในการถ่ายภาพ หรือบันทึกภาพนั้นซับซ้อนและมีความยุ่งยากในการใช้งาน รวมทั้งราคาสูง แต่อย่างไรก็ตามได้มีการพัฒนาปรับปรุงเทคโนโลยีเพื่อตอบรับกับความต้องการที่เพิ่มขึ้น ทำให้อุปกรณ์ถ่ายภาพในปัจจุบันมีราคาถูกลง และสะดวกต่อการใช้งานมาก และพัฒนาต่อเนื่องไปเป็นอุปกรณ์ถ่ายภาพแบบดิจิทัล (digital) ทำให้ภาพถ่ายภาพดิจิทัล มีจำนวนเพิ่มมากขึ้นอย่างต่อเนื่อง ด้วยเหตุนี้ทำให้เกิดปัญหาในการจัดเก็บข้อมูลภาพรวมถึงการค้นคืนข้อมูลภาพ (image retrieval) ที่เพิ่มมากขึ้นอย่างต่อเนื่อง จนทำให้เกิดคำถามเดิม ๆว่าควรจะใช้วิธีการใดจึงจะสามารถจัดเก็บอย่างมีระบบและสามารถที่จะค้นหาภาพที่ต้องการได้อย่างถูกต้อง

การประมวลผลภาพดิจิทัล (digital image processing) เป็นสาขาที่ได้รับความนิยมในการวิจัยค่อนข้างมาก ทั้งทางด้านภาพ (still image) และ วีดีโอ (video) เพราะสามารถที่จะนำไปประยุกต์ได้หลายด้าน เช่น ทางด้านการแพทย์ ทางด้านอุตสาหกรรม หรือนำมาประยุกต์ใช้ในส่วนของทางด้านการจราจร ในการตรวจจับทะเบียนรถยนต์ เป็นต้น หัวข้อหลักในงานวิจัยแขนงนี้จะเกี่ยวเนื่องกันทั้งทางด้านการสร้างกระบวนการ (algorithm) เพื่อนำมาใช้ในการจำแนกหรือสืบค้นข้อมูลภาพ, การแทนความหมายของข้อมูลภาพ (image representation) และวิธีการใช้เครื่องมือในการจำแนกภาพ (classification method) [Pedro,2007] ซึ่งปัญหาทั้งหมดที่เกิดขึ้นเพียงเพื่อนำมาใช้ในการจำแนกหรือสืบค้นข้อมูลภาพ เพื่อให้ได้ภาพคำตอบที่ตรงตามความต้องการมากที่สุด สิ่งที่กล่าวมาข้างต้นเป็นเพียงส่วนหนึ่งที่นำการประมวลผลภาพดิจิทัลเข้ามาปรับปรุงและไปประยุกต์ใช้งาน ปัจจุบันมีงานวิจัยหลายกลุ่ม [Qian Huang,1995][Vailaya A.,2001][W. Ma,1997][C. Carson, 1999] พยายามนำเทคนิคต่าง ๆเข้ามาใช้ในการค้นหาภาพเพื่อให้ผลลัพธ์ตรงกับความต้องการของผู้ใช้ ไม่ว่าจะเป็นการสืบค้นข้อมูลโดยใช้การประมวลผลภาพระดับต่ำ (low level image processing) เป็นการนำคุณลักษณะของภาพต่าง ๆหรือฟีเจอร์ (feature) ที่ถูกสกัด (extraction) อย่างเช่น สี (color)

ลวดลาย (texture) หรือ รูปทรง (shape) ด้วยวิธีการ อัลกอริทึมต่างๆ เพื่อใช้สำหรับการสืบค้นข้อมูล

ความพยายามในการค้นหาข้อมูลภาพเริ่มมาจากการสกัดคุณลักษณะข้อมูลภายในภาพ (feature extraction) ออกมาเป็นข้อมูลที่สามารถคำนวณได้ ข้อมูลส่วนนี้มักจะถูกเรียกกันว่าข้อมูลภาพระดับต่ำหรือ low-level features โดยข้อมูลนี้จะถูกนำมาใช้สำหรับการสืบค้นข้อมูลภาพแต่อย่างไรก็ตาม การวิจัยในส่วนแรกนี้ยังคงมีการปรับปรุงพัฒนาอย่างต่อเนื่อง แต่อย่างไรก็ตามได้มีงานวิจัยอีกกลุ่ม ที่พยายามจะใช้เทคนิคของการเข้าใจความหมายของภาพแทน การสืบค้นแบบข้างต้นเรียกว่า การประมวลผลภาพระดับสูง (high level image processing) งานวิจัยในกลุ่มนี้พยายามที่จะพิจารณาข้อมูลบนภาพเป็นวัตถุ (object) ที่มีความหมาย [Benitez A.B,2002] และแทนวัตถุนั้นๆ ด้วยคำหลัก (keyword) ลงบนภาพ เช่น “building”, “car” , “ball” เป็นต้น และใช้ความหมายหรือคำหลักนั้นเพื่อทำการสืบค้นข้อมูลแทน เมื่อเปรียบเทียบผลลัพธ์ของการค้นคืนที่ค่อนข้างแม่นยำมากกว่าการใช้ฟีเจอร์ระดับต่ำ แต่อย่างไรก็ตามยังขึ้นอยู่กับว่าอัลกอริทึมที่ถูกนำมาใช้ร่วมด้วยนั้นจะเป็นลักษณะใด เช่น การใช้เทคนิคหาความสัมพันธ์ของแท็กชื่อวัตถุบนภาพ [Benitez A.B,2001][R. Zhao,2002][Philippe Mulhem,2002] ซึ่งเป็นการสอดคล้องกันด้วยความหมายตามพจนานุกรม การใช้ความสัมพันธ์ของความหมายที่เหมือนกันของคำ (synonym) การค้นหาภาพด้วยเทคนิคนี้จะได้ผลลัพธ์ ที่ค่อนข้างขึ้นกับคำหลักที่ถูกให้ความหมายไว้ในภาพหรือ วัตถุบนภาพส่วนใหญ่ เพราะฉะนั้น ผลลัพธ์ของการค้นคืนภาพ คือการค้นหาคำหลักที่อยู่ในภาพนั้นเอง จึงทำให้ความหมายภาพที่แท้จริงจะไม่สามารถเกิดขึ้นได้เลย

นักวิจัยบางกลุ่มได้มีการสร้างความสัมพันธ์ของวัตถุบนภาพด้วยโมเดลเหตุการณ์ (event model) [Joo-Hwee Lim 2003] เช่นการเชื่อมวัตถุด้วยคำต่างๆ เช่น “touch”, “ontop” เป็นต้น แต่ผลของการทดลองยังไม่ดีนัก บางงานวิจัยได้พยายามใส่ข้อความที่บรรยายความหมายของภาพ (context) [Mathias Lux 2003] [Mathias Lux 2009] ในหัวข้อที่สอดคล้องกับภาพ เช่น “birthday party of uncle Adam” หรือ “a picture showing a barking dog” แต่อย่างไรก็ตาม การสืบค้นภาพที่ถูกบรรยายนั้นส่วนใหญ่จะเป็นจะภาพส่วนตัว (personal images) บางงานวิจัยพยายามที่จะใส่ข้อมูลเหตุการณ์ต่างๆจนครบถ้วนในรูปแบบของ ใคร ทำอะไร ที่ไหน เมื่อไหร่ (who, what, when, where) ในการใส่ข้อมูลบางครั้ง เป็นข้อมูลที่นอกเหนือ หรือเกินความจำเป็นโดยใช่เหตุ ข้อมูลเหล่านี้อาจจะไม่มีความจำเป็นเลยสำหรับการสืบค้นข้อมูลหรือ แทนความหมายของภาพจริงๆ หรือบางครั้งอาจจะเป็นข้อมูลที่มีความเป็นส่วนตัวจนเกินไป และใช้คำศัพท์ซ้ำซ้อน ฟุ่มเฟือย

และมีนักวิจัยบางกลุ่มพยายามที่แก้ไขถึงความยุ่งยากลำบากในการใส่ข้อมูลบนภาพ (image annotation) และจำกัดขอบเขตคำศัพท์ของภาพให้รัดกุมยิ่งขึ้น จึงได้พัฒนาซอฟต์แวร์เพื่อให้เกิดความสะดวกสบายในการใส่ข้อมูล และการใส่ข้อมูลที่มีแบบแผนแน่นอน ยกตัวอย่างงานวิจัยเช่น Caliph & Emir [Mathias Lux 2009], Annosearch [Xin-Jing 2008], CAMEL [Apostol Paul Natsev 2001] เป็นต้น แต่อย่างไรซอพต์แวร์ส่วนใหญ่นั้น จะทำการสืบค้นภาพในรูปแบบของการเทียบคำศัพท์เป็นคำต่อคำตามที่มีการเก็บข้อมูลไว้เท่านั้น เพราะฉะนั้นคำตอบที่ได้จะไม่ได้ขึ้นกับความหมายของภาพ แต่ขึ้นกับคำหลักที่มีการเก็บข้อมูลลงไปเท่านั้น ความหมายของภาพยังไม่ได้อย่างแท้จริง ในความเป็นจริงแล้วความหมายของภาพจะเกิดจากความสัมพันธ์ที่คล้ายกันของวัตถุในหมวดหมู่เดียวกัน เพราะฉะนั้นผลลัพธ์ของกลุ่มภาพที่ได้จะไม่ได้ขึ้นกับความหมายของภาพอย่างแท้จริง แต่ขึ้นกับคำหลักที่มีการเก็บข้อมูลลงไปบนภาพเท่านั้น จึงทำให้ความหมายของภาพโดยรวมนั้นยังไม่ได้สื่อออกมาให้เห็นอย่างเด่นชัด ซึ่งในความเป็นจริงแล้วความหมายของภาพจะเกิดจากการแปลความหมายรวมของทุกวัตถุที่ปรากฏบนภาพ ดังนั้นในงานวิจัยนี้ได้นำเสนอในรูปแบบของการแทนข้อมูลภาพ ด้วยความสัมพันธ์ของข้อมูลวัตถุภายในภาพ หรือเรียกว่า แนวคิดกราฟ (Conceptual Graph) ในลักษณะของกราฟที่นำเสนอนั้นจะมีแสดงความสัมพันธ์ของข้อมูลภายในภาพของวัตถุที่เกิดขึ้นทั้งหมดในรูปแบบของกราฟรวมทั้งความสัมพันธ์ (Relationship) ระหว่างวัตถุที่เกิดขึ้น และจะมีการวัดค่าความสัมพันธ์ของความเหมือนกันด้วยวิธีการที่เรียกว่า การจับคู่กราฟ (Graph Matching) ของข้อมูลภาพที่ถูกแทนค่าความหมายไว้ ดังนั้นภาพที่มีความหมาย เหมือนกันจะมีได้ค่าของความเหมือนกันมากที่สุด

จากที่กล่าวมาข้างต้นในงานวิจัยนี้จึงได้นำเสนอการปรับปรุงการสืบค้นข้อมูลภาพด้วยความหมายภาพเพื่อให้ได้ผลลัพธ์ของภาพที่มีอย่างแท้จริง โดยใช้วิธีการแทนข้อมูลภาพแบบคอนเซปชวลกราฟ (Conceptual Graph) และทำเปรียบเทียบความเหมือนกันของความหมายภาพด้วยการหาความเหมือนของภาพทั้งหมด 4 วิธีการ โครงข่ายประสาทเทียมแบบหลายชั้น (Multiple Feedforward Neural Network) แผนผังการจัดระบบตัวเอง (Self-organizing maps) เครือข่ายแบบเบย์ (Bayesian Network) และการวัดความคล้ายกันของกราฟแบบจับคู่ (Similarity Measure with Conceptual Graph Matching)

1.2 วัตถุประสงค์

1. เพื่อค้นหาเทคนิคใหม่ที่สามารถนำมาใช้วัดความคล้ายกันของกราฟแบบจับคู่
2. เพื่อสร้างโมเดลรูปแบบใหม่ในการแทนข้อมูลภาพ
3. เพื่อสร้างแบบจำลองวิเคราะห์การค้นหาภาพรูปแบบใหม่
4. เพื่อพัฒนาเทคนิคในการหาความเหมือนกันของภาพ
5. สามารถนำไปประยุกต์กับการสืบค้นภาพได้

1.3 สมมติฐาน

งานวิจัยนี้จึงเสนอวิธีการแปลความหมายภาพด้วย วิธีวัดความคล้ายกันของกราฟแบบจับคู่ โดยใช้หลักทฤษฎี การเชื่อมโยงประสานกับวัตถุบนภาพให้สอดคล้องตามความหมายของกระบวนการคิดแปลความหมายภาพของมนุษย์ และได้ทำการเปรียบเทียบความเหมือนกันของความหมายภาพด้วยการหาความเหมือนของภาพทั้งหมด 4 วิธีการ โครงข่ายประสาทเทียมแบบหลายชั้น (Multiple Feedforward Neural Network) แผนผังการจัดระบบตัวเอง (Self-organizing maps) เครือข่ายแบบเบย์ (Bayesian Network) และการวัดวิธีการวัดความคล้ายกันของกราฟแบบจับคู่ (Similarity Measure with Conceptual Graph Matching)

1.4 ขอบเขตของการวิจัย

1. ข้อมูลภาพที่นำมาวิเคราะห์เป็นข้อมูลภาพดิจิทัล เฉพาะภาพธรรมชาติ (natural images) เช่น ภาพทิวทัศน์ ชายทะเล ภูเขา หรือ สวนสาธารณะ ภาพภายในอาคาร ภายในบ้าน สำนักงาน หรือ โรงงาน เป็นต้น
2. ภาพที่มีวัตถุเกิดปรากฏขึ้น สำหรับภาพที่ไม่มีความหมายหรือเป็นภาพที่มีความหมายกำกวม หรือภาพที่สามารถแปลได้หลายความหมาย หรือภาพที่โฟกัสระยะใกล้ จะไม่นำภาพนั้นเข้ามาวิเคราะห์สำหรับการหาความหมายภาพข้อมูลที่นำมาวิเคราะห์เป็นข้อมูลภาพดิจิทัลที่เห็นวัตถุชัดเจน
3. ข้อมูลหรือวัตถุบนภาพที่เป็นอินพุตบนภาพจะถูกแท็ก เป็นคำศัพท์และขนาดมาก่อนด้วยขบวนการ รู้จำของการประมวลผลภาพในระดับต่ำ (low-level image)

4. ในการสร้างแบบจำลองภาพ (image representation) ข้อมูลจะถูกจัดเก็บมาในรูปแบบของฟีเจอร์ข้อมูลไว้แล้ว
5. ภาพจากคลังภาพจะถูกตั้งค่าเริ่มต้นโดยการใส่ข้อมูลวัตถุบนภาพพร้อมความสัมพันธ์ภาพไว้ก่อนแล้ว
6. ภาพที่ถูกนำมาใช้เป็นข้อมูลทดลองจากคลังภาพ [11]-[13]
7. การเก็บผลข้อมูลเบื้องต้นของความหมายของภาพ ใช้กลุ่มนักศึกษา และบุคคลทั่วไปเป็นกลุ่มบุคคลที่ตัดสิน ความหมายของภาพสำหรับการทดลองในคลังภาพที่กล่าวมา

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 ผลต่อสังคม

1. เพื่อนำเอากระบวนการนี้มาประยุกต์ใช้กับการค้นหาภาพบนห้องสมุดดิจิทัลได้
2. เพื่อได้แนวทางการแปลและการตีความหมายภาพ
3. ช่วยให้การใช้คำศัพท์ในการค้นหาข้อมูลภาพได้ผลตามความหมายของภาพมากขึ้น

1.5.2 ผลต่อมหาวิทยาลัย

1. สร้างกลุ่มนักวิจัยที่เป็นลักษณะของสาขาทางด้าน Images Processing ในสาขาย่อย Semantic Image Processing
2. เพื่อสนับสนุนและเสริมสร้างความรู้ความสามารถของการพัฒนาตนเอง และ กลุ่มงานในสาขาเทคโนโลยีสารสนเทศ
3. เพื่อเพิ่มตัวชี้วัดให้กับองค์กร สถาบันในส่วนของงานวิจัยและพัฒนา
4. สร้างชื่อเสียงให้กับมหาวิทยาลัย เมื่อมีบทความวิจัยลงในวารสารต่างประเทศ บทความวิจัยที่นำเสนอในการประชุมระดับชาติ และนานาชาติ

1.5.3 ผลต่อกลุ่มผู้วิจัย

1. พบแนวทางในการคิดค้นสิ่งใหม่ ๆในการจัดกลุ่มและกระบวนการแปลความหมายของภาพ
2. สามารถนำผลวิจัยมาเขียนบทความลงวารสารนานาชาติ และร่วมประชุมวิชาการระดับชาติและนานาชาติ
3. พัฒนาความรู้ใหม่ ๆ หลักการแนวคิดใหม่ ให้เกิดขึ้นในกลุ่มของนักวิจัย

1.6 นิยามคำศัพท์

1. การประมวลผลภาพ (Image processing) หมายถึง การนำภาพมาผ่านกระบวนการเพื่อประมวลผลสัญญาณบนสัญญาณ 2 มิติ เช่น ภาพนิ่ง (ภาพถ่าย) หรือภาพวีดิทัศน์(วีดิโอ) และนำมาใช้งาน
2. การจำแนกประเภทข้อมูลภาพ (Image Classification) หมายถึง การแยกแยะภาพที่มีคุณลักษณะเดียวกันออกเป็นกลุ่ม ๆ
3. การจำแนกความหมายของภาพ (Semantic Image Classification) หมายถึง การนำภาพมาแบ่งเป็นกลุ่มโดยพิจารณาจากความหมายโดยรวมของภาพ
4. การค้นคืนข้อมูล (Image Retrieval) หมายถึง การค้นหาข้อมูลของภาพจากระบบฐานข้อมูล
5. วัตถุ (Object) หมายถึง ส่วนของวัตถุบนภาพ ยกตัวอย่างเช่น คน, ต้นไม้ เป็นต้น
6. คำหลัก (Keyword) หมายถึง คำที่มีความหมายได้ใจความใช้แทนวัตถุบนภาพ
7. แท็ก (Tag) หมายถึง การกำหนดคำ หรือ คำศัพท์บนภาพ หรือเรียกว่าการ labeled
8. การแทนที่ข้อมูลภาพด้วยกราฟ (Graph representation) คือการแทนความหมายของภาพโดยใช้กราฟ
9. การเปรียบเทียบความเหมือนของกราฟ (Graph Similarity) คือการวัดค่าความคล้ายกันของกราฟ
10. คุณลักษณะข้อมูล หรือ ฟีเจอร์ (Feature) หรือตัวแปร ที่ถูกสกัดออกมาจากภาพ เช่น สี (color) ลวดลาย (texture) หรือ รูปทรง (shape) รวมทั้ง วัตถุ ที่ปรากฏบนภาพเพื่อนำมาใช้ในการสืบค้นข้อมูลต่อไป
11. โครงข่ายประสาทเทียม (Artificial neural network) หมายถึง เทคนิคการพยากรณ์เหมาะสำหรับลักษณะข้อมูลซึ่งไม่ใช่เชิงเส้น