# บทที่ 1

# บทนำ

## ความเป็นมาของปัญหา

ในปัจจุบันข้อมูลมัลติมีเดียได้กลายเป็นสิ่งจำเป็นและเป็นที่ต้องการเพิ่มมากขึ้น โดยเฉพาะกับหน่วยงานทั้งภาครัฐและเอกชนที่มีความจำเป็นต้องใช้ข้อมูลมัลติมีเดียเหล่านี้ทั้งที่ใช้ภายในองค์กรและที่ใช้แลกเปลี่ยนกันผ่านเครือข่ายอินเตอร์เน็ต (Internet) รวมถึงบุคคลทั่วไปที่ต้องการค้นหาข้อมูลเหล่านี้บนเครือข่ายอินเตอร์เน็ตโดยเฉพาะการค้นหาข้อมูลมัลติมีเดียประเภทภาพที่กำลังได้รับความนิยมเพิ่มมากขึ้นในปัจจุบัน

จากความต้องการที่เพิ่มมากขึ้นดังกล่าวจึงได้มีการพัฒนาเครื่องมือสำหรับค้นหาข้อมูลภาพบนเครือข่ายอินเตอร์เน็ตจากผู้วิจัยต่าง ๆ มากมายโดยแต่ละระบบก็จะมีเทคนิควิธีการที่แตกต่างกันออกไปซึ่งระบบหนึ่งที่มีประสิทธิภาพและกำลังได้รับความนิยมอยู่ในปัจจุบันคือ การค้นคืนภาพโดยยึดหลัก Content ของภาพ [Content-Based Image Retrieval (CBIR)] ซึ่งเป็นวิธีที่สามารถเข้าถึง (Access) ข้อมูลเชิงเสมือน (Visual Data) ได้โดยวิธีนี้จะค้นคืนภาพ (Retrieval) โดยใช้เนื้อหาเชิงเสมือนของภาพเหล่านั้นเช่น สี (Color) , รูปร่าง (Shape) , พื้นผิว (Texture) และโครงสร้าง (Structure) ของภาพ เป็นต้น แต่เนื่องจากระบบดังกล่าวนี้ยังไม่สามารถให้ผลลัพธ์ของการค้นคืนที่ตรงกับความต้องการของผู้ใช้ได้ดีพอจึงมีการคิดค้นวิธีการใหม่ที่ปรับปรุงเพิ่มเติมจากวิธีเดิมคือ วิธีที่ยึดหลักของการเรียนรู้ (Learning-Based Approach) โดยวิธีการนี้จะให้ผู้ใช้เป็นผู้ที่คอยสอนระบบโดยตรงว่าข้อมูลใดของภาพที่เขาต้องการและความเหมือนของภาพนั้นคืออะไร เป็นต้น โดยระหว่างที่มีการโต้ตอบระหว่างผู้ใช้กับระบบนั้นระบบจะส่งรายการภาพที่ได้จากการค้นหามาให้ผู้ใช้เลือกว่าภาพใดที่ผู้ใช้ต้องการและเหมือนกับภาพที่ใช้ค้นหา (Query Image) จากนั้นระบบก็จะทำการเรียนรู้ (Learning) จากข้อมูลการเลือกของผู้ใช้แล้วนำไปปรับแก้ Query เพื่อใช้ในการค้นหาในรอบต่อไปซึ่งวิธีการดังกล่าวนี้จะช่วยเพิ่มประสิทธิภาพในการค้นหาที่ตรงกับความต้องการของผู้ใช้มากขึ้นซึ่งเทคนิคดังกล่าวนี้เรียกว่า Relevance Feedback (RF)

เนื่องจากวิธีการที่ยึดหลักของการเรียนรู้ (Learning-Based Approach) นั้นเป็นลักษณะที่ให้ผู้ใช้เป็นผู้ที่คอยสอนระบบจึงยังมีปัญหาในเรื่องของการโต้ตอบกันระหว่างผู้ใช้กับระบบซึ่งผู้ใช้นั้นต้องคอยโต้ตอบกับระบบอยู่ตลอดเวลาทำให้ผู้ใช้ต้องทำงานมากขึ้นและเนื่องจาก

วิธีดังกล่าวนี้มีพื้นฐานการทำงานแบบ CBIR ซึ่งระบบนี้ส่วนใหญ่ถูกออกแบบมาเป็นเครื่องมือในการค้นหาภาพบนเครือข่ายอินเตอร์เน็ตซึ่งเทคโนโลยีเครือข่ายอินเตอร์เน็ตในปัจจุบันจะถูกออกแบบมาในลักษณะ Client-Server ซึ่งจะต้องมีการส่งข้อมูล และการแสดงภาพผ่านเครือข่ายจึงทำให้วิธีการ Learning-Based Approach ดังกล่าวนี้ประสบกับปัญหาในเรื่องของการแสดงผลข้อมูลและการติดต่อกันระหว่างผู้ใช้กับระบบซึ่งผู้ใช้กับระบบจะต้องมีการโต้ตอบกันตลอดเวลาในช่วงที่มีการค้นหาข้อมูลภาพซึ่งส่งผลกระทบในเรื่องความรวดเร็วในการทำงานผ่านเครือข่ายรวมถึง Traffic ที่เกิดขึ้นด้วย

ในเรื่องของการค้นคืนข้อมูลภาพนั้นยังมีปัญหาหลักที่สำคัญอันหนึ่งซึ่งในงานวิจัยที่เกี่ยวข้องหลาย ๆ งานวิจัยจะต้องพบและมักจะกล่าวถึงปัญหานี้ซึ่งปัญหาดังกล่าวคือ ช่องว่างระหว่างมุมมองหรือที่เรียกว่า Semantic Gab ซึ่งเป็นปัญหาที่เกิดขึ้นระหว่างมุมมองที่ต่างกันของมนุษย์ที่มีมุมมองในระดับสูง (High-Level) กับคอมพิวเตอร์ที่มีมุมมองในระดับต่ำ (Low-Level) และปัญหาด้านความรู้สึกนึกคิดของผู้ใช้ (Subjectivity of User) โดยปัญหาทั้ง 2 ส่วนนี้ยังคงเป็นประเด็นที่สำคัญในสาขาที่เกี่ยวข้องกับการประมวลผลภาพซึ่งส่งผลถึงประสิทธิภาพในการค้นคืนภาพที่ตรงกับความต้องการของมนุษย์และนักวิจัยต่าง ๆ ก็พยายามที่จะหาวิธีที่จะแก้ปัญหาดังกล่าวนี้เรื่อยมาจนถึงปัจจุบัน

เนื่องจากในปัจจุบันข้อมูลมัลติมีเดียได้กลายเป็นสิ่งจำเป็นในชีวิตประจำวันของมนุษย์ดังที่ได้กล่าวไปแล้วและเนื่องจากข้อมูลมัลติมีเดียนั้นมีรูปแบบและชนิดที่แตกต่างกันออกไปขึ้นอยู่กับวิธีการสร้างหรือการจัดเก็บข้อมูลเหล่านั้นของผู้ที่เป็นเจ้าของข้อมูลในแต่ละรายจึงทำให้เกิดปัญหาเกี่ยวกับการแลกเปลี่ยนข้อมูลระหว่างกันโดยเฉพาะการแลกเปลี่ยนข้อมูลบนเครือข่ายอินเตอร์เน็ตเพื่อแก้ปัญหาดังกล่าวจึงได้มีการกำหนดมาตรฐานเกี่ยวกับข้อกำหนดของข้อมูลมัลติมีเดียเหล่านี้ขึ้นโดยองค์กรที่ชื่อ MPEG ซึ่งหนึ่งในมาตรฐานดังกล่าวนี้คือ มาตรฐาน MPEG-7 ซึ่งจะกล่าวถึงข้อกำหนดเกี่ยวกับรูปแบบของการอธิบายและการจัดการ Content ของข้อมูลมัลติมีเดียโดยรายละเอียดเพิ่มเติมของมาตรฐานของข้อมูลมัลติมีเดียนั้นสามารถดูเพิ่มเติมได้ในบทที่ 2

MPEG-7 เป็นมาตรฐานที่ถูกกำหนดขึ้นโดยมีจุดประสงค์เพื่อกำหนดรูปแบบและวิธีการในการอธิบาย Content ของข้อมูลมัลติมีเดียต่าง ๆ ให้มีรูปแบบและมาตรฐานเดียวกันเพื่อให้ผู้ใช้สามารถใช้ข้อมูลดังกล่าวร่วมกันได้รวมถึงสามารถแลกเปลี่ยนข้อมูลดังกล่าวผ่านเครือข่ายชนิดต่าง ๆ ได้สะดวกยิ่งขึ้นและช่วยให้ผู้ใช้สามารถใช้ข้อมูลดังกล่าวได้อย่างมีประสิทธิภาพและเกิดประโยชน์มากที่สุดซึ่งจากมาตรฐานดังกล่าวนี้จะช่วยให้ผู้ใช้สามารถที่จะทำการ Browsing ,

Editing , Searching และอื่น ๆ เป็นต้น และช่วยให้องค์กรต่าง ๆ รวมถึงผู้ใช้ทั่วไปที่เป็นเจ้าของข้อมูลสามารถที่จะจัดการกับข้อมูลของตนได้อย่างมีประสิทธิภาพและเกิดประโยชน์มากยิ่งขึ้น ยกตัวอย่างเช่น การทำดัชนี (Indexing) การจัดกลุ่มข้อมูล (Data Clustering) การจัดประเภท (Classification) การจัดเก็บ (Storage) การบีบอัด (Compression) เป็นต้น จากประโยชน์ของมาตรฐาน MPEG-7 ดังกล่าวทำให้ MPEG-7 ได้กลายเป็นมาตรฐานที่สำคัญและเป็นประโยชน์ต่อสาขาอาชีพต่าง ๆ อาทิเช่น การศึกษา การแพทย์ การให้บริการสืบค้นข้อมูลต่าง ๆ เป็นต้น ส่งผลให้ MPEG-7 กลายเป็นมาตรฐานหนึ่งที่มีบทบาทที่สำคัญต่อระบบมัลติมีเดียในปัจจุบันและจะกลายเป็นมาตรฐานที่เป็นที่ยอมรับไปทั่วโลกในอนาคต

จากมาตรฐาน MPEG-7 ดังกล่าวจะเห็นว่า MPEG-7 นั้นจะกำหนดมาตรฐานเฉพาะส่วนของรูปแบบและวิธีในการอธิบายเนื้อหา (Content) ของข้อมูลมัลติมีเดียเท่านั้นโดยไม่ได้กำหนดมาตรฐานในส่วนของวิธีการดึงเนื้อหา วิธีการใช้งานและตัวโปรแกรมที่ใช้ โดยมีเหตุผลที่สำคัญที่ว่าเพื่อให้ผู้ผลิตสามารถที่จะพัฒนาระบบหรือโปรแกรมในการจัดการข้อมูลที่เป็นลักษณะเฉพาะของตนเองได้เพื่อให้เกิดความหลากหลายในการพัฒนาระบบหรือโปรแกรมดังกล่าว ส่งผลให้ไม่เกิดการผูกขาดทางการค้าและเกิดการแข่งขันในการที่จะพัฒนาระบบของตนให้มีประสิทธิภาพออกสู่ท้องตลาดมากขึ้นและเพื่อให้ผู้ใช้สามารถที่จะเลือกใช้ระบบหรือโปรแกรมต่าง ๆ ตามที่ตนเองสนใจได้มากขึ้นจึงเป็นหน้าที่ของผู้พัฒนาที่ต้องพัฒนาระบบในการจัดการกับข้อมูลดังกล่าวขึ้นเอง จากเหตุผลดังกล่าวในปัจจุบันจึงเริ่มมีการพัฒนาระบบสำหรับการจัดการข้อมูลและระบบที่ให้บริการในการใช้ข้อมูลดังกล่าวแก่ผู้ใช้ทั่วไปมากขึ้นโดยเฉพาะระบบที่ให้บริการในการค้นหาข้อมูลภาพผ่านเครือข่ายอินเตอร์เน็ต (Internet) ซึ่งกำลังเป็นที่นิยมของผู้ใช้ทั่วไปอยู่ในปัจจุบัน

จากเหตุผลทั้งหมดดังกล่าวข้างต้นในปัจจุบันจึงมีหลาย ๆ นักวิจัยและผู้พัฒนาระบบต่าง ๆ พยายามที่จะพัฒนาระบบค้นคืนภาพที่มีประสิทธิภาพเพิ่มมากขึ้นโดยพยายามที่จะหาเทคนิคหรือวิธีการใหม่ ๆ มาใช้กับระบบของตนเพื่อให้สามารถแก้ไขปัญหาที่เกิดขึ้นทั้งในเรื่องของความถูกต้องของข้อมูล ความตรงต่อความต้องการของผู้ใช้ การลดภาระการทำงานของผู้ใช้ และการแก้ไขปัญหาช่องว่างระหว่างมุมมองดังกล่าวรวมถึงการพัฒนาระบบของตนให้สามารถสนับสนุนและรองรับมาตรฐาน MPEG-7 ที่กำลังจะกลายเป็นมาตรฐานที่สำคัญของระบบมัลติมีเดียประเภทต่าง ๆ ในอนาคตได้

จากสิ่งที่กล่าวมาทั้งหมดข้างต้นในงานวิจัยนี้จึงมีจุดมุ่งหมายเพื่อพัฒนาระบบค้นคืนภาพเพื่อให้บริการในการสืบค้นข้อมูลภาพแก่ผู้ใช้ผ่านเครือข่ายอินเตอร์เน็ต (Internet) โดยได้

เสนออัลกอริทึมแบบใหม่สำหรับการค้นหาข้อมูลภาพเพื่อแก้ไขปัญหาความตรงต่อความต้องการของผู้ใช้ การลดภาระการทำงานของผู้ใช้และช่องว่างระหว่างมุมมองรวมถึงการออกแบบระบบค้นคืนภาพให้สอดคล้องและรองรับกับมาตรฐาน MPEG-7 ดังกล่าวโดยในงานวิจัยนี้จะนำสถาปัตยกรรมของเครือข่ายเส้นประสาท (Neural Network) มาประยุกต์ใช้สำหรับการเรียนรู้ความเหมือนในระดับสูงและใช้อัลกอริทึมของ K-Mean สำหรับการจัดกลุ่มความเหมือนในระดับต่ำด้วยการใช้ความสามารถในการเรียนรู้ของเครื่องแบบไม่มีผู้สอน (Unsupervised Machine Learning) โดยมีรูปแบบของการทำงานอยู่บนพื้นฐานของ RF ที่เป็นลักษณะประยุกต์ เพื่อลดภาระการทำงานของผู้ใช้ที่ต้องคอยโต้ตอบกับระบบโดยในส่วนของกระบวนการวิเคราะห์ การเรียนรู้ความเหมือนและการจัดกลุ่มความเหมือนนั้น ในงานวิจัยนี้ได้ใช้เทคนิคแบบผสมผสานระหว่างการค้นหาโดยใช้ข้อมูลของภาพในระดับต่ำและการค้นหาโดยใช้ข้อมูลของภาพในระดับสูงเพื่อแก้ไขปัญหาช่องว่างระหว่างมุมมองและความรู้สึกนึกคิดของผู้ใช้และช่วยให้ผลลัพธ์จากการค้นหาตรงต่อความต้องการของผู้ใช้มากขึ้น

## วัตถุประสงค์ของงานวิจัย

1. เพื่อพัฒนาระบบสืบค้นข้อมูลภาพผ่านเครือข่ายอินเตอร์เน็ต (Internet) ที่สามารถรองรับและสนับสนุนมาตรฐาน MPEG-7
2. เพื่อแก้ไขปัญหาช่องว่างระหว่างมุมมอง (Semantic Gab) และความรู้สึกนึกคิดของผู้ใช้ (Subjectivity of User) ที่เกิดขึ้นในระบบสืบค้นข้อมูลภาพ
3. เพื่อลดขั้นตอนการทำงานของผู้ใช้ในการตอบสนองกับระบบในระหว่างกระบวนการค้นหาข้อมูล

## ขอบเขตงานวิจัย

ในการทำการวิจัยนี้ ผู้วิจัยจะทำการศึกษาและออกแบบระบบการค้นคืนภาพด้วยการใช้ความสามารถในการเรียนรู้ของเครื่องแบบไม่มีผู้สอน (Unsupervised Machine Learning) โดยใช้ K-Mean Algorithm ในการจัดกลุ่มและเรียนรู้ความเหมือนของภาพในมุมมองระดับต่ำและใช้หลักการของ RBFN มาประยุกต์สำหรับการเรียนรู้ความเหมือนของภาพในมุมมองระดับสูง โดยจะมีการปรับปรุงวิธีการบางส่วนเพื่อให้เหมาะสมกับวัตถุประสงค์ของงานวิจัยนี้ โดยการค้นหาข้อมูลภาพของทั้งสองส่วนนั้นจะเป็นการทำงานแบบขนาน ซึ่งในงานวิจัยนี้จะกระทำเฉพาะส่วน

ของการค้นคืนข้อมูลภาพกับฐานข้อมูลภาพประเภทภาพในชีวิตจริง (Real Photography) โดยจะทำการทดลองการทำงานของระบบในเครื่องพีซีแทนการทดลองบนอินเตอร์เน็ตซึ่งในส่วนที่มีการปรับปรุงเพิ่มเติมคือการปรับปรุงวิธีการทำงานให้สามารถสนับสนุนและรองรับกับมาตรฐาน MPEG-7 โดยมีการปรับปรุงอยู่ 2 ส่วนหลัก ๆ คือในส่วนของกระบวนการค้นคืนภาพและส่วนของฐานข้อมูลที่ใช้จัดเก็บข้อมูลภาพโดยมีรายละเอียดดังนี้

ในส่วนของกระบวนการค้นคืนภาพจะมีอยู่ 2 ส่วนที่มีลักษณะการทำงานแบบผสมผสานกันระหว่างกระบวนการในการค้นคืนภาพในมุมมองระดับต่ำและกระบวนการในการค้นคืนภาพในมุมมองระดับสูงโดยในงานวิจัยนี้จะใช้ฟีเจอร์ (Feature) ในมุมมองระดับต่ำจำนวน 3 ฟีเจอร์ คือ Color Layout Descriptor , Region Shape Descriptor , Edge Histogram Descriptor ที่ถูกใช้โดย K-Mean Algorithm ซึ่งทั้งสาม Visual Descriptor นี้จะเป็นการเรียนรู้แบบขนานด้วยและอีกส่วนคือใช้ ฟีเจอร์ประเภท Keyword ในมุมมองระดับสูงที่ถูกใช้โดย Single-pass RBFN แบบประยุกต์

โดยในส่วนของโปรแกรมจะเป็นหน้าจอสำหรับค้นหาข้อมูลภาพแบบง่าย ๆ โดยจะมีส่วนของรายการ Keyword ที่ใช้ในการค้นหาในระดับสูงแสดงอยู่ด้านซ้ายของหน้าจอและรายการภาพที่แสดงอยู่ตรงกลางของหน้าจอเพื่อให้ผู้ใช้เลือกซึ่งส่วนของผลลัพธ์หลังจากที่ผู้ใช้เลือกภาพหรือ Keyword ที่ใช้ในการค้นหาแล้วจะแสดงให้ผู้ใช้เห็นผ่านหน้าจอ

ในส่วนของฐานข้อมูลภาพจะทำการจำลองฐานข้อมูลขึ้นใหม่โดยอ้างอิงกับมาตรฐาน MPEG – 7 โดยการเพิ่มข้อมูลที่เป็นเนื้อหาของภาพที่อยู่ในรูปของไฟล์ XML ลงในฐานข้อมูลแบบเชิงสัมพันธ์ (Relational) ซึ่งในงานวิจัยนี้จะใช้โปรแกรม MySQL

ในการสร้างไฟล์ XML นั้นในงานวิจัยนี้จะใช้เครื่องมืออยู่ 2 โปรแกรมคือ M-OnToMat-Annotizer [14] สำหรับสร้าง Content ของภาพในส่วนของมุมมองระดับต่ำและ CaliphEmirSetup_v0.9.21_build327.exe [15] สำหรับสร้าง Content ของภาพในส่วนของมุมมองระดับสูงซึ่งเป็นเครื่องมือที่พัฒนาขึ้นโดยอ้างอิงตามมาตรฐาน MPEG-7

ภาพที่ใช้ในการทดลองนี้จะเป็นภาพที่เกี่ยวกับสภาพแวดล้อมในชีวิตจริง เช่น ภาพธรรมชาติ ภาพคน เป็นต้น เพื่อใช้ประเมินประสิทธิภาพในเรื่องของความถูกต้องของการค้นคืนภาพและความสามารถในการค้นหาที่ครอบคลุมกับความต้องการของผู้ใช้ซึ่งอยู่ในมุมมองระดับสูงโดยแต่ละภาพมีการอธิบายเนื้อหาในระดับสูงไว้ตามประเภทของ Ground Truth เพื่อให้ผู้พัฒนาระบบค้นคืนภาพต่าง ๆ ที่มีการนำฐานข้อมูลภาพเหล่านี้ไปใช้งานสามารถใช้เป็นเกณฑ์ในการวัดประสิทธิภาพการทำงานของระบบของตน

ในส่วนของฟีเจอร์ เครื่องมือ และ Visual Descriptor ที่ใช้จะอยู่ในมาตรฐานของ MPEG-7 โดยในงานวิจัยนี้ในส่วนของมุมมองระดับต่ำจะใช้ ฟีเจอร์ของภาพอยู่ 3 ประเภทคือ สี (Color) รูปร่าง (Shape) และพื้นผิว (Texture) โดยแต่ละประเภทจะใช้ Visual Descriptor เพียง 1 ตัวคือ Color Layout Descriptor สำหรับฟีเซอร์ประเภทสี (Color Feature) Region-Based Shape Descriptor สำหรับฟีเซอร์ประเภทรูปร่าง (Shape Feature) และ Edge Histogram Descriptor สำหรับฟีเซอร์ประเภทพื้นผิว (Texture Feature) ซึ่งเหตุผลที่ใช้เพียงประเภทละหนึ่ง Visual Descriptor นั้นเพียงเพื่อแสดงให้เห็นถึงความเป็นไปได้ของวิธีการและอัลกอริทึมที่เสนอว่ามีประสิทธิภาพมากน้อยแค่ไหนซึ่ง Visual Descriptor อื่น ๆ นั้นสามารถเพิ่มเข้ามาในระบบได้ในอนาคต และในส่วนของมุมมองระดับสูงในงานวิจัยนี้จะใช้ Keyword ประเภท WhatObject ซึ่งเป็น Content ประเภทหนึ่งสำหรับการค้นหาในมุมมองระดับสูง โดยรายละเอียดของมาตรฐาน MPEG-7 และรายละเอียดของ Visual Descriptor ที่ใช้ จะอธิบายอยู่ในบทที่ 2 และ 3 ตามลำดับ

ในส่วนของการประเมินประสิทธิภาพการทำงานของระบบจะแบ่งเป็น 2 ส่วนคือ

1. ส่วนของการวัดประสิทธิภาพการทำงานโดยในส่วนนี้จะมีส่วนที่จะพิจารณาคือเรื่องของอัตราการถูกต้องของการค้นคืน (Retrieval Precision Rate) [1] สำหรับใช้วัดประสิทธิภาพของความถูกต้องในการค้นคืน
2. ส่วนของการพิจารณาความเหมือนกันของภาพจะดู 2 เกณฑ์ [1] คือ
   2.1 Ground Truth
   2.2 Subjectivity of User

โดยในส่วนของการทดลองในงานวิจัยนี้จะมีการเชิญบุคคลมากลุ่มหนึ่งประมาณ 3-5 คนเพื่อให้ทดลองใช้ระบบที่พัฒนาขึ้นเพื่อทำการประเมินประสิทธิภาพการทำงานของระบบในส่วนของการพิจารณาความเหมือนของภาพในเกณฑ์ของ Subjectivity of User เพื่อวิเคราะห์ประสิทธิภาพของระบบเมื่อมีการเพิ่มกระบวนการในการค้นคืนในระดับสูงเข้ามาดังที่ได้กล่าวไว้ในตอนต้นซึ่งเป็นส่วนที่อยู่ในมุมมองเดียวกันกับมนุษย์ว่าระบบนั้นสามารถที่จะแสดงข้อมูลภาพได้ใกล้เคียงกับความต้องการของผู้ใช้ได้มากน้อยเพียงใดโดยจะมีส่วนที่ใช้สำหรับให้บุคคลกลุ่มดังกล่าวระบุว่าผลลัพธ์ที่ได้นั้นตรงกับความต้องการของตนหรือไม่เพียงใดเพื่อนำข้อมูลที่ได้มาวิเคราะห์หาประสิทธิภาพของระบบ และในส่วนของ Ground Truth จะพิจารณาจากค่าของอัตราการค้นคืนโดยพิจารณาจาก Keyword ที่พบจากภาพที่แสดงกับ Keyword ของ Query โดย

ส่วนของรายละเอียดของวิธีการพิจารณาความเหมือนของภาพทั้ง 2 ส่วนดังกล่าวจะขอกล่าวถึงในหัวข้อของวิธีการดำเนินงานวิจัยในบทที่ 3 ต่อไป

**นิยามศัพท์**

1. เครือข่ายเส้นประสาท (Neural Network) คือ เครือข่ายเส้นประสาทซึ่งหมายถึงระบบคอมพิวเตอร์ที่สร้างขึ้นเลียนแบบเส้นใยประสาทของมนุษย์ที่เรียกว่า นิวโรน ในสมองมนุษย์ซึ่งมีนับจำนวนล้าน ๆ ตัวการประมวลผลกิจกรรมต่าง ๆ จะออกมาในลักษณะแบบขนานไม่ใช่แบบอนุกรม หมายถึง ทำงานพร้อม ๆ กันเพื่อให้ได้คำตอบอันเดียวทั้งนี้หมายถึง การที่ต้องสามารถทำงานที่สลับซับซ้อนมาก ๆ โดยเรียนรู้จากประสบการณ์ได้เอง ฉะนั้นการที่เครือข่ายแบบนิวโรนนี้จะให้คำตอบได้ก็ต่อเมื่อมีการส่งปัญหาต่าง ๆ เข้าไปให้ลองทำมาก ๆ เครือข่ายแบบนี้จะเป็นประโยชน์เฉพาะในงานประมวลผลงานบางประเภท เช่น เรื่องการประมวลผลเกี่ยวกับการซื้อขายหุ้นในตลาดหรือการหาความคล้ายคลึงระหว่างภาพหลาย ๆ ภาพ เช่น ลายนิ้วมือ เป็นต้น

2. Radial Basis Function Network คือ เป็นโครงสร้างเครือข่ายเส้นประสาท (Neural Network) ประเภทหนึ่งที่มีการทำงานบนพื้นฐานของรัศมี (Radial) โดยมีลักษณะของการทำงานเป็นแบบ 2 ระยะ (Two Stage) ในการประมวลผล

3. Self-Organizing Tree Map (SOTM) คือ โครงสร้างข้อมูลแบบแผนที่ (Map) ที่เป็นลักษณะแบบต้นไม้ (Tree) สามารถจัดการกับข้อมูลและโครงสร้างข้อมูลได้ด้วยตัวเอง (Self-Organizing)

4. การค้นคืนภาพโดยยึดหลักเนื้อหาภาพ (Content-Based Image Retrieval) คือ วิธีการในการสืบค้นข้อมูลภาพแบบหนึ่งซึ่งจะค้นหาข้อมูลภาพโดยพิจารณาจากเนื้อหา (Content) ของภาพเป็นหลัก

5. ช่องว่างทางความหมาย (Semantic Gab) คือ ความบกพร่องของการเกิดขึ้นพร้อมกันระหว่างข้อมูล (Information) ที่ส่วนหนึ่งสามารถดึง (Extract) มาจาก Visual Data และอีกส่วนหนึ่งเป็นความหมายที่ได้จากการแปลความหมายจากผู้ใช้ ซึ่งทั้งสองส่วนได้มาจากการกระทำบนข้อมูล (Data) ชุดเดียวกัน

6. MPEG-7 คือ มาตรฐาน ISO/IEC ที่ถูกพัฒนาโดย Moving Pictures Experts Group (MPEG) มีชื่ออย่างเป็นทางการคือ Multimedia Content Description Interface โดยมีจุดมุ่งหมายเพื่ออธิบายเนื้อหา (Content) ของข้อมูลมัลติมีเดีย (Multimedia Data) โดยการผูก Metadata เข้ากับ Multimedia Content

7. Extensible Markup Language (XML) คือ ภาษาที่ใช้แสดงข้อมูลต่าง ๆ บนเว็บ เช่นเดียวกับภาษา HTML แต่มีคุณสมบัติที่แตกต่างกันคือภาษา XML สามารถอธิบายถึงความหมายทางโครงสร้างของข้อมูลได้

8. การตอบสนองความสัมพันธ์ [Relevance Feedback (RF)] คือ สถาปัตยกรรมที่ใช้สำหรับจับ (Capture) เป้าหมายในการค้นหาของผู้ใช้ซึ่งเป็นลักษณะของกระบวนการเรียนรู้ (Learning Process)

9. อัลกอริทึมเคมีน (K-Mean Algorithm) คือ อัลกอริทึมสำหรับการจัดการข้อมูล เช่น การจัดกลุ่มข้อมูล (Data Clustering) การจัดประเภทข้อมูล (Data Classification) เพื่อใช้ประโยชน์ในด้านใดด้านหนึ่ง เป็นต้น

**ประโยชน์ที่จะได้รับ**

1. ระบบสืบค้นข้อมูลภาพที่พัฒนาสามารถสนับสนุนและรองรับมาตรฐานของ MPEG-7 ได้

2. ระบบสืบค้นข้อมูลภาพที่พัฒนาสามารถค้นหาข้อมูลภาพได้ตรงกับความต้องการของผู้ใช้มากขึ้น

3. ระบบสืบค้นข้อมูลภาพที่พัฒนาสามารถที่จะช่วยลดปัญหาช่องว่างระหว่างมุมมองของคอมพิวเตอร์กับมนุษย์ (Semantic Gab) และความรู้สึกนึกคิดของผู้ใช้ (Subjectivity of User) ให้น้อยลงได้

4. ระบบสืบค้นข้อมูลภาพที่พัฒนาสามารถที่จะช่วยลดภาระการทำงานของผู้ใช้ที่ต้องคอยโต้ตอบกับระบบในระหว่างกระบวนการค้นหาให้ลดน้อยลงได้

5. ระบบสืบค้นข้อมูลภาพที่พัฒนาสามารถที่จะวิเคราะห์และเรียนรู้ความเหมือนของภาพได้ด้วยตนเองเพื่อให้การค้นหาข้อมูลภาพมีประสิทธิภาพมากขึ้น