บทที่ 3

วิธีดำเนินการวิจัย

ในการทำวิจัยครั้งนี้ผู้วิจัยได้ดำเนินการพัฒนาโปรแกรมสำหรับการค้นคืนภาพโดยได้ทำการศึกษาแนวคิดหลักการรวมถึงเอกสารและงานวิจัยที่เกี่ยวข้องตามที่ได้กล่าวไว้ในบทที่ 2 เพื่อเป็นพื้นฐานความรู้สำหรับการทำวิจัย และได้นำความรู้ที่ได้มาทำการวิเคราะห์ ออกแบบและพัฒนาโปรแกรมเพื่อให้สอดคล้องกับวัตถุประสงค์การดำเนินงานวิจัยโดยมีวิธีการดำเนินงานตามขั้นตอนดังต่อไปนี้

1. การศึกษา
2. การวิเคราะห์
3. การออกแบบ
4. การพัฒนาโปรแกรม
5. การทดสอบโปรแกรม
6. การวัดประสิทธิภาพในการค้นคืนภาพและประเมินผล

1. การศึกษา

ในงานวิจัยนี้ได้ทำการศึกษาเกี่ยวกับแนวคิด ทฤษฎี หลักการและเอกสารงานวิจัยที่เกี่ยวข้องต่าง ๆ ดังที่ได้เสนอไว้ในบทที่ 2 เพื่อนำข้อมูลความรู้ที่ได้มาใช้เป็นพื้นฐานสำหรับการดำเนินงานวิจัยในขั้นตอนอื่น ๆ ถัดไปโดยสรุปรายการของเอกสารงานวิจัยที่เกี่ยวข้องได้ดังต่อไปนี้

1.1 ศึกษาระบบการค้นคืนภาพโดยยึดหลักเนื้อหา [Content-Based Image Retrieval (CBIR) System]

1.2 ศึกษาหลักการของการตอบสนองความสัมพันธ์ [Relevance Feedback (RF)]

1.3 ศึกษาหลักการของโครงข่ายแบบฟังก์ชันพื้นฐานรัศมี [Radial Basis Function Network (RBFN)]

1.4 ศึกษาหลักการของอัลกอริทึมเคมีน (K-Mean Algorithm)

1.5 ศึกษาระบบการโต้ตอบของเครื่องแบบอัตโนมัติ (Automatic Machine Interaction System)

1.6 ศึกษามาตรฐานของ MPEG-7

1.7 ศึกษางานวิจัยเรื่องเครื่องมือสรุปวีดีโอแบบอัตโนมัติ (Automatic Video Summarizing Tool using MPEG-7 Descriptor for Personal Video Recorder)

1.8 ศึกษางานวิจัยเรื่องการใช้ตัวอธิบายเส้นแบบโลคอลที่มีประสิทธิภาพ (Efficient Use of Local Edge Histogram Descriptor)

2. การวิเคราะห์

เพื่อนำข้อมูลที่ได้จากการศึกษามาทำการวิเคราะห์ถึงความเป็นไปได้เพื่อหารูปแบบและวิธีการในการค้นคืนภาพที่เหมาะสม มีประสิทธิภาพและเป็นไปตามมาตรฐานของข้อมูลมัลติมีเดียที่กำหนดคือ MPEG-7 ซึ่งประกอบด้วยการวิเคราะห์วิธีการดังต่อไปนี้

2.1 เทคนิค วิธีการ และหลักการทำงานของระบบค้นคืนภาพ

ในส่วนนี้จะเป็นการนำข้อมูลที่ได้จากการศึกษาเกี่ยวกับเทคนิค วิธีการ และหลักการทำงานของระบบค้นคืนภาพต่าง ๆ มาทำการวิเคราะห์หาข้อดีข้อเสียของเทคนิคหรือระบบเหล่านั้นเพื่อหาเทคนิค วิธีการใหม่ที่มีประสิทธิภาพมากขึ้นและสามารถแก้ปัญหาต่าง ๆ ที่เกี่ยวข้องกับการทำงานให้ลดน้อยลงได้ โดยเทคนิค วิธีการ และหลักการทำงานของระบบค้นคืนภาพที่ได้จากการวิเคราะห์ของงานวิจัยนี้จะอธิบายอยู่ในหัวข้อของ 3.1 หลักการทำงานของระบบค้นคืนภาพ

2.2 การทำงานของโปรแกรม

เนื่องจากการทดลองในงานวิจัยนี้จะดำเนินการบนระบบ Standalone โดยจะเน้นให้เห็นถึงความเป็นไปได้และประสิทธิภาพของระบบค้นคืนภาพที่เสนอเท่านั้นผู้วิจัยจึงใช้โปรแกรมภาษา Visual Basic ในการสร้างส่วนของการติดต่อกับผู้ใช้ การประมวลผลข้อมูล และการแสดงผลข้อมูลรวมถึงการวัดประสิทธิภาพในการค้นคืนข้อมูลด้วย โดยรายละเอียดของการทำงานของโปรแกรมจะอธิบายอยู่ในหัวข้อของ 3.2 การทำงานของโปรแกรม

2.3 ข้อมูล

ข้อมูลในกรณีของงานวิจัยนี้จะหมายถึงข้อมูลภาพที่ใช้สำหรับการค้นหาซึ่งข้อมูลภาพเหล่านี้จะถูกดำเนินการหรือใช้งานโดยอ้างอิงตามมาตรฐาน MPEG-7 ซึ่งเป็นวัตถุประสงค์หลักอันหนึ่งของงานวิจัยนี้ทั้งในเรื่องของการสืบค้น การเปรียบเทียบ การจัดกลุ่ม ความสัมพันธ์ และการทำ Feature Extraction เป็นต้นโดยข้อมูลเหล่านี้จะถูกดำเนินการโดยใช้เครื่องมือต่าง ๆ ที่ถูกพัฒนาขึ้นตามมาตรฐาน MPEG-7 และจะจัดเก็บในฐานข้อมูลโดยอยู่ในรูปของ Code XML

2.4 ระบบฐานข้อมูล

เนื่องจากข้อมูลภาพที่จัดเก็บนั้นค่อนข้างมีจำนวนมากดังนั้นเพื่อให้สามารถรองรับกับข้อมูลจำนวนมากเหล่านี้ได้ในงานวิจัยนี้จึงใช้ MySQL ซึ่งเป็นระบบฐานข้อมูลที่สามารถรองรับข้อมูลที่มีจำนวนมากได้ดีโดยรูปแบบของฐานข้อมูลในการจัดเก็บข้อมูลนั้นจะอ้างอิงกับมาตรฐาน MPEG-7 โดยรายละเอียดของระบบฐานข้อมูลนั้นจะอธิบายอยู่ในหัวข้อของ 3.3 ระบบฐานข้อมูล

2.5 การวัดประสิทธิภาพ

เพื่อให้สอดคล้องกับวัตถุประสงค์ของงานวิจัยและหลักการโดยทั่วไปในการค้นคืนภาพ ในการวัดประสิทธิภาพในการค้นคืนนั้นจะพิจารณาจาก อัตราความแม่นยำเฉลี่ย (Average Precision Rate (APR)) โดยจะอาศัยเกณฑ์ของการพิจารณาความเหมือนหรือความสัมพันธ์ของภาพอยู่ 2 เกณฑ์คือ Ground Truth และ Subjectivity of Human

สำหรับการวัดประสิทธิภาพโดยพิจารณาจากการตรงต่อความต้องการของผู้ใช้จะใช้เกณฑ์ของ Subjectivity ในการพิจารณาโดยจะเชิญบุคคลมาร่วมการทดลองในการใช้โปรแกรมประมาณ 3-5 คนแล้วให้บันทึกผลการประเมินหลังจากใช้งานระบบแล้ว โดยในส่วนของการพิจารณานั้นจะดูว่าภาพที่แสดงนั้นตรงกับความต้องการของผู้ใช้มากเพียงใดโดยดูจากอัตราการค้นคืนเฉลี่ย โดยในส่วนนี้จะใช้กับฐานข้อมูลประเภท Real-Life Photography Database

สำหรับการวัดประสิทธิภาพโดยพิจารณาจากเกณฑ์ของ Ground Truth นั้นจะดูว่าภาพที่แสดงนั้นตรงกับภาพที่ใช้ค้นหาถ้าภาพนั้นอยู่ใน Class เดียวกันกับภาพที่ใช้ค้นหาโดยดูจากอัตราการค้นคืนเฉลี่ย โดยวิธีการวัดประสิทธิภาพนั้นจะอธิบายอยู่ในขั้นตอนที่ 6 การวัดประสิทธิภาพในการค้นคืนภาพและประเมินผล

3. การออกแบบ

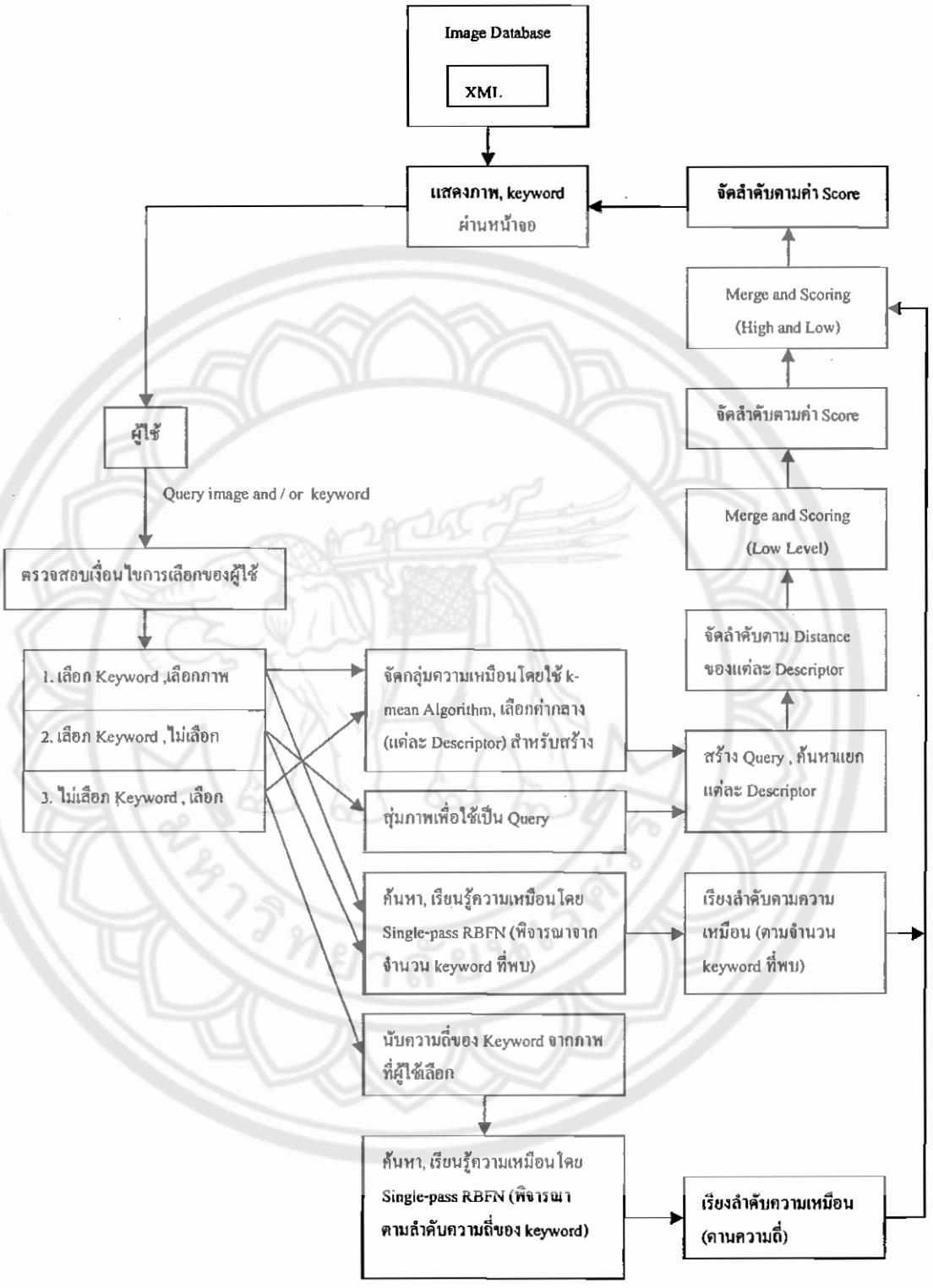
เป็นการนำข้อมูลที่ได้จากการวิเคราะห์มาทำการออกแบบระบบค้นคืนภาพโดยมีรูปแบบและวิธีการทำงานที่สอดคล้องกับวัตถุประสงค์ของงานวิจัยและอยู่ในมาตรฐานที่กำหนดโดยแบ่งออกได้เป็น 3 ส่วนคือ 3.1 หลักการทำงานของระบบค้นคืนภาพ 3.2 การทำงานของโปรแกรม และ 3.3 ระบบฐานข้อมูล โดยมีรายละเอียดดังต่อไปนี้

3.1 หลักการทำงานของระบบค้นคืนภาพ

ในงานวิจัยนี้จะนำเอาหลักการและเทคนิคต่าง ๆ ที่เกี่ยวข้องในการค้นคืนภาพ ดังที่อธิบายในบทที่ 2 มาประยุกต์ใช้ร่วมกับมาตรฐาน MPEG-7 เพื่อพัฒนาระบบการค้นคืนภาพให้เป็นไปตามมาตรฐานที่กำหนดในปัจจุบันโดยได้ออกแบบเทคนิคใหม่ที่ใช้ในระบบการค้นคืนภาพดังกล่าวเพื่อเพิ่มประสิทธิภาพในการค้นคืนภาพโดยรายละเอียดของระบบการค้นคืนภาพที่เสนออธิบายได้ดังต่อไปนี้

3.1.1 โครงสร้างระบบค้นคืนภาพ

โครงสร้างของระบบค้นคืนภาพที่เสนอในงานวิจัยนี้แสดงได้ดังภาพ 21 ข้างล่างนี้



ภาพ 21 แสดงโครงสร้างระบบค้นคืนภาพ

3.1.2 กระบวนการทำงานของระบบค้นคืนภาพ

กระบวนการทำงานโดยรวมของระบบค้นคืนภาพสามารถอธิบายได้ตามลำดับขั้นตอนดังนี้

1) ระบบทำการสุ่มภาพจากฐานข้อมูลภาพเพื่อแสดงผ่านหน้าจอให้ผู้เลือก

2) ระบบทำการดึง Feature ประเภท Keyword จากรายการภาพที่ได้จากข้อ 1 แล้วแสดงรายการ Keyword ทั้งหมดผ่านหน้าจอให้ผู้ใช้เลือก

3) ผู้ใช้ทำการเลือกภาพจากรายการภาพหรือ Keyword จากรายการ Keyword หรือเลือกทั้งสองอย่างตามต้องการ

4) ระบบตรวจสอบเงื่อนไขที่ผู้ใช้เลือกโดยมีเงื่อนไขดังนี้

4.1) ผู้ใช้เลือก Keyword และภาพ ให้ดำเนินการต่อในข้อ 5) และ 11)

4.2) ผู้ใช้เลือก Keyword แต่ไม่เลือกภาพ ให้ดำเนินการต่อในข้อ 6) และ 11)

4.3) ผู้ใช้ไม่เลือก Keyword แต่เลือกภาพ ให้ดำเนินการต่อในข้อ 5) และ 13)

5) ระบบนำภาพที่ผู้ใช้เลือกมาทำการดึง Feature Content ของภาพในระดับต่ำ (Low-Level) ได้แก่ Visual Descriptor Value ของ Visual Descriptor ทั้ง 3 ตัวคือ CLD, EHD และ RSD แล้วนำ Visual Descriptor Value ที่ได้มาจัดกลุ่มความเหมือนโดยใช้ K-Mean Algorithm แล้วเลือกค่ากลางของกลุ่มที่เหมือนที่สุดเพื่อใช้สำหรับสร้าง Query ต่อไปซึ่งกระบวนการทำงานในส่วนของการดึง Feature การจัดกลุ่มและการเลือกค่ากลางของ Visual Descriptor ทั้ง 3 ตัวนี้จะเป็นลักษณะแบบขนาน แล้วดำเนินการต่อในข้อ 7)

6) ระบบทำการสุ่มภาพมา 1) ภาพแล้วดึง Visual Descriptor Value ของ Visual Descriptor ทั้ง 3 ตัวคือ CLD, EHD และ RSD เพื่อใช้สร้าง Query ของ Visual Descriptor แต่ละตัวต่อไป แล้วดำเนินการต่อในข้อ 7)

7) นำ Visual Descriptor Value ของ Visual Descriptor ทั้ง 3 ตัวมาสร้าง Query และทำการค้นหารายการภาพในฐานข้อมูลที่มีค่า Visual Descriptor Value ที่เหมือนกับ Query โดยกระบวนการทำงานของทั้ง 3 Visual Descriptor จะเป็นแบบขนาน

8) นำรายการภาพที่ได้จากการค้นหามาจัดลำดับตามค่า Distance จากน้อยไปมากโดยดำเนินการแยกกันระหว่าง 3 Visual Descriptor ซึ่งจะได้รายการภาพมา 3 กลุ่ม โดยแต่ละกลุ่มจะเลือกรายการภาพที่เหมือนที่สุดมาจำนวนหนึ่ง

9) นำรายการภาพที่ได้ในข้อ 8) ทั้ง 3 กลุ่มมารวม (merge) กันแล้วนับค่า Score ตามความถี่ที่พบซึ่งมี 3 ค่าตามลำดับความสำคัญจากน้อยสุดไปมากสุดคือ 1, 2 และ 3 ตามลำดับ

10) จัดลำดับความเหมือนตามลำดับความสำคัญของค่า Score โดยเรียงจากค่า Score สูงสุดไปน้อยสุดคือ 3, 2 และ 1 ตามลำดับแล้วเลือกรายการภาพที่เหมือนที่สุดมาจำนวนหนึ่ง แล้วดำเนินการต่อในข้อ 16)

11) นำ Keyword จากรายการ Keyword และรายการภาพที่ผู้ใช้เลือกหรือภาพที่ได้จากการสุ่มมาสร้าง Query เพื่อค้นหาภาพในฐานข้อมูลที่มี Keyword ที่เหมือนกับ Query โดยใช้ Single-Pass RBFN ในการเรียนรู้ความเหมือน

12) นำรายการภาพที่ได้จากข้อ 11) มาเรียงลำดับความเหมือนโดยพิจารณาจากจำนวน Keyword ที่พบเช่น ภาพใดมีจำนวน Keyword ที่พบมากสุดถือว่าเหมือนมากสุดแล้วเลือกรายการภาพที่เหมือนที่สุดมาจำนวนหนึ่งแล้วดำเนินการต่อในข้อ 16)

13) นับความถี่ของ Keyword จากรายการภาพที่ผู้ใช้เลือกโดยการดึง Feature ประเภท Keyword ของภาพเหล่านั้นจากฐานข้อมูลแล้วทำการจัดเรียงรายการ Keyword ตามความถี่จากมากไปน้อยโดยความถี่มากสุดถือว่าผู้ใช้สนใจมากที่สุด

14) นำรายการ Keyword ที่ได้จากข้อ 13) มาสร้าง Query เพื่อค้นหาภาพในฐานข้อมูลที่มี Keyword ที่เหมือนกับ Query โดยใช้ Single-Pass RBFN ในการเรียนรู้ความเหมือน

15) นำรายการภาพที่ได้จากข้อ 14) มาเรียงลำดับความเหมือนโดยพิจารณาจากลำดับความถี่ของ Keyword ที่ใช้สร้าง Query แล้วเลือกรายการภาพที่เหมือนที่สุดมาจำนวนหนึ่ง แล้วดำเนินการต่อในข้อ 16)

16) นำรายการภาพที่ได้ทั้ง 2 กลุ่มคือรายการภาพที่ได้จากการค้นหาโดยใช้ Feature Content ในระดับต่ำและรายการภาพที่ได้จากการค้นหาโดยใช้ Feature Content ในระดับสูงมาทำการรวม (Merge) แล้วนับค่า Score ตามความถี่ที่พบซึ่งมี 2 ค่าตามลำดับความสำคัญจากน้อยไปมากคือ 1 และ 2 ตามลำดับ

17) จัดลำดับความเหมือนตามลำดับความสำคัญของค่า Score โดยเรียงจากค่า Score สูงสุดไปน้อยสุดคือ 2 และ 1 ตามลำดับแล้วเลือกรายการภาพที่เหมือนที่สุดมาจำนวนหนึ่ง

18) ทำการดึง Feature Content ประเภท Keyword ของรายการภาพที่ได้จากข้อ 17) จากฐานข้อมูลแล้วแสดงรายการภาพและ Keyword ผ่านหน้าจอ

19) ผู้ใช้ทำการเลือกภาพหรือ Keyword จากรายการที่แสดง

20) เริ่มกระบวนในการค้นหาในข้อที่ 4) ใหม่จนกว่าผู้ใช้จะยุติการค้นหา

3.1.3 กระบวนการดึง Feature Content ของภาพ

ในงานวิจัยนี้จะทำการดึง Feature Content ของภาพมาจากฐานข้อมูลซึ่งถูกเก็บอยู่ในรูปของไฟล์ XML ตามมาตรฐาน MPEG-7 ซึ่งรายละเอียดในการสร้างไฟล์ XML นั้นจะอธิบายอยู่ในหัวข้อ 3.1.10 การสร้างไฟล์ XML จากภาพ โดยวิธีการทำงานของระบบในส่วนนี้จะแบ่งออกเป็น 2 ส่วนหลักคือ

1) การดึง Feature ในระดับ Low Level ซึ่งแบ่งเป็น 3 ส่วนย่อย คือ

1.1) การดึง Feature ประเภทสี (Color Feature) โดยใช้ Color Layout Descriptor

1.2) การดึง Feature ประเภทรูปร่าง (Shape Feature) โดยใช้ Region-Shape Descriptor

1.3) การดึง Feature ประเภทพื้นผิว (Texture Feature) โดยใช้ Edge Histogram Descriptor

ซึ่งค่าของ Feature ของ Visual Descriptor แต่ละตัวที่ได้นั้นจะเรียกว่า Visual Descriptor Value โดยกระบวนการทั้ง 3 ส่วนย่อยนี้จะแยกเป็นอิสระจากกันและจะกระทำไปพร้อม ๆ กันในลักษณะขนาน

2) การดึง Feature ในระดับ High-level ในส่วนนี้จะเป็นการดึงข้อมูลภาพประเภท Keyword ต่าง ๆ ซึ่งในงานวิจัยนี้จะใช้ Keyword ประเภท WhatObject ซึ่งเป็น Feature Content ย่อยตัวหนึ่งของ Content ประเภท StructuredAnnotation ซึ่งสามารถดูตัวอย่างของ Content ประเภทต่าง ๆ ได้ในภาคผนวก ค

สำหรับการดึง Feature ของทั้ง 2 กระบวนการหลัก (High and Low Level) ในข้อ 1) และ 2) ดังที่อธิบายข้างบนจะดำเนินการเป็นแบบขนานและเป็นลักษณะ Real

Time โดยกระบวนการดึง Feature Content ของภาพนี้จะใช้สำหรับการดึง Feature Content ของภาพที่ใช้เป็น Query และภาพในฐานข้อมูลเพื่อใช้ในการประมวลผลและการแสดงผลต่อไป

3.1.4 กระบวนการค้นหา

จากรายการภาพและ Keyword ที่ผู้ใช้เลือกระบบจะทำการดึง Feature Content ที่ได้จากกระบวนการของการดึง Feature Content ในข้อ 1) การดึง Feature ในระดับ Low Level มาใช้สำหรับสร้าง Query เพื่อค้นหา Feature Content ของภาพในฐานข้อมูลซึ่งถูกจัดเก็บอยู่ในรูปของไฟล์ XML โดย Feature Descriptor Set ที่ใช้ในกระบวนการนี้คือ Color Layout Descriptor (CLD) ,Edge Histogram Descriptor (EHD) และ Region Based-Shape Descriptor (RSD) สำหรับการค้นคืนภาพในมุมมองระดับต่ำโดยแต่ละ Feature descriptor จะกระทำไปพร้อม ๆ กันและใช้ Keyword สำหรับการค้นคืนภาพในมุมมองระดับสูงโดยทั้งสองส่วนจะกระทำไปพร้อมกันเป็นแบบขนานโดยมีวิธีการเปรียบเทียบความเหมือนและการเรียนรู้ความเหมือนดังที่อธิบายในหัวข้อที่ 3.1.5 การเปรียบเทียบความเหมือน และ 3.1.6 การเรียนรู้ความเหมือน ตามลำดับ

3.1.5 การเปรียบเทียบความเหมือน

สำหรับการเปรียบเทียบความเหมือนที่ใช้ในกระบวนการค้นหานั้นสามารถแบ่งเป็น 2 ส่วนหลักคือ

1) การเปรียบเทียบความเหมือนในระดับ Low Level

จะพิจารณาจากค่าความแตกต่าง (Distance) ของ Descriptor ระหว่าง 2 ภาพ (ภาพที่ใช้ค้นหากับภาพในฐานข้อมูล) ซึ่งแบ่งเป็น 3 ส่วนย่อยคือ

1.1) การเปรียบเทียบความเหมือนโดยใช้ Color Layout Descripto จะใช้สมการ ในงานวิจัยเรื่อง Automatic Video Summerizing Tool Using MPEG-7 Descriptor for Personal Video Recorder ที่เสนอโดย Jae-Ho Lee, Gwang-Gook Lee และ Whoi-Yul Kim [8] ในการเปรียบเทียบค่า Distance ระหว่าง 2 Color Layout Descriptor ดังที่อธิบายไว้ในบทที่ 2

1.2) การเปรียบเทียบความเหมือนโดยใช้ Region Based-Shape Descriptor จะใช้สมการใน [13] สำหรับการเปรียบเทียบค่า Distance ตามสมการ 15 ดังนี้

$$D = \sum_{i=1}^{35} \| Q_i - I_i \| \quad \text{.....(15)}$$

โดยที่ Region Based-Shape Descriptor ตามมาตรฐานของ MPEG-7 จะกำหนดจำนวนสัมประสิทธิ์เท่ากับ 35 ส่วนตัวแปร Q คือภาพที่ใช้ Query และตัวแปร I คือภาพในฐานข้อมูล

1.3) การเปรียบเทียบความเหมือนโดยใช้ Edge Histogram Descriptor ซึ่งจะใช้สมการใน [13] สำหรับการเปรียบเทียบค่า Distance ตามสมการ 16 ดังนี้

$$D = S_L = \sum_{i=0}^{79} \| Q_i - I_i \| + S_G = \sum_{i=0}^{4} \| Q_i - I_i \| + S_{SG} = \sum_{i=0}^{64} \| Q_i - I_i \| \quad \text{.....(16)}$$

โดยที่ตัวแปร S_L คือผลรวมของ Local Edge Histogram ที่มีจำนวนสัมประสิทธิ์เท่ากับ 80 ส่วนตัวแปร S_G คือผลรวมของ Global Edge Histogram ที่มีจำนวนสัมประสิทธิ์เท่ากับ 5 และตัวแปร S_{SG} คือผลรวมของ Semi-Global Edge Histogram ที่มีจำนวนสัมประสิทธิ์เท่ากับ 65

สำหรับการสร้าง Local Edge Histogram ในงานวิจัยนี้จะดึงมาจากค่าของ Visual Descriptor Value ของ Edge Histogram Descriptor ที่อยู่ในไฟล์ XML ในฐานข้อมูลที่ได้ออกแบบไว้ส่วน Global Edge Histogram และ Semi-Global Edge Histogram จะสร้างขึ้นโดยใช้หลักการที่อธิบายอยู่ในงานวิจัยเรื่อง Efficient Use of Local Edge Histogram Descriptor ที่ถูกเสนอโดย Dong Kwon Park, Yoon Seok Jeon และ Chee Sun Won ซึ่งอธิบายอยู่ในบทที่ 2

2) การเปรียบเทียบความเหมือนในระดับ High Level

ซึ่งมีหลักการง่าย ๆ คือการเปรียบเทียบความเหมือนของคำ (Keyword) ตรง ๆ โดยใช้วิธีพื้นฐานทั่วไป

3.1.6 การเรียนรู้ความเหมือน

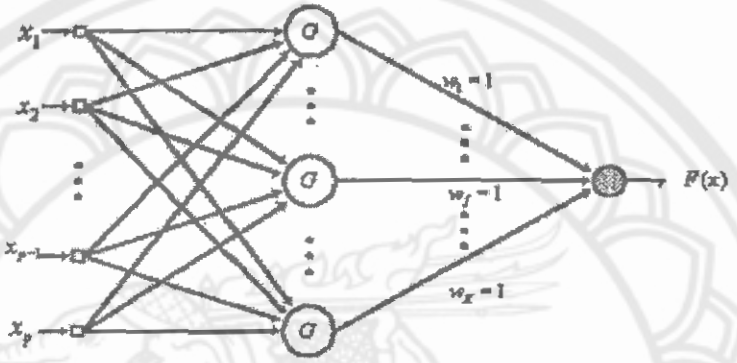
ในส่วนนี้จะแบ่งได้เป็น 2 ส่วนคือ

1) การเรียนรู้ความเหมือนในมุมมองระดับต่ำ

ในงานวิจัยนี้จะใช้ K-Mean Algorithm ที่ถูกพัฒนาขึ้นโดย Kardi Teknomo [5] ดังที่อธิบายในบทที่ 2 สำหรับการจัดกลุ่มความเหมือนของ Visual Descriptor ทั้ง 3 ตัว คือ CLD, EHD และ RSD โดยดำเนินการแยกกันเป็นลักษณะขนานโดยจะนำค่ากลางของกลุ่มที่เหมือนที่สุดของแต่ละ Descriptor ไปใช้สำหรับสร้าง Query สำหรับการค้นหาต่อไปโดยในงานวิจัยนี้จะกำหนดจำนวนกลุ่มที่ใช้ในอัลกอริทึมไว้ที่ 2 กลุ่มสำหรับ Source Code ของ K-Mean Algorithm ที่ใช้นั้นจะแสดงอยู่ในภาคผนวก ข

2) การเรียนรู้ความเหมือนในมุมมองระดับสูง

ในงานวิจัยนี้จะใช้ RBFN [1] ดังที่อธิบายในบทที่ 2 มาประยุกต์ใช้สำหรับการเรียนรู้ความเหมือนในระดับสูงเพื่อให้เหมาะสมกับจุดประสงค์ของงานวิจัยนี้โดยโครงสร้างของ Network ดังกล่าวนี้สามารถแสดงได้ดังภาพ 22 ข้างล่างนี้



ภาพ 22 แสดงโครงสร้างของ RBFN [1]

เพื่อให้สอดคล้องกับจุดประสงค์ของงานวิจัยเราจะทำการปรับปรุงกระบวนการทำงานบางส่วนของ Network ดังกล่าวโดยแบ่งเป็น 2 กรณีซึ่งขึ้นอยู่กับการเลือกของผู้ใช้ดังนี้

1. กรณีผู้ใช้เลือก Keyword จากรายการ

RBFN จะมีหลักการเรียนรู้คือ Keyword ที่ผู้ใช้เลือกจะกำหนดให้เป็นรายการ Keyword ด้วย $Q_K = \{Q_i, Q_i, \ldots, Q_N\}$ ซึ่งจะถูกกำหนดให้กับศูนย์กลางของแต่ละ Neural เพื่อใช้สำหรับการเรียนรู้ความเหมือนและส่วนของ Input คือรายการ Keyword ของภาพในฐานข้อมูลที่ต้องการค้นหา

สำหรับ Hidden Layer แต่ละโหนดจะเป็นส่วนของการตรวจสอบเงื่อนไขของความเหมือนของแต่ละคำ (Word) โดยจำนวนโหนดจะขึ้นอยู่กับจำนวนคำที่ใช้ค้นหา (Query Keyword) ซึ่งสามารถกำหนดเงื่อนไขของความเหมือนและฟังก์ชันในการตรวจสอบเงื่อนไขดังกล่าวได้ดังนี้

กำหนดให้ $S_m(x,w_m)$ เป็นฟังก์ชันในการตรวจสอบเงื่อนไขความเหมือน โดยมีเงื่อนไขดังสมการ 17 คือ

$$S_m(x, w_m) = \begin{cases} 1, Similar \\ 0, Nonsimilar \end{cases} \quad \text{.....(17)}$$

โดย w คือ คำที่ใช้ค้นหา (Query Keyword) , m คือ ดัชนีของรายการคำซึ่งจะมีจำนวนตามจำนวนของคำที่ผู้ใช้เลือก สำหรับการตรวจสอบเงื่อนไขนั้นฟังก์ชัน $S_m(x,w_m)$ จะมีค่าเป็น 1 ถ้าในรายการคำของ x ซึ่งเป็นภาพในฐานข้อมูลมีคำที่เหมือนกับคำ (w) ที่ใช้ค้นหา และ $S_m(x,w_m)$ จะมีค่าเป็น 0 ถ้าไม่มีคำเหมือน โดยค่าความเหมือนของแต่ละคำทั้งหมดจะถูกรวมกันในลักษณะ linear ด้วยฟังก์ชัน F(x) ดังสมการ 18 ดังนี้

$$F(x) = \sum_{m=1}^{M} S_m(x, w_m) \quad \text{.....(18)}$$

ผลของฟังก์ชัน F(x) จะแสดงถึงลำดับความเหมือนของแต่ละภาพซึ่งจะได้รายการของภาพที่เรียงตามลำดับความเหมือนโดยค่าของ F(x) จะมี 2 กรณีคือ มากกว่า 0 หรือเท่ากับ 0 ซึ่งในงานวิจัยนี้จะเลือกภาพจากรายการดังกล่าวมาจำนวนหนึ่งเพื่อใช้ในกระบวนการถัดไป

การจัดลำดับความเหมือนในส่วนนี้จะพิจารณาจากจำนวนของคำของภาพในฐานข้อมูลที่เหมือนกับคำในรายการ Q_K ซึ่งภาพใดที่มีจำนวนคำที่เหมือนมากสุดถือว่าภาพนั้นสำคัญที่สุด

2. กรณีผู้ใช้ไม่เลือก Keyword จากรายการแต่เลือกภาพ

ในส่วนของ RBFN จะมีหลักการเรียนรู้คือระบบจะทำการดึง Feature Content ประเภท Keyword ของรายการภาพที่ผู้ใช้เลือกแล้วกำหนดค่าความถี่ให้กับ Keyword แต่ละตัวโดยพิจารณาจากความถี่ที่พบแล้วเรียงลำดับ Keyword ตามความถี่จากมากไปน้อยซึ่งค่าความถี่มากสันนิษฐานว่าผู้ใช้ให้ความสนใจกับ Keyword นั้นมากโดยรายการ Keyword นี้จะกำหนดให้เป็นรายการ Keyword ด้วย $Q_KF = \{Q_i, Q_i, ..., Q_N\}$ ซึ่งจะถูกกำหนดให้กับศูนย์กลางของแต่ละ Neural เพื่อใช้สำหรับการเรียนรู้ความเหมือนและส่วนของ Input คือรายการ Keyword ของภาพในฐานข้อมูลที่ต้องการค้นหา

สำหรับ Hidden Layer แต่ละโหนดจะเป็นส่วนของการตรวจสอบเงื่อนไขของความเหมือนของแต่ละคำ (Word) โดยจำนวนโหนดจะขึ้นอยู่กับจำนวนคำของแต่ละภาพที่ดึงมาจากฐานข้อมูล (Image Database) และศูนย์กลางของทุก Neural ในการเปรียบเทียบในแต่ละรอบ

จะเป็น Keyword ที่ i จากรายการ Q_KF ซึ่งสามารถกำหนดเงื่อนไขของความเหมือนและฟังก์ชันในการตรวจสอบเงื่อนไขดังกล่าวได้ดังนี้

กำหนดให้ $S_m(w,x_m)$ เป็นฟังก์ชันในการตรวจสอบเงื่อนไขความเหมือนโดยมีเงื่อนไขดังสมการ 19 คือ

$$S_m(w, x_m) = \begin{cases} 1, similar \\ 0, nonsimilar \end{cases} \quad \text{.....(19)}$$

โดย x คือ คำของภาพ O ในฐานข้อมูล (Image Database) , m คือ ดัชนีของรายการคำซึ่งจะมีจำนวนตามจำนวนของคำของภาพ O ในฐานข้อมูลส่วน w คือ Keyword ที่ i จากรายการ Q_KF สำหรับการตรวจสอบเงื่อนไขนั้นฟังก์ชัน $S_m(w,x_m)$ จะมีค่าเป็น 1 ถ้าในรายการคำของ x ซึ่งเป็นภาพในฐานข้อมูลมีคำที่เหมือนกับคำ (w) ที่ใช้ค้นหา และ $S_m(w,x_m)$ จะมีค่าเป็น 0 ถ้าไม่มีคำเหมือน โดยค่าความเหมือนของแต่ละคำทั้งหมดจะถูกรวมกันในลักษณะ linear ด้วยฟังก์ชัน F(O) ดังสมการ 20 ดังนี้

$$F(O) = \sum_{m=1}^{M} S_m(w, x_m) \quad \text{.....(20)}$$

ผลของฟังก์ชัน F(O) จะแสดงถึงความเหมือนของแต่ละภาพซึ่งมีอยู่ 2 กรณีเท่านั้น คือ F(O) มีค่าเท่ากับ 1 แสดงว่าภาพนั้นมีคำที่เหมือนกับคำที่ i ในรายการ Q_KF และ F(O) มีค่าเท่ากับ 0 แสดงว่าภาพนั้นไม่มีคำเหมือนกับคำที่ i ในรายการ Q_KF โดยหลักการทำงานของ Single-Pass RBFN ในส่วนนี้จะเป็นลักษณะวนซ้ำตามจำนวนของคำในรายการ Q_KF ตามลำดับความถี่เพื่อหาภาพที่มีคำเหมือนซึ่งจะได้รายการของภาพที่เรียงตามลำดับความเหมือนตามลำดับความถี่ที่ใช้ค้นหาซึ่งในงานวิจัยนี้จะเลือกภาพจากรายการดังกล่าวมาจำนวนหนึ่งเพื่อใช้ในกระบวนการถัดไป

การจัดลำดับความเหมือนในส่วนนี้จะพิจารณาจากลำดับความถี่ของคำในรายการ Q_K ที่ใช้ค้นหาโดยภาพในฐานข้อมูลใดที่มีคำเหมือนกับคำในรายการ Q_K ซึ่งมีค่าความถี่สูงสุดถือว่าภาพนั้นสำคัญที่สุด

3.1.7 ผลลัพธ์จากกระบวนการค้นหา

จากกระบวนการค้นหาในข้อ 3.1.4 กระบวนการค้นหา นี้จะได้กลุ่มของภาพมา 2 กลุ่มหลักคือ

1) กลุ่มของภาพที่ได้จากการค้นหาในมุมมองระดับต่ำ ซึ่งแบ่งออกได้เป็น 3 ส่วนย่อยคือ

1.1) รายการภาพจากการค้นหาโดยใช้ Color Layout Descriptor กำหนดเป็น $C_L_D = \{O_i, O_i, \ldots, O_N\}$

1.2) รายการภาพจากการค้นหาโดยใช้ Region –Based Shape Descriptor กำหนดเป็น $R_S_D = \{O_i, O_i, \ldots, O_N\}$

1.3) รายการภาพจากการค้นหาโดยใช้ Edge Histogram Descriptor กำหนดเป็น $E_H_D = \{O_i, O_i, \ldots, O_N\}$

โดยรายการภาพของทั้ง 3 ส่วนย่อยดังกล่าวนี้จะมีการพิจารณาลำดับความเหมือนตามกฎของเพื่อนบ้านที่ใกล้ที่สุด (Nearest Neighbor Rule) [1] ของภาพ Query Q ซึ่งภาพที่มีค่า Distance น้อยแสดงว่ามีความเหมือนมากโดยภาพที่เหมือนที่สุดจากการค้นหาลำดับที่ K คือ $O_1, O_2, \ldots, O_K$ สามารถจัดระยะ (Arrange) ได้ดังสมการ 21 [1] ดังนี้

$$D(Q,O_1) \leq \ldots \leq D(Q,O_k) \leq \ldots \leq D(Q,O_K) \quad \ldots\ldots(21)$$

โดยค่า Distance ระหว่าง 2 ภาพสามารถนิยามได้ดังที่แสดงอยู่ในหัวข้อการเปรียบเทียบความเหมือนในมุมมองระดับต่ำทั้ง 3 ข้อดังที่ได้อธิบายอยู่ใน หัวข้อที่ 3.1.5 การเปรียบเทียบความเหมือน โดยแต่ละรายการภาพดังกล่าวในงานวิจัยนี้จะใช้ภาพเพียง 30 ภาพแรกจากการค้นหาโดยถือว่าเป็นกลุ่มของภาพที่เหมือนที่สุด

2) กลุ่มของภาพที่ได้จากการค้นหาในมุมมองระดับสูง

สำหรับกลุ่มของภาพที่ได้จากการค้นหาในส่วนนี้ในงานวิจัยนี้จะใช้จำนวนภาพเพียง 30 ภาพแรกจากการค้นหาโดยถือว่าเป็นกลุ่มของภาพที่เหมือนที่สุดโดยในที่นี้จะกำหนดสัญลักษณ์ของรายการภาพดังกล่าวด้วย $IH = \{O_i, O_i, \ldots, O_N\}$

3.1.8 การรวม (Merge) และการนับค่า Score

ในการรวม (Merge) และการนับค่า Score นั้นจะแบ่งเป็น 2 ส่วนคือ

1) การรวม (Merge) และการนับค่า Score ในมุมมองระดับต่ำ

ในที่นี้คือการนำรายการภาพที่ได้จากการค้นหาในมุมมองระดับต่ำทั้ง 3 รายการจากการค้นหาโดยใช้ 3 Descriptor คือ C_L_D, R_S_D และ E_H_D มารวมกันโดยทำการนับค่า Score การเกิดของภาพแต่ละภาพซึ่งค่า Score นั้นมี 3 ค่าเรียงตามลำดับความเหมือนจากค่ามากไปหาค่าน้อยคือ 3, 2 และ 1 ตามลำดับซึ่งภาพใดที่มีค่า Score มากสุดแสดงว่าภาพนั้นมีความเหมือนมากที่สุดเช่น ภาพ A มีค่า Score เท่ากับ 3 แสดงว่าภาพ A นี้ถูกค้นพบจากการใช้ Descriptor ทั้ง 3 ตัวซึ่งหมายความว่าภาพ A นี้มีความเหมือนมากที่สุดซึ่งเปรียบเทียบกับภาพ B ที่มีค่า Score เท่ากับ 2 เป็นต้น จากนั้นจะเรียงลำดับภาพตามค่า Score จากมากไปน้อยแล้วเลือกรายการภาพที่เหมือนที่สุดมาจำนวนหนึ่งในที่นี้กำหนดให้เป็น IL = $\{O_i, O_i, ..., O_N\}$ เพื่อใช้ประมวลผลในกระบวนการถัดไป

2) การรวม (Merge) และการนับค่า Score ในระหว่างมุมมองระดับต่ำและระดับสูง

ในที่นี้คือการนำรายการภาพที่ได้จากการค้นหาในมุมมองระดับต่ำคือ IL = $\{O_i, O_i, ..., O_N\}$ และรายการภาพที่ได้จากการค้นหาในมุมมองระดับสูงคือ IH = $\{O_i, O_i, ..., O_N\}$ มารวมกันโดยทำการนับค่า Score การเกิดของภาพแต่ละภาพซึ่งค่า Score นั้นมี 2 ระดับเรียงตามลำดับความเหมือนจากมากไปน้อยคือ 2 และ 1 ตามลำดับซึ่งภาพใดที่มีค่า Score มากสุดแสดงว่าภาพนั้นมีความเหมือนมากที่สุดจากนั้นจะเรียงลำดับภาพตามค่า Score จากค่ามากไปหาค่าน้อยแล้วเลือกรายการภาพที่เหมือนที่สุดมาจำนวนหนึ่งเพื่อใช้แสดงผลผ่านหน้าจอและการประมวลผลในกระบวนการถัดไป

3.1.9 เงื่อนไขการประมวลผล

หลังจากที่ผู้ใช้เลือกภาพที่ต้องการแล้วและต้องการค้นหาต่อระบบจะทำการจัดเก็บประวัติการค้นหาของผู้ใช้ไว้ในรายการในฐานข้อมูลเพื่อใช้ในการปรับปรุงกระบวนการทำงานของระบบสำหรับการค้นหาในรอบถัดไปโดยรายละเอียดของข้อมูลการค้นหาที่จัดเก็บนั้นสามารถยกตัวอย่างดังแสดงได้ดังตาราง 3 ข้างล่างนี้

ตาราง 3 รายการประวัติการค้นหาของภาพ

Image List	...	Show Image
A	...	1
B	...	1
C	...	1
...	...	0
...	...	...

จากตาราง 3 แสดงจะมีข้อมูลที่จัดเก็บอยู่ 2 ส่วนหลัก ๆ คือ

1. Image List คือรายการภาพในฐานข้อมูล
2. Show Image คือรายการภาพที่แสดงผ่านหน้าจอ

โดยส่วนของระบบที่มีการอัพเดตนั้นจะมี 2 ส่วนคือ

1. การอัพเดตรายการภาพที่ใช้สำหรับค้นหาโดยมีเงื่อนไขคือภาพในรายการนี้จะถูกสุ่มขึ้นมาให้ผู้ใช้เลือกเพื่อใช้สำหรับค้นหาดังนั้นภาพที่แสดงให้ผู้ใช้เห็นและภาพที่ผู้ใช้ได้เลือกแล้วจะไม่ถูกแสดงขึ้นมาในรายการนี้อีก เพราะถือว่าเป็นภาพที่ผ่านการพิจารณาจากผู้ใช้แล้ว

2. การตรวจสอบรายการภาพในช่วงของกระบวนการค้นหาโดยมีเงื่อนไขคือภาพที่แสดงให้ผู้ใช้เห็นและภาพที่ผู้ใช้เลือกแล้วจะไม่นำมาประมวลผลอีกเพราะถือว่าเป็นภาพที่ผ่านการพิจารณาจากผู้ใช้แล้ว

เหตุผลในการกำหนดเงื่อนไขที่ว่าภาพที่แสดงให้ผู้ใช้เห็นและภาพที่ผู้ใช้เลือกแล้วจะไม่นำมาประมวลผลอีกเพราะถือว่าได้ผ่านการพิจารณาจากผู้ใช้แล้วและเพื่อเพิ่มประสิทธิภาพของการทำงานให้เร็วขึ้นซึ่งจะเป็นผลดีทางอ้อมเมื่อนำไปใช้สำหรับให้บริการสืบค้นข้อมูลภาพบนเว็บ

3.1.10 การสร้างไฟล์ XML จากภาพ

ในงานวิจัยนี้จะใช้ Feature Content ตามมาตรฐาน MPEG-7 ทั้งหมด 3 Feature ในระดับ Low Level คือ Color, Shape และ Texture ซึ่งเป็น Feature หลักของภาพในระดับ Low Level และในระดับ High Level คือ Keyword ใน Content ประเภท WhatObject โดย Feature Descriptor ที่ใช้สำหรับอธิบาย Feature Content ทั้ง 3 ตัวในระดับ Low Level คือ Color Layout Descriptor (CLD) , Edge Histogram Descriptor (EHD) และ Region-Based Shape Descriptor (RSD) โดยเครื่องมือที่ใช้ในการดึง Feature ทั้งหมด

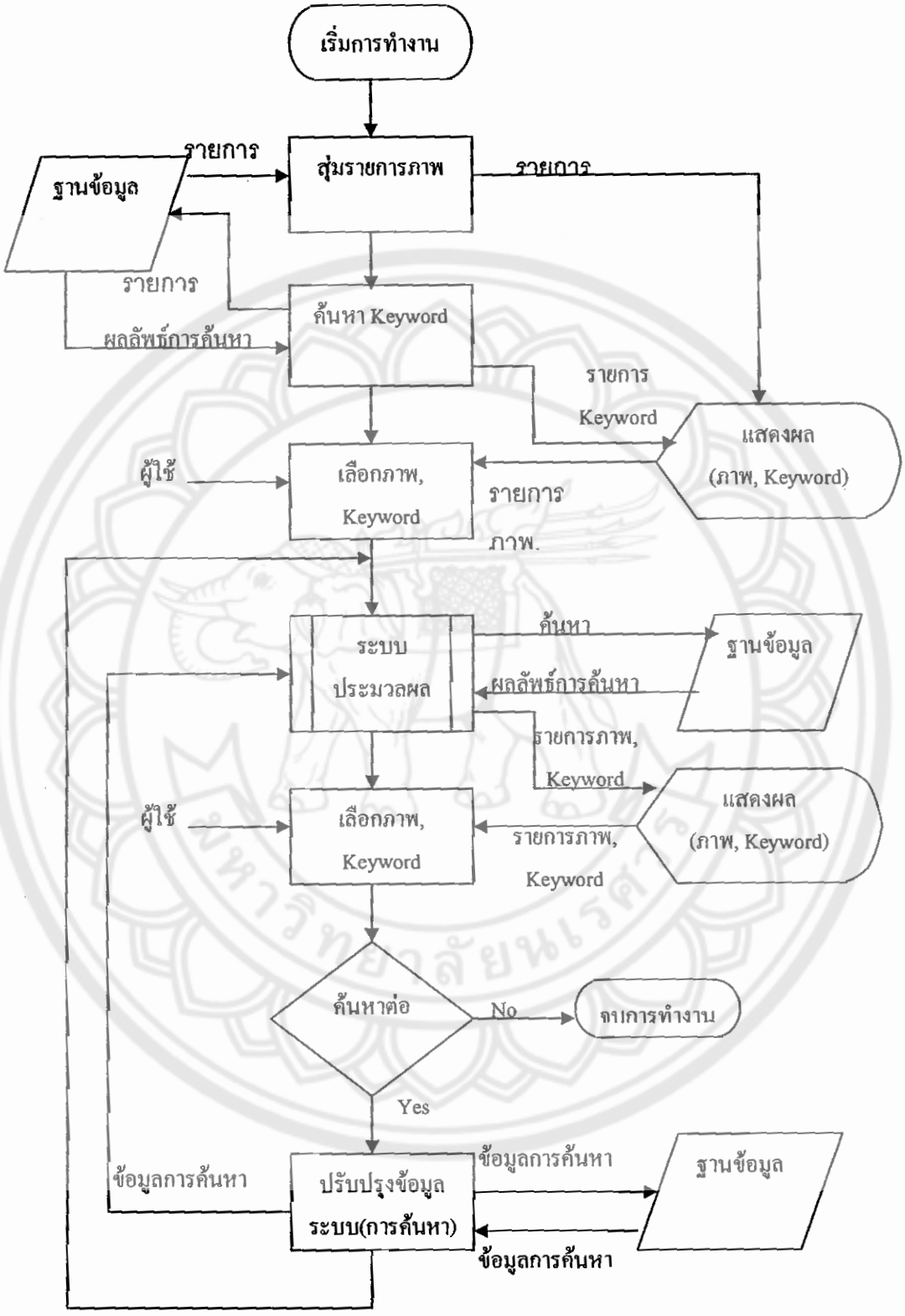
ดังกล่าวจะใช้เครื่องมือที่พัฒนาขึ้นตามมาตรฐานของ MPEG-7 โดยมี 2 ตัวคือ M-OnToMat-Annotizer [14] ซึ่งใช้สำหรับการดึง Feature Content ในระดับ Low Level และ CaliphEmirSetup_v0.9.21_build327.exe ซึ่งใช้สำหรับการดึง Feature Content ในระดับ High Level ซึ่งตัวอย่างของไฟล์ XML ของ 3 Visual Descriptor และ Keyword ที่ได้จากทั้ง 2 โปรแกรมนั้นจะแสดงอยู่ในภาคผนวก ค และรายละเอียดของ Feature Descriptor ที่ใช้จะแสดงอยู่ในตาราง 4

ตาราง 4 แสดงคำอธิบายของ Visual Descriptor

Color Layout Descriptor	อธิบายถึงการกระจายตัวของสี (Spatial Distribution of color) ในลักษณะ visual signals ที่มีรูปแบบที่กระชับ (Compact Form) ซึ่งความกระชับนี้จะช่วยให้การทำ Matching ในการค้นคืน (Retrieval) นั้นมีประสิทธิภาพมากขึ้นในขณะที่มีค่า Cost ที่ต่ำ
Region-Based Shape Descriptor	อธิบายถึงรูปร่างของวัตถุ (Object) ซึ่งอาจประกอบด้วยขอบเขต (Region) เพียงขอบเขตเดียวหรือประกอบด้วยกลุ่มของขอบเขตหลายขอบเขตหรือขอบเขตที่ประกอบด้วยหลุมและพื้นที่ (Area) ที่ไม่เชื่อมต่อกัน (Disjointed Areas) ซึ่ง Region-Based Shape Descriptor จะทำงานได้ดีกับภาพที่มีขอบเขตที่เชื่อมต่อกันเป็นวัตถุเดียวกัน
Edge Histogram Descriptor	อธิบายถึงการกระจายตัวของเส้น (Edge) ทั้ง 5 ประเภทคือเส้นที่มีทิศทาง (Directional Edges) ได้แก่แนวตั้ง (Vertical) แนวนอน (Horizontal) 45 องศา (45 Degrees Diagonal) และ 135 องศา (135 Degrees Diagonal) และเส้นที่ไม่มีทิศทาง (Non-Directional Edge) เนื่องจากเส้นนั้นจะแสดงถึงลักษณะสำคัญของภาพดังนั้นจึงสามารถใช้สำหรับค้นหาภาพที่มีความหมายในเชิงความหมาย (Semantic Meaning) ที่คล้ายกันได้จึงเหมาะกับการค้นหาแบบ Image-to-Image Matching เช่นการค้นหาโดยใช้ตัวอย่าง (by Example) หรือโดยใช้ภาพสเก็ต (by Sketch) โดยเฉพาะภาพธรรมชาติ (Natural Images) ที่มีการกระจายตัวเส้นที่ไม่คงที่ (Non-Uniform Edge Distribution)

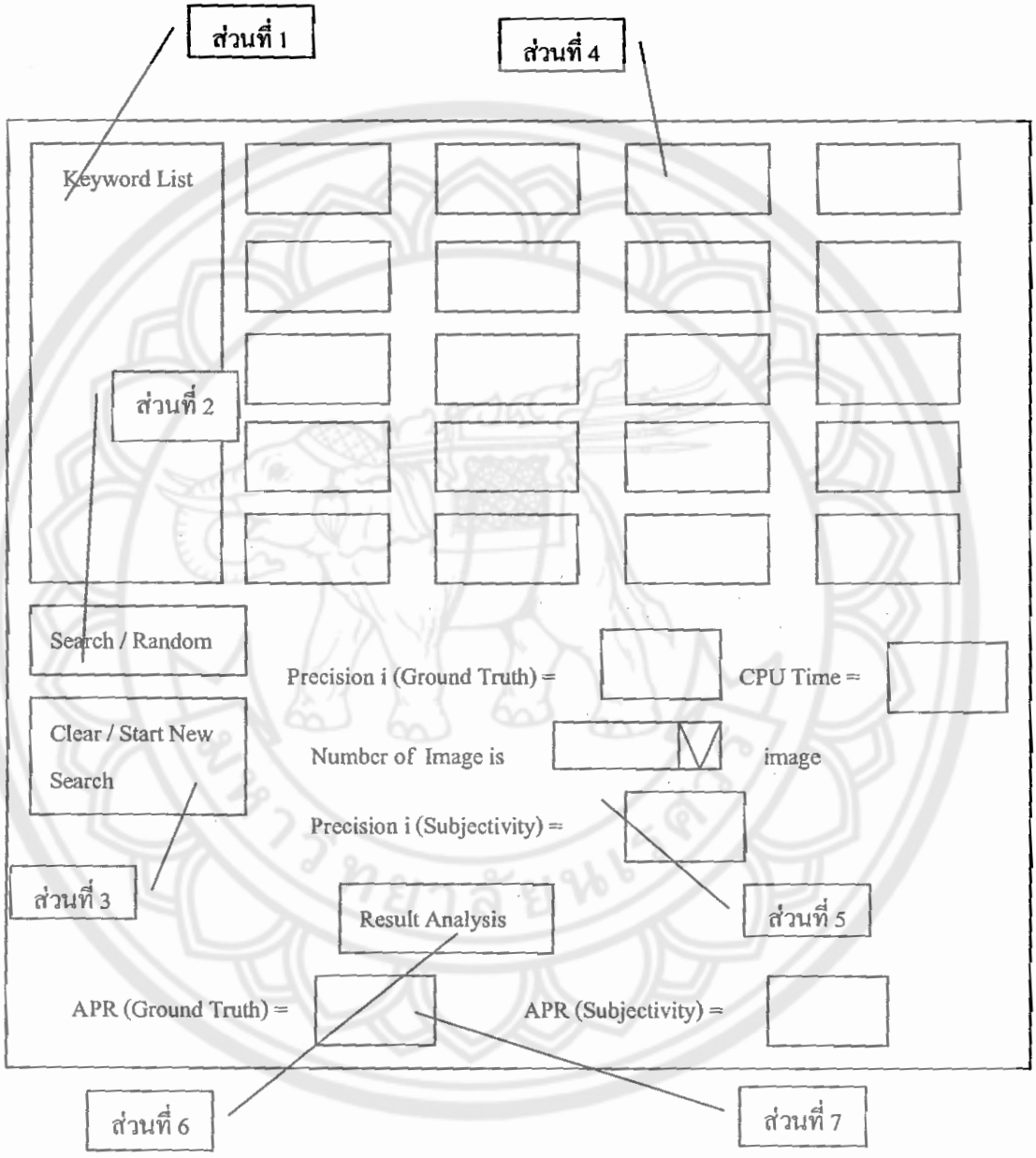
3.2 การทำงานของโปรแกรม

กระบวนการทำงานของโปรแกรมจะเริ่มจากระบบทำการสุ่มภาพมาจำนวนหนึ่งและดึง Keyword ของภาพเหล่านั้นแล้วแสดงภาพและ Keyword ให้ผู้ใช้เลือกผ่านหน้าจอหลังจากนั้นผู้ใช้เลือกภาพและ Keyword ที่ต้องการจากรายการภาพและรายการ Keyword ดังกล่าว จากนั้นระบบจะนำภาพและ Keyword ดังกล่าวมาทำการประมวลผลตามหลักการทำงานที่ออกแบบไว้ในหัวข้อ 3.1 หลักการทำงานของระบบค้นคืนภาพ โดยระบบจะทำการประมวลผลทั้งข้อมูลในระดับต่ำ (Low Level) และข้อมูลในระดับสูง (High Level) ไปพร้อม ๆ กันในลักษณะขนานหลังจากเสร็จสิ้นการประมวลผลระบบจะทำการแสดงรายการภาพและ Keyword ที่ได้จากการค้นคืนให้ผู้ใช้เห็นผ่านหน้าจอหลังจากนั้นผู้ใช้จะทำการเลือกภาพที่ต้องการหากผู้ใช้ต้องการค้นหาต่อก็สามารถทำการค้นหาใหม่ได้โดยระบบจะทำการอัพเดตข้อมูลการค้นหาและเริ่มกระบวนการค้นหาในรอบถัดไปซึ่งกระบวนการทำงานของระบบโดยรวมสามารถอธิบายดังที่แสดงอยู่ในภาพ 23



ภาพ 23 แสดงกระบวนการทำงานของโปรแกรม

ส่วนของโครงสร้างหน้าจอการทำงานของโปรแกรมที่ออกแบบนั้นจะประกอบด้วยโครงสร้างแสดงได้ดังภาพ 24



ภาพ 24 แสดงหน้าจอสำหรับค้นหา

หน้าจอจะประกอบด้วย 7 ส่วนหลักคือ

ส่วนที่ 1 คือส่วนของรายการ Keyword สำหรับผู้ใช้เลือก

ส่วนที่ 2 คือปุ่มสำหรับค้นหาและการสุ่ม

ส่วนที่ 3 คือปุ่มสำหรับเริ่มกระบวนการค้นหาใหม่

ส่วนที่ 4 คือรายการภาพที่แสดงให้ผู้ใช้เห็นและเลือก

ส่วนที่ 5 คือส่วนที่แสดงผลความแม่นยำของการค้นหาในแต่ละรอบและเวลาในการทำงาน

ส่วนที่ 6 คือปุ่มที่ใช้วิเคราะห์ประสิทธิภาพจากการค้นหา

ส่วนที่ 7 คือส่วนที่ใช้แสดงผลของอัตราความแม่นยำเฉลี่ยของการค้นหา

โดยในส่วนนี้จะทดสอบกับฐานข้อมูลประเภท Real-Life Photography Database เท่านั้นเนื่องจากภาพประเภทนี้สามารถใช้วัดประสิทธิภาพการค้นคืนในมุมมองระดับสูงได้ดีซึ่งเป็นมุมมองเดียวกับมนุษย์โดยจะให้ผู้ใช้ระบุจำนวนภาพที่ตรงกับที่ต้องการจากนั้นจึงทำการวัดประสิทธิภาพโดยกดปุ่ม "Analysis" โดยผลการวัดประสิทธิภาพจะแสดงอยู่ด้านล่างของส่วนนี้

3.3 ระบบฐานข้อมูล

เพื่อให้ระบบฐานข้อมูลสามารถจัดเก็บข้อมูลที่อยู่ในมาตรฐานของ MPEG-7 ได้และเพื่อให้สอดคล้องกับวัตถุประสงค์ของงานวิจัยในการวัดประสิทธิภาพทั้งในเรื่องของความครอบคลุมกับ Content ของภาพในระดับต่ำและความตรงกับความต้องการของผู้ใช้ในระดับสูงในงานวิจัยนี้จึงได้มีการออกแบบฐานข้อมูลใหม่เพื่อให้สามารถจัดเก็บข้อมูลดังกล่าวได้ซึ่งในฐานข้อมูลนี้จะประกอบด้วยข้อมูล 3 ส่วนหลักคือ ส่วนแรกคือรายการภาพ ส่วนที่สองคือ Code XML File ที่เก็บ Content ของภาพเหล่านั้นได้แก่ Visual Descriptor Value และ Keyword และส่วนที่สามคือ ประวัติการค้นหา ซึ่งสองส่วนหลังจะเป็นสองส่วนหลักที่ใช้ในกระบวนการค้นหา เปรียบเทียบความเหมือน และการจัดกลุ่มความสัมพันธ์ ทั้งในระดับต่ำและระดับสูงเป็นต้นซึ่ง Content ของภาพดังกล่าวนี้จะถูกสร้างโดยใช้เครื่องมือที่อยู่ในมาตรฐาน MPEG-7 ดังที่อธิบายในหัวข้อ 3.1.10 การสร้างไฟล์ XML จากภาพ โดยโครงสร้างของตารางในระบบฐานข้อมูลนี้จะมีโครงสร้างแบบง่าย ๆ ซึ่งประกอบด้วยฐานข้อมูลที่เก็บภาพประเภทภาพในชีวิตจริง (Real-Life Photography) โดยตารางจะมีรายละเอียดของโครงสร้างดังนี้

1. Real-Life Photography Table

เป็นตารางที่เก็บข้อมูลภาพในชีวิตจริงโดยมีโครงสร้างดังตาราง 5 ดังนี้

ตาราง 5 โครงสร้างตารางฐานข้อมูล

Name	Description	Type
Id	รหัสของภาพ	Number
CLD	Code XML, Content ของภาพประเภท color	File
RSD	Code XML, Content ของภาพประเภท shape	File
EHD	Code XML, Content ของภาพประเภท texture	File
H_L_D	Code XML, Content ของภาพประเภท Keyword	File
Show Image	ค่าการแสดงผลของภาพ(1=แสดง, 2=ไม่แสดง)	Number
ImageName	ชื่อของภาพ	Text
GroundTruth	รายการ Keyword ของภาพ	Text

4. การพัฒนาโปรแกรม

ขั้นตอนนี้จะเป็นส่วนของการพัฒนาตัวโปรแกรมทั้งในส่วนของตัวโปรแกรมสำหรับค้นหาข้อมูลและระบบฐานข้อมูลตามที่ได้ออกแบบไว้จากขั้นตอนการออกแบบโดยใช้โปรแกรมต่าง ๆ ตามที่ได้อธิบายไว้ในขั้นตอนของการวิเคราะห์

5. การทดสอบโปรแกรม

เป็นส่วนของการทดสอบการทำงานของโปรแกรมว่ามีความผิดพลาดในการทำงานในส่วนใดบ้างถูกต้องหรือไม่และเป็นไปตามที่ได้ออกแบบไว้หรือไม่เช่น การแสดงรายการภาพที่ใช้ค้นหา การประมวลผล หรือการแสดงรายการภาพที่ได้จากการค้นหา เป็นต้น

6. การวัดประสิทธิภาพในการค้นคืนภาพและประเมินผล

เป็นขั้นตอนของการวัดประสิทธิภาพของการค้นคืนข้อมูลของโปรแกรมตามเกณฑ์ในการวัดประสิทธิภาพที่ได้กำหนดไว้โดยมี 2 ขั้นตอนคือ

6.1 การทดสอบโดยผู้วิจัยก่อนเพื่อดูว่าระบบนั้นทำงานถูกต้องหรือไม่และได้ผลมากน้อยเพียงใดในด้านของประสิทธิภาพการค้นคืนซึ่งการทดสอบโดยผู้วิจัยนี้ถือว่าเป็นการทดสอบความผิดพลาดของโปรแกรมในอีกระดับหนึ่งเพื่อแก้ไขข้อผิดพลาดดังกล่าวเพื่อให้โปรแกรมมีความสมบูรณ์มากขึ้นก่อนที่จะนำไปให้ผู้ใช้ใช้งาน

6.2 การทดสอบโดยเชิญบุคคลมากลุ่มหนึ่ง (ประมาณ 3 – 5 คน) เพื่อมาทดลองใช้งานระบบและวัดประสิทธิภาพของผลการค้นคืนโดยพิจารณาจาก 2 เกณฑ์ที่กำหนดไว้คือ Ground Truth และ Subjectivity of User

สำหรับการวัดประสิทธิภาพทั้ง 2 เกณฑ์นั้นในงานวิจัยนี้จะพิจารณาเฉพาะในมุมมองระดับสูงเท่านั้นโดยในการวัดประสิทธิภาพโดยใช้เกณฑ์ Ground Truth จะพิจารณาจาก Keyword ของภาพที่แสดงว่าอยู่ใน Class เดียวกับภาพหรือ Keyword ที่ใช้ค้นหาหรือไม่ถ้าอยู่ใน Class เดียวกันแสดงว่าภาพนั้นมี Keyword ที่เหมือนกับ Keyword ของภาพหรือ Keyword ที่ใช้ค้นหาซึ่งการวัดโดยใช้เกณฑ์นี้จะวัดทุกครั้งของการค้นหาโดยระบบจะตรวจสอบจาก Class ของ Keyword ที่เก็บอยู่ในฟิลด์ GroundTruth ในตารางฐานข้อมูลดังตาราง 5

ส่วนการวัดประสิทธิภาพโดยใช้เกณฑ์ Subjectivity นั้นจะวัดโดยให้ผู้ใช้ระบุจำนวนภาพที่ตรงกับที่ต้องการในแต่ละครั้งของการค้นหา

สำหรับการวัดประสิทธิภาพของทั้ง 2 เกณฑ์นั้นจะพิจารณาจาก อัตราความแม่นยำเฉลี่ย (Average Precision Rate (APR)) ซึ่งมีสมการดังนี้

$$APR = \frac{\left[\sum_{i=1}^{M} P_i\right]}{M} \quad \text{.....(22)}$$

โดย Precision (P_i) นิยามได้ดังนี้

$$\text{Precision}(P_i) = \frac{\text{no.Relevant image, Nr}}{\text{no.Retrievaled image, } N_{RT}} \times 100\% \quad \text{.....(23)}$$

ซึ่ง M คือจำนวนรอบของการค้นหา N_r คือจำนวนภาพที่ตรงกับที่ค้นหาและ N_{RT} คือจำนวนภาพที่แสดงซึ่งค่าของ N_{RT} นี้จะลดลงตามจำนวนภาพที่แสดงในแต่ละครั้งของการค้นหาตามเงื่อนไขในการค้นหาที่ตั้งไว้คือภาพที่แสดงแล้วจะไม่นำมาประมวลผลอีกดังที่อธิบายในหัวข้อ 3.1.9 เงื่อนไขการประมวลผล

สำหรับการวัดประสิทธิภาพในด้านของเวลาในการทำงานนั้นในงานวิจัยนี้จะวัดจากเวลาของ CPU Time ซึ่งจะกำหนดอยู่ในตัวโปรแกรมที่พัฒนา

อุปกรณ์และเครื่องมือ

ในการศึกษา ออกแบบและพัฒนาระบบค้นคืนภาพดังที่อธิบายในงานวิจัยนี้ ผู้วิจัยได้ใช้อุปกรณ์และเครื่องมือต่าง ๆ ในการดำเนินงานดังที่อธิบายไว้ในหัวข้อต่อไปนี้

1. Hardware มีรายละเอียดดังนี้
   1.1 เครื่องไมโครคอมพิวเตอร์ Intel Pentium3
   1.2 Hard disk 20 GB
   1.3 Ram 256 Mb
   1.4 อุปกรณ์ต่อพ่วง (คอมพิวเตอร์) ต่าง ๆ
   1.5 เครื่องพิมพ์ HP LaserJet
   1.6 อื่น ๆ เช่น กระดาษ A4, แผ่น Disket, อุปกรณ์สำนักงานต่าง ๆ ฯลฯ
2. Software
   2.1 โปรแกรมภาษา Microsoft Visual Basic 6.0
   2.2 โปรแกรม MySQL 4.1
   2.3 โปรแกรม Microsoft XP (Professional)
   2.4 โปรแกรมในชุด Microsoft Office XP
   2.5 เครือข่าย Internet

เครื่องมือที่ใช้ในมาตรฐาน MPEG-7 ได้แก่

1. โปรแกรม M-OnToMat-Annotizer
2. โปรแกรม CaliphEmirSetup_v0.9.21_build327.exe