

Digital Library Curriculum Development

Module (7-a) Image Retrieval

(Draft Last Modified: 12/09/2009)

1) Module Name

Image Retrieval

2) Scope

The module covers basic explanation of Image Retrieval, various techniques used, and its working in existing systems.

3) Learning Objectives

At the end of the module, student will be able to

- 1) Explain the basics of Image Retrieval and its needs in digital libraries.
- 2) Explain the different Image Retrieval types.
- 3) Classify Content Based Image Retrieval based on various features and differentiate between the techniques being used.
- 4) Explain the various Image Retrieval systems in existence and evaluation strategies.

4) 5S Characteristics of the Module

- **Streams:** Text and visual queries which are inputs and images that serve as the output of an Image Retrieval system. These are sequences of characters or bits.
- **Structures:** Images are stored in dynamic organizational structures such as R Trees and R* Trees aiding in easy retrieval.
- **Spaces:** Feature spaces for images that are indexed. The screen of an interface, images in 2D spaces.
- **Society:** End users of the system such as librarians, journalists, and curator of image collections and software agents who process the information.

- **Scenarios:** Situations where users submit images, create collections, search and retrieve images, invoke relevance feedback and submit queries by sample images.

5) Level of Effort Required:

- a. Prior to Class: 2 hours of reading
- b. In Class: 3 hours
 - a) 2 hours for learning the concepts of image retrieval
 - b) 1 hour for class discussions and activities

6) Relationship with other modules:

Close connections with:

- *6-a: Info Needs, Relevance*

6-a explains how information needs and the state of people's minds evolve when looking for information, browsing and retrieving information. It is related to 7-a (1) where the user searches for images based on their information needs.
- *6-b: Online info seeking and search strategy*

6-b covers the theories, models, and practices related to online information seeking behaviors in different settings of digital libraries. This is related to 7-a (1) where it details the search strategies for images.

7) Prerequisite knowledge required:

Basic knowledge of internet, images and computer systems is required.

8) Introductory Remedial Instruction:

None

9) Body of Knowledge

I. Image Retrieval and its Needs

The below architecture explains the architecture of an image retrieval system. The user uses the query interface to submit the query which is processed and browses the image collection to extract the visual features or the texts. This is based on the type of the image retrieval system being used.

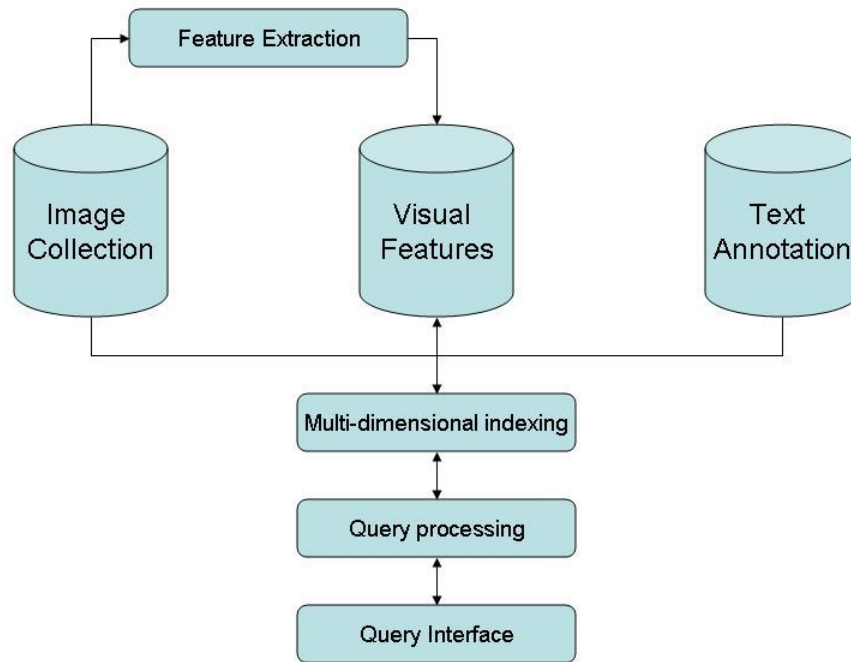


Figure 1: Architecture of an Image Retrieval System

a) Need for image retrieval systems

The need for a desired image is by groups such as artists, designers, teachers, historians, advertising agencies, photographers, engineers and journalists. The requirement for images and corresponding use varies considerably among the groups mentioned above. Before the advent of digitized images, librarians and archivists were providing access to images through text descriptors and classification codes manually.

b) Challenges faced by Image Retrieval technologies

The basic features that users look for in images include color, shape or texture [2]. With the huge amount of information present over the internet, it needs to be organized efficiently for effective browsing search and retrieval. The knowledge on how the systems interact with the visual information is required to further develop an understanding of Image Retrieval systems.

II: Basic Image Retrieval Types

The image retrieval techniques focus on two aspects of image research. They are text based and content based image retrieval techniques. These topics are discussed in detail in the following sub-sections.

(a) Text based Image Retrieval

(i) How Text based Image Retrieval works

Text based image retrieval has been employed from the 1970s. Collections of images are annotated. The annotation of images employs keywords describing the image, caption of an image, text surrounding/embedded in it, complete text of the containing page, and its filename. These systems use text based Database Management systems to retrieve them. These solutions assume a high degree of relevance of the text with the image.

(ii) Textual descriptors

The annotations are based on semantic reasoning that is done in two ways, 1) made by human manually with the aid of tools and 2) made by automatically relating the semantics of the terms with relevant descriptions. Another approach followed in adding text to images is to tag images based on its metadata. Metadata (data about data) contains attributes of the image such as image creation date, image creator, format and simple descriptions of the digital object.

(iii) Images in databases

After annotation, the various categories and the images are indexed for easy retrieval. For example, Getty's Art and Architecture Thesaurus maintained over 150,000 terms under topics such as history, architecture and cultural objects and images. Similar collections were used as databases to retrieve

images. Running a search query on these databases fetches images that are related to query terms.

(iv) Problems faced in Text-based Image Retrieval

Manually adding annotations and textual attributes is difficult and the amount of man-hours it consumes is huge. Also, the probability of disagreement between indexers is high and this adds to the existing problems in text based image retrieval. For example, an image could be interpreted in many ways and that could lead to conflicts in indexing. Though there have been examples of automatic assignment of attributes, lack of captions or accompanying text accounts as a disadvantage for the text based system. Studies show that user's needs for images is predominantly related to visual features and hence the need arises for an image retrieval system based on visual features such as shape, color and texture arises.

The following sub-sections describe the various content based techniques in existence.

(b) Content-based Image Retrieval

The queries for CBIR systems include image attributes or visual examples such as a sketch, texture palette, etc. Following are some of the common image needs and patterns of use such as searches for a particular image, pictures that are related to a story or an article, searches for images that illustrate a document, similarity searches, etc. The primitive features in images that are of importance to image retrieval are color, shape and texture [2]. Some of the other features include spatial and face recognition feature extractions [10].

(i) Color feature based retrieval

Image retrieval systems extract the color feature and similarity measures from images.

- *Color Histogram*

Color Histogram is a commonly used color feature representation. It basically calculates the intensity levels of the three color channels (Red, Green and Blue) and denotes their joint probability. Histograms can be represented as one 3-D histogram or three separate 1-D histograms. These

color representations holds irrespective of the rotation and translation variants of the image.

Using the mathematical representations of these histograms, the similarity of the queried image and the stored images in the database is calculated based on two methods.

- *Histogram Intersection*

This method calculates the normalized color histograms of each image in the database with the normalized color histograms of the queried image. The similarity value that ranges between 0 and 1 determines the extent of the relationship between the images.

- *Distance Method*

Calculation of the Euclidean distance between the feature vectors calculated from the histograms can also be used as a similarity measure. The lower the value, the higher the similarity between the images.

- *Color Moments and Color Sets*

The lower order moments such as first moment (mean) and the second and third moments (variance and skewness) are extracted from the images and checked for similarity. Additionally, binary feature vectors such as Color sets which are created by the transforming the color space into uniformed space such as HSV (Hue-Saturation Value) and then quantizing the color space into bins called Color sets. These color sets are binary vectors and binary tree search can be used for faster execution.

(iii) Shape based image retrieval

Shape is an additional feature that should be extracted if the colors of the images are the same throughout. Shape features are based on two methods, Boundary based and Region based. Boundary based methods uses the outer boundary characteristics of the image/objects while the region-based involve the entire region. Shape features may also be local or global. A shape feature is local if it is derived from some proper subpart of an object, while it is global if it is derived from the entire object.

- *Histogram method*

Identification of features such as boundaries, lines, circularities and edge detections are primary areas concerning boundary based methods. Similar to color based techniques, these methods use histograms depicting the edge directions. A histogram intersection technique matches the parts of the images and is used to calculate the similarity values.

(c) Texture based image retrieval

Texture based techniques are classified into three categories; probabilistic, spectral and structural. In probabilistic, texture patterns are treated as certain random fields; features are extracted from these fields. Spectral methods involve decomposing of images into channels and analyzing them to extract relevant features. Structural techniques model texture features based on heuristic rules of primitive image elements that mimics human perception of textural patterns.

Visual textual features such as coarseness, contrast and roughness are used in quantifying the textual patterns in an image.

- *Image segmentation*

Images containing textures usually contain multiple objects. Hence, images are segmented into regions of similarity using image segmentation algorithms or tools. Each region represents a set of pixels coherent in their local texture properties and its corresponding texture descriptor is defined. These descriptors are used to determine the similarity between the regions. Image segmentation is also used in Color based schemes.

III: Indexing in Image Retrieval

Indexing techniques are useful when images are stored in large databases and size of the images stored grows exponentially. Images are represented as feature vectors in image retrieval systems.

- *Multi-dimensional indexing*

Multi-dimensional feature vectors are stored in trees that contain special properties. They are represented in structures that contain aspects such as:

- 1) Leaves at the same level contain entries of the form (reference, feature vector) where reference is the pointer to the image in the database.

- 2) Presence of intermediate nodes including area representation.
 - 3) Every node of the tree has a specification of the capacity it can hold.
- R-Trees are dynamic organizational tree structures that conform to the above aspects such that each node has a variable number of entries and each non-leaf node contains a child node and a bounding box.
 - R* Trees are a variant of R-Trees where they contain better storage efficiency by means of margin minimization and better spatial distribution of the tree entries. QBIC is an example Image Retrieval system using this structure.
 - SS-Tree is a dynamic organization where the area field of the node entry is indicated with a centroid and radius, thus accounting for less space with more area information.
 - M-Tree or metric tree is a data structure that suits indexing of generic metric spaces.
 - Feature vectors also use Incremental Clustering Techniques for dynamic information retrieval. They have a dynamic structure which is capable of handling high dimensional data and deals with Non-Euclidean similarity measures.

Tools such as Self Organization Map (SOM) are used to construct the tree indexing structure that aids in efficient retrieval. SOM supports dynamic clustering and similarity measures.

IV: Image Retrieval systems in existence

Among the different systems, some commercial and research prototypes such as QBIC [6], VisualSEEK [3, 9], Virage [3], MARS [15] that encompass the above said features are discussed below.

(a) QBIC (Query by Image Content)

QBIC is a content based image retrieval system by IBM [6]. The queries used in the system include example sketches, images, drawings and color patterns with textures. It uses RGB color features. The texture retrieval features take into consideration, characteristics such as coarseness, contrast and directionality. Image segmentation is employed to identify interesting objects in the images and retrieve them. QBIC incorporates high dimensional feature indexing and R* Trees

is the dynamic organization data structure being used. Another feature is that it integrates both text based keyword search and content based similarity search in its interface.

(b) VisualSEEK

VisualSEEK is a content based image query system integrating feature based image indexing and spatial query methods [3]. It employs color sets which is useful in automated region extraction through color set-back projection. Another advantage of using color sets is it is easily indexed and retrieval costs are low. VisualSEEK improves image retrieval efficiency by automating the extraction of localized regions and its features and by supporting feature extraction of compressed data. VisualSEEK tools can be ported to an application aiding in the search of images that is specific to the application [9]. The retrieval process of the images is made fast by use of binary tree based indexing algorithms such as R Trees and M Trees.

(c) VIRAGE

Virage developed by Virage Inc. is a content based image search engine supporting visual queries [3]. It uses features like color, composition, layout, boundary information and texture. The difference between Virage and other systems lie in its ability to provide weights and to lets users adjust the weights based on their preference. It supports both general features such as color, shape, and textures and domain specific features like face recognition and cell detection.

(d) MARS

MARS stands for Multimedia Analysis and Retrieval System. It focuses on organizing the various visual features of the image rather than concentrating one particular visual feature such as color, shape, or texture by seamlessly integrating DBMS and ranked retrieval technologies. It follows an integrated Relevance Feedback Architecture [15].

The system achieves relevance feedback at feature representation and similarity measure levels. The user queries are refined based on the relevance feedback provided by the user.

MARS provides options to retrieve images based on a specific feature. For example, the user can retrieve images that follow a color space (HSV, YIQ, and

LUV) or a particular color feature (color moments, color histogram), texture features or image scales.

V: Trends and Improvements

a. *Linguistic indexing*

Associating pictures with textual descriptions can be further developed using automatic image annotation or linguistic indexing. Images that lack reliable metadata may require linguistic indexing. Content based Image retrieval focuses more on the features and fails when user needs an image related to a particular content. Hence, automatic annotations facilitates content based image search using text.

b. *Stories, Aesthetics in pictures*

Certain concepts described in an article or stories are mapped with relevant pictures. An average image retrieval system fails images are not stored in proper indexing schemes. Hence, the representative pictures needs to be indexed based on a ranking scheme that discriminates the relevant pictures for a given text.

In aesthetics, image retrieval system developers need to consider the most desirable image to a query. This aspect of an image is calculated by measuring its quality. Quality is perceived at levels such as size, aspect ratio and color depth.

c. *Human Interaction*

Interaction and feedback in image retrieval has caused major changes to the way the image retrieval systems have been developed. Techniques such as user interactive segmentation and dynamic feature vector re-computation are based on relevance feedback provided by the users.

d. *Web Oriented Engines*

With the advent of the World Wide Web, the ability to store images has increased manifold. The current image search engines are predominantly text based and subject browsing based. The major concerns regarding web oriented content based engines are due to indexing features and the concern of interoperability.

10) Resources

a) Reading list for Students:

Yong Rui, Thomas S. Huang, Shih-fu Chang (1997), Image Retrieval: Past, Present, And Future, *Journal of Visual Communication and Image Representation*, Page 12-17, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.40.5149>

Abby. A. Goodrum (2000), Image Information Retrieval: An Overview of Current Research, *Informing Science - Special Issue on Information Science Research*, Vol. 3 No 2, Page 1-5, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.26.8589>

Yong Rui, Thomas S. Huang, Shih-fu Chang (1999), Image Retrieval: Current Techniques, Promising Directions and Open Issues, *Journal of Visual Communication and Image Representation*, Pages 39-62, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.32.7819>

b) Reading list for Instructors:

Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang (2006), Image Retrieval: Ideas, Influences, and Trends of the New Age, *ACM Computing Survey*, Volume 39, Pages 1-60, <http://citeseerx.ist.psu.edu/viewdoc/versions?doi=10.1.1.91.7203>

Anil K. Jain, Aditya Vailaya (1996), Image Retrieval using Color and Shape, *Patter Recognition*, Volume 29, Issue 8, Pages 1233-1244, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.47.9699>

M. Flickner, H.Sawhney, W.Niblack, J. Ashley, Q. Huang, D. Steele (1995), Query by Image and Video Content: The QBIC System, *Computer*, Vol. 28, No. 9, Pages 23-32, <http://doi.ieeecomputersociety.org/10.1109/2.410146>

J.Chen, T. Pappas, B. Rogowitz (2002), Adaptive Image Segmentation Based On Color and Texture (2002), *International Conference on Image Processing (ICIP)* Volume 2, Pages 789-792, <http://citeseerx.ist.psu.edu/viewdoc/versions?doi=10.1.1.14.2102>

S.B.Chad, C. Carson, H.Greenspan, J.Malik (1998), Color and Texture-Based Image Segmentation Using EM and Its Application to Content-Based Image Retrieval, *Proceedings of the Sixth International Conference on Computer Vision*, Pages 675-682, <http://portal.acm.org/citation.cfm?id=939161>

VisualSEEK Image retrieval system, <http://www.ee.columbia.edu/ln/dvmm/researchProjects/MultimediaIndexing/VisualSEEK/VisualSEEk.htm>

Grosky, W.I. (n.d.). Image retrieval - existing techniques, content-based (cbir) systems. Retrieved from <http://encyclopedia.jrank.org/articles/pages/6763/Image-Retrieval.html>

Zhuge, H. (n.d.). Semantic-based web image retrieval. Retrieved from <http://www2003.org/cdrom/papers/poster/p172/p172-zhuge/p172-zhuge.htm>

A Content Based Image Search Engine, <http://www.tiltomo.com/>

Anaktisi – Content based Image Search, <http://orpheus.ee.duth.gr/anaktisi/>

Celentano, A., & Sabbadin, S. (n.d.). *Multiple features indexing in image retrieval systems*. Retrieved from http://www.dsi.unive.it/~auce/docs/celentano_cbmi99.pdf

MARS Image Retrieval System, <http://www.ifp.illinois.edu/~qitian/MARS.html>

12) Exercise/Learning Activities

In-Class exercise 1 (30 minutes)

- Split the class into groups of 4 members each.
- Have them use the following web based image retrieval search engines and make them perform the following activities.

The web sites are

- 1) <http://www.tiltomo.com/> “Tiltomo” - A Content Based Image Search Engine
- 2) <http://orpheus.ee.duth.gr/anaktisi/> “Anaktisi – Content based Image Search”

- Provide textual/visual queries to each of the tools.
- Identify which content based image retrieval scheme is being used (mentioned in the curriculum)
- analyze the advantages and disadvantages of each scheme and document the results.

In-Class exercise 2 (15 minutes)

- Use Google ‘Similar Images’ feature at <http://images.google.com> and identify how accurate is the similarity between the retrieved results.
- Provide three examples of queries that return at least one non-similar result for an image.
- Each of the features, color, shape and texture, should be represented in the examples.
- Analyze the dissimilarity of the retrieved result and explain what could be the reason for the behavior.

In-Class Exercise 3 (15 minutes)

- The below link contains demo to a sample R-Tree structure
 - o <http://donar.umiacs.umd.edu/quadtree/points/rtrees.html>
- Using the demo, study how insertion and deletion of nodes work.
- Document the steps and submit.

- The idea of this learning activity is to help students understand the intricacies of these structures.

13) Evaluation of Learning Outcomes

- A) Do students know the fundamental concepts, and techniques of Image Retrieval?
- B) Do students effectively distinguish the features of various content based image retrieval techniques?

14) Glossary

- **Collections:** A group of items, often documents.
- **Feature:** Information extracted from an object and used during query processing
- **Index:** A data structure built for the images to speed up searching.

- **Information Retrieval:** Part of computer science which studies retrieval of information (not data) from a collection of written documents.
- **Metadata:** Attributes of a data or a document
- **Query:** The expression of the user information need in the input language provided by the information system
- **Repository:** A physical or digital place where objects are stored for a period of time, from which individual objects can be obtained if they are requested.
- **Relevance feedback:** An interactive process of obtaining information from the user about the relevance and the non-relevance of retrieved documents.
- **Feature vector:** An n-dimensional vector of numerical features that represent some object.
- **Feature space:** The vector space associated with feature vectors.
- **Color Histogram:** It is the representation of distribution of colors in an image.
- **Euclidean Distance:** It is the ordinary distance between two points in a vector space.

15) Additional Useful Links:

16) Contributors:

Authors:

Nagarajan Kuppuswami

Dr. Edward Fox

Reviewers:

Seungwon Yang

Tarek Kanan

Venkatasubramaniam Ganesan

John Ewers