A STUDY OF MULTIMEDIA INFORMATION RETRIEVAL SYSTEMS

D. Saravanan  
Assistant Professors  
Sathyabama University

V. SomaSundaram  
Assistant Professors  
Sathyabama University

N. Nalini  
Assistant Professors  
Vels University

K. Arun  
Assistant Professors  

ABSTRACT

Today, many of our business and cultural information are being recorded in electronic media and stored in multiple electronic databases. In order to make these information available to an information seeker, there must be an electronic information retrieval system, that facilitates location and retrieval of documents, that are relevant to the information seeker's questions. Since information can be recorded on various media types, such as tables, images, text, audio, and video. The retrieval system must be able to retrieve information from various media representations, giving rise to the concept of Multimedia Information Retrieval Systems - MIRS. In this paper we try to focus different techniques to retrieve multimedia data.

Key Words: Image Retrieval, Multimedia data base, Image mining, Information Retrieval, Data , Databases.
1. INTRODUCTION

1.1 Data Collections and Databases

Related data are commonly collected and stored in files or archives, administered by public or private organizations, such as libraries, museums, statistical offices, government agencies, business organizations, as well as in private collections. Current storage media include books, file cabinets, CDs and electronic (computer) databases. Among the early terms used to describe electronic databases were: files, data registers, data archives, and data banks. Many definitions highlight some important characteristic of the DB and prefix the term database with an attribute that emphasizes that characteristic. Examples include the:

- Underlying data model: hierarchic, network, relational, object-oriented,
- Primary data-type: tabular, text, image, map/spatial, audio, video,
- Content: document, record, multimedia, bibliographic, statistical, geographic, ...
- Primary DB usage: administrative, library, geographic, museum, ...,
- Architecture: centralized, distributed, homogeneous, heterogeneous, or

1.2 Characteristics of Text Documents

Text documents can be considered semi-structured in the sense that they are constructed from a defined term vocabulary according to known grammatical rules for forming sentences and paragraphs. As documents, they also share some common elements such as an author, title, date of origin, a (set of) title(s), and perhaps a recipient. Text document can be described from 3 perspectives:

1. Semantic content of the document, i.e. representation of its meaning.
2. Context of the document, e.g. its author, publisher
3. Structure of the document, e.g. its language, style, length.

1.3 Media Databases:

Media databases are databases that contain the representations of a collection of media objects, such as the articles published in a Web-based newspaper or the scanned images from a museum collection. Multimedia applications frequently contain several media databases containing text, image and/or audio/visual data. For example, a video database may consist of an image stream/DB and 1 or more audio stream/DBs, which must be coordinated in presentation to give proper meaning for the viewer. Advances in computer capacity and the functionality of data management systems have made possible storage of large objects - LOBs, suitable for storage of media data. With these systems, it is possible to create multimedia databases containing representations of related media objects.

2. Image Mining:

Image mining deals with the extraction of image patterns from a large collection of images. Clearly, image mining is different from low-level computer vision and image processing techniques because the focus of image mining is in extraction of patterns from large collection of images, whereas the focus of computer vision and image processing techniques is in understanding and/or extracting specific features from a single image. While there seems to be some overlaps between image mining and content-based retrieval (both are dealing with large collection of images), image mining goes
beyond the problem of retrieving relevant images. In image mining, the goal is the discovery of image patterns that are significant in a given collection of images. Perhaps, the most common misconception of image mining is that image mining is nothing more than just applying existing data mining algorithms on images.

2.1 Characteristics of Image Documents

Image documents are represented as unstructured data, actually by a bit or pixel (picture element) string. As opposed to text documents, they do not have a standard 'vocabulary' or grammar that can be used for automatic interpretation of the semantic content, or meaning of the image. Instead, most image management systems use manual annotations, such as title/caption, keywords and/or text descriptions, to capture semantic interpretation. In addition to text annotations, the semantic content of images can be described by content descriptors, frequently called features. Image features can be classified according to their level of abstraction, or distance from the actual physical content of the image.

Fig 1. Image Data Mining Steps

3. Multimedia Data Base:

3.1 Contents of MMDB

An MMDB needs to manage several different types of information pertaining to the actual multimedia data. They are:

3.1.1 Media data - This is the actual data representing images, audio, video that are captured, digitized, processes, compressed and stored.

3.1.2 Media format data - This contains information pertaining to the format of the media data after it goes through the acquisition, processing, and encoding phases. For instance, this consists of information such as the sampling rate, resolution, frame rate, encoding scheme etc.
3.1.3 Media keyword data - This contains the keyword descriptions, usually relating to the generation of the media data. For example, for a video, this might include the date, time, and place of recording, the person who recorded, the scene that is recorded, etc. This is also called as content descriptive data.

3.1.4 Media feature data - This contains the features derived from the media data. A feature characterizes the media contents. For example, this could contain information about the distribution of colors, the kinds of textures and the different shapes present in an image. This is also referred to as content dependent data.

4. IR Evaluation
The evaluation of information retrieval system is the process of assessing how well a system meets the information needs of its users. Traditional evaluation metrics, designed for Boolean retrieval or top-k retrieval, include precision and recall.

Precision is the fraction of retrieved documents that are relevant to the query:

\[
\text{precision} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{retrieved documents}\}|}
\]

Recall is the fraction of the documents relevant to the query that are successfully retrieved:

\[
\text{recall} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{relevant documents}\}|}
\]

For modern (Web-scale) information retrieval, recall is no longer a meaningful metric, as many queries have thousands of relevant documents, and few users will be interested in reading all of them.

4.1 Multimedia Information Retrieval (MMIR)
MMIS is extracting semantic information from multimedia data sources. Data sources include directly perceivable media such as audio, image, and video, indirectly perceivable sources such as text, biosignals as well as not perceivable sources such as bioinformation, stock prices, etc. The methodology of MMIR can be organized in three groups:

1. Methods for the summarization of media content (feature extraction). The result of feature extraction is a description.
2. Methods for the filtering of media descriptions (for example, elimination of redundancy)
3. Methods for the categorization of media descriptions into classes.

4.1.1 Feature Extraction Methods
Feature extraction is motivated by the sheer size of multimedia objects as well as their redundancy and, possibly, noisiness. Generally, two possible goals can be achieved by feature extraction: summarization of media content. Methods for summarization include in the audio domain, for example, Mel Frequency Cepstral Coefficients.

4.1.2 Merging and Filtering Methods
Merging can be performed by simple concatenation if the descriptions are of fixed size. Variable-sized descriptions - as they frequently occur in motion description - have to be normalized to a fixed length first. Frequently used methods for description filtering include factor analysis (e.g. by PCA), singular value decomposition (e.g. as latent semantic indexing in text retrieval)
4.1.3 Categorization Methods

Generally, all forms of machine learning can be employed for the categorization of multimedia descriptions though some methods are more frequently used in one area than another. For example, Hidden Markov models are state-of-the-art in speech recognition. The list of applicable classifiers includes the following:

1. Metric approaches (Cluster Analysis, Vector Space Model, Minkowski Distances, Dynamic Alignment)
3. Risk Minimization (Support Vector Regression, Support Vector Machine, Linear Discriminant Analysis)
4. Density-based Methods (Bayes Nets, Markov Processes, Mixture Models)
5. Neural Networks (Perceptron, Associative Memories, Spiking Nets)
6. Heuristics (Decision Trees, Random Forests, etc.) The selection of the best classifier for a given problem (test set with descriptions and class)

5. Conclusion:

In general, the information retrieval (IR) refers to the retrieval of unstructured data relevance is like a tree of knowledge. IR tools have been designed for IR professionals to enable them effectively and efficiently retrieve information from a source. It is assumed that the information exists in the source and that a well-formed query will retrieve it. Information seeking may be understood as a more human-oriented and open-ended process than information retrieval. Here we projected some of the relevance retrieval methods for text, image, and multimedia data types. In future more general type of data retrieval structured adopted for particular event.

7. References:


***