The decreasing costs of consumer electronic devices such as digital cameras and digital camcorders, along with the ease of transportation facilitated by the Internet, has lead to a phenomenal rise in the amount of multimedia data. Now multimedia data comprising of images, audio, and video is becoming increasingly common. Given that this trend of increased use of multimedia data is likely to accelerate, there is an urgent need for providing a clear means of capturing, storing, indexing, retrieving, analyzing, and summarizing such data.

Image data, for example, is a very commonly used multimedia data. The early image retrieval systems are based on manually annotated descriptions, called text-based image retrieval (TBIR). TBIR is a great leap forward, but has several inherent drawbacks. First, textual description is not capable of capturing the visual contents of an image accurately, and in many circumstances, textual annotations are not available. Second, different people may describe the content of an image in different ways, which limits the recall performance of textual-based image retrieval systems. Third, for some images there is something that no words can convey. To resolve these problems, content-based image retrieval (CBIR) systems are designed to support image retrieval, and have been used since the early 1990s. Also, some novel approaches (e.g., relevance feedback, semantic understanding, semantic annotation, and semantic retrieval of images) have been developed in the last decade to improve image retrieval and satisfy the advanced requirements of image retrieval.

Multimedia data retrieval is closely related to multimedia data management. Multimedia data management facilitates the manipulation of multimedia data such as representation, storage, index, retrieval, maintenance, and so on. Multimedia data retrieval is the key to implementing multimedia data management on one hand. On the other hand, multimedia data retrieval should be carried out based on multimedia data representation, storage, and index, which are the major tasks of multimedia data management. Databases are designed to support the data storage, processing, and retrieval activities related to data management, and database management systems can provide efficient task support and tremendous gain in productivity is hereby accomplished using these technologies. There is no doubt that database systems play an important role in multimedia data management, and multimedia data management requires database technique support. Multimedia databases, which have become the repositories of large volumes of multimedia data, are emerging.

Multimedia databases play a crucial role in multimedia data management, which provide the mechanisms for storing and retrieving multimedia data efficiently and naturally. Being a special kind of databases, multimedia databases have been developed and used in many application fields. Many researchers have been concentrating on multimedia data management using multimedia databases. The research and development of multimedia data management using multimedia databases are receiving increasing attention. By means of multimedia databases, large volumes of multimedia data can be stored...
and indexed and then retrieved effectively and naturally from multimedia databases. Intelligent multimedia data retrieval systems are built based on multimedia databases to support various problem solving and decision making. Thus, intelligent multimedia databases and information retrieval is a field that must be investigated by academic researchers together with developers both from CBIR and AI fields.

The book has two focuses on multimedia data retrieval and multimedia databases, aiming at providing a single account of technologies and practices in multimedia data management. The objective of the book is to provide the state of the art information to academics, researchers, and industry practitioners who are involved or interested in the study, use, design, and development of advanced and emerging multimedia data retrieval and management with ultimate aim to empower individuals and organizations in building competencies for exploiting the opportunities of the knowledge society. This book presents the latest research and application results in multimedia data retrieval and management. The different chapters in the book have been contributed by different authors and provide possible solutions for the different types of technological problems concerning multimedia data retrieval and management.

INTRODUCTION

This book, which consists of fifteen chapters, is organized into two major sections. The first section discusses the feature and semantics of multimedia data as well as their usage in multimedia information retrieval in the first eight chapters. The next seven chapters covering database and intelligence technologies for multimedia data management comprise the second section.

First of all, we take a look at the issues of the feature and semantics of multimedia data as well as their usage in multimedia information retrieval.

Imad EL-Zakhem et al. concentrate on building a user profile according to his own perception of colors for image retrieving. They develop a dynamic construction of the user profile, which will increase their satisfaction by being more personalized and accommodated to their particular needs. They suggest two methods to define the perception and transform it into a profile: the first one is achieved by querying the user and getting answers and the second one is achieved by comparing different subjects and ending up by an appropriate aggregation. They also present a method recalculating the amount of colors in the image based on another set of parameters, and the colorimetric profile of the image is being modified accordingly. Avoiding the repetition of the process at the pixel level is the main target of this phase, because reprocessing each image is time consuming and not feasible.

In content-based image retrieval, different kinds of features (e.g., texture features, color features and shape features) may be used jointly, and feature integration is hereby one of crucial issues in content-based image retrieval. Gang Zhang et al. develop an approach of integrating shape and texture features and investigate if integration features are more discriminative than single features. Single feature extraction and description is foundation of the feature integration. They apply Gabor wavelet transform with minimum information redundancy to extract texture features, which are used for feature analyses. Fourier descriptor approach with brightness is used to extract shape features. Then both features are integrated together by weights. They make the comparisons among the integration features, the texture features, and the shape features so that the discrimination of the integration features can be testified.

The research domain of automatic image annotation and search from low-level descriptors analysis has considerably evolved in the last 15 years. Since then, this domain has reached a level of maturity where only small improvements are brought in new models and systems. Jean Martinet and Ismail Elsayad propose a classification of image descriptors, from low-level descriptors to high-level
descriptors, introducing the notion of mid-level descriptors for image representation. A mid-level descriptor is described as an intermediate representation between low-level descriptors (derived from the signal) and high-level descriptors (conveying semantics associated to the image). Mid-level descriptors are built for the purpose of yielding a finer representation for a particular set of documents. They describe a number of image representation techniques from a mid-level description perspective.

There are hundreds of millions of images available on the current World Wide Web, and the demand for image retrieval and browsing online is growing dramatically. The typical keyword-based retrieval methods for multimedia documents assume that the user has an exact goal in mind in searching a set of images whereas users normally do not know what they want, or the user faces a repository of images whose domain is less known and content is semantically complicated. In these cases it is difficult to decide what keywords to use for the query. Lisa Fan and Botang Li present an approach of the user-driven ontology guided image retrieval. It combines (a) the certain reasoning techniques based on logic inside ontology and (b) the uncertain reasoning technique based on Bayesian Network to provide users the enhanced image retrieval on the Web. Their approach is for easily plugging in an external ontology in the distributed environment and assists user searching for a set of images effectively. In addition, to obtain a faster real-time search result, the ontology query and BN computation should be run on the off-line mode, and the results should be stored into the indexing record.

A large number of digital medical images have been produced in hospitals in the last decade. These medical images are stored in large-scale image databases and can facilitate medical doctors, professionals, researchers, and college students to diagnose current patients and provide valuable information for their studies. Image annotation is considered as a vital task for searching, and indexing large collections of medical images. Chia-Hung Wei and Sherry Y Chen present a complete scheme for automatic annotation on mammograms. Firstly, they present the feature extraction methods based on BI-RADS standards. This ensures that the meaning and interpretation of mammograms are clearly characterized and can be reliably used for feature extraction. Secondly, they propose the SVM classification approach to image annotation. Finally, their experimental results demonstrate that the scheme can achieve fair performance on image annotation.

Digital image storage and retrieval is gaining more popularity due to the rapidly advancing technology and the large number of vital applications, in addition to flexibility in managing personal collections of images. Traditional approaches employ keyword based indexing which is not very effective. Content based methods are more attractive though challenging and require considerable effort for automated feature extraction. Görkem Aşıloğlu et al. present a hybrid method for extracting features from images using a combination of already established methods, allowing them to be compared to a given input image as seen in other query-by-example methods. First, the image features are calculated using edge orientation autocorrelograms and color correlograms. Then, distances of the images to the original image are calculated using the L1 distance feature separately for both features. The distance sets are then be merged according to a weight supplied by the user.

Disadvantages with text-based image retrieval have provoked growing interest in the development of Content-Based Image Retrieval (CBIR). In CBIR, instead of being manually annotated by text-based keywords, images are indexed by their visual content, such as color, texture, etc. Ling Shao surveys content-based image retrieval techniques on representing and extracting visual features, such as color, shape, and texture. The feature representation and extraction approaches are first classified and discussed. Then, he summarizes several classical CBIR systems which rely on either global features or features detected on segmented regions. The inefficiency and disadvantages of those narrow-domain systems
are also presented. Finally, he discusses two recent trends on image retrieval, namely semantic based methods and local invariant regions based methods, and proposes directions for future work.

With the rapid growth of digital videos, efficient tools are essential to facilitate content indexing, searching, retrieving, browsing, skimming, and summarization. Sport video analysis aims to identify what excites audiences. Previous methods rely mainly on video decomposition, using domain specific knowledge and lacking the ability to produce personalized semantics especially in highlight detection. Research on suitable and efficient techniques for sport video analysis has been conducted extensively over the last decade. Chia-Hung Yeh et al. review the development of sport video analysis and explore solutions to the challenge of extracting high-level semantics in sport videos. They propose a method to analyze baseball videos via the concept of gap length. Use-interaction may be a solution to achieve personalization in semantics extraction. The techniques introduced can be wildly applied to many fields, such as indexing, searching, retrieving, summarization, skimming, training, and entertainment.

The second section deals with the issues of database and intelligence technologies for multimedia data management.

The last decades have witnessed a considerable rise in the amount of multimedia data. Data models and database management systems (DBMSs) can play a crucial role in the storage and management of multimedia data. Being a special kind of database systems, multimedia databases have been developed and used in many application fields. Shi Kuo Chang, Vincenzo Deufemia, and Giuseppe Polese present normal forms for the design of multimedia database schemes with reduced manipulation anomalies. They first discuss how to describe the semantics of multimedia attributes based upon the concept of generalized icons, already used in the modeling of multimedia languages. They then introduce new extended dependencies involving different types of multimedia data. Based upon these new dependencies, they define five normal forms for multimedia databases, some focusing on the level of segmentation of multimedia attributes, others on the level of fragmentation of tables. Thus a normalization framework for multimedia databases is developed, which provides proper design guidelines to improve the quality of multimedia database schemes.

Multimedia data is a challenge for data management. The semantics of traditional alphanumeric data are mostly explicit, unique, and self-contained, but the semantics of multimedia data are usually dynamic, diversiform, and varying from one user’s perspective to another’s. Dawen Jia and Mengchi Liu introduce a new model, titled the Information Networking Model (INM). It provides a strong semantic modeling mechanism that allows modeling of the real world in a natural and direct way. With INM, users can model multimedia data, which consists of dynamic semantics. The context-dependency and media-independency features of multimedia data can easily be represented by INM. In addition, multimedia multiple classifications are naturally supported. Based on INM, they propose a multimedia data modeling mechanism in which users can take advantage of basic multimedia metadata, semantic relationships, and contextual semantic information to search multimedia data.

With increasing use of multimedia in various domains, several metadata standards appeared these last decades in order to facilitate the manipulation of multimedia contents. These standards help consumers to search content they desire and to adapt the retrieved content according to consumers’ profiles and preferences. However, in order to extract information from a given standard, user must have a pre-knowledge about this latest. This condition is not easy to satisfy due to the increasing number of available standards. Samir Amir et al. first give an overview about existing multimedia metadata standards and CAM4Home project initiative that covers a wide area of information related to multimedia delivery and includes multimedia content description, user preference and profile description, and devices’ characteristic description. Then they relate about multimedia and generic integration issues
by discussing the work done by W3C working group in order to integrate heterogeneous metadata and some generic approaches providing mapping between ontologies. They also consecrate to the illustration of the proposal of a new architecture for the multimedia metadata integration system and discuss about challenges of its realization.

Semantic characterization is necessary for developing intelligent multimedia databases, because humans tend to search for media content based on their inherent semantics. However, automated inference of semantic concepts derived from media components stored in a database is still a challenge. Ranjan Parekh and Nalin Sharda demonstrate how layered architectures and visual keywords can be used to develop intelligent search systems for multimedia databases. The layered architecture is used to extract meta-data from multimedia components at various layers of abstractions. To access the various abstracted features, a query schema is presented which provides a single point of access while establishing hierarchical pathways between feature-classes. Minimization of the semantic gap is addressed using the concept of visual keyword (VK). Semantic information is however predominantly expressed in textual form, and hence is susceptible to the limitations of textual descriptors—viz. ambiguities related to synonyms, homonyms, hypernyms, and hyponyms. To handle such ambiguities they propose a domain specific ontology-based layer on top of the semantic layer, to increase the effectiveness of the search process.

Fuzzy set theory has been extensively applied to the representation and processing of imprecise and uncertain data. Image data is becoming a kind of important data resources with rapid growth in the number of large-scale image repositories. But image data is fuzzy in nature and imprecision and vagueness may exist in both image descriptions and query specifications. Li Yan and Z. M. Ma review some major work of image retrieval with fuzzy logic in the literature, including fuzzy content-based image retrieval and database support for fuzzy image retrieval. For the fuzzy content-based image retrieval, they present how fuzzy sets are applied for the extraction and representation of visual (colors, shapes, textures) features, similarity measures and indexing, relevance feedback, and retrieval systems. For the fuzzy image database retrieval, they present how fuzzy sets are applied for fuzzy image query processing based on a defined database models, and how various fuzzy database models can support image data management.

Project portfolio management of multimedia production and use emerges today as a challenge both for the enrichment of traditional classroom based teaching and for distance education offering. In this way, Joni A. Amorim, Rosana G. S. Miskulin, and Mauro S. Miskulin intend to answer the following question: “Which are the fundamental aspects to be considered in the management of projects on educational multimedia production and use?” They present a proposal of a project management model for digital content production and use. The model, the methodology and the implementation are named EduPMO (Educational Project Management Office). The model, the methodology and the implementation should be understood as related but independent entities. This interdisciplinary investigation involves different topics, going from metadata and interoperability to intellectual property and process improvement.

Latent Semantic Analysis (LSA) or Latent Semantic Indexing (LSI), when applied to information retrieval, has been a major analysis approach in text mining. It is an extension of the vector space method in information retrieval, representing documents as numerical vectors, but using a more sophisticated mathematical approach to characterize the essential features of the documents and reduce the number of features in the search space. Anne Kao et al. summarize several major approaches to this dimensionality reduction, each of which has strengths and weaknesses, and describe recent breakthroughs and advances. They show how the constructs and products of LSA applications can be made user-interpretable and review applications of LSA beyond information retrieval, in particular, to text information visualization. While the major application of LSA is for text mining, it is also highly applicable to cross-language information retrieval, Web mining, and analysis of text transcribed from speech and textual information in video.