Preface

Among the new challenges of multimedia computing and distributed technologies, mechanisms for content-based information retrieval in multimedia databases seems to be the most difficult issue. Unlike traditional database systems, which allow query specification based on keywords and numerical comparisons, an image database system requires a sophisticated retrieval mechanism. Current approaches of image retrieval rely on color, texture, shape, and object spatial relations. However, the most difficult issue of image content-based retrieval is the investigation of friendly visual specification techniques. How to visually describe the need of a user is a very difficult problem. On the other hand, content-based retrieval of video records not only involves the objects in a video, but the timing of object movement is also considered. Scene identification and object tracing are basic techniques that only solve part of the problem. Yet, tools for semantic analysis of video contents are still under development. Content-based video retrieval may rely on speech detection and recognition, which are also used in the automatic retrieval of audio information. Part I is an introduction. Part II and III of this book will discuss current research status and solutions to image and video databases.

Part IV of this book discusses watermark technologies. This part of the discussion is relatively new in the area of multimedia computing. With the growing number of Internet users, it is likely that entertainment data, such as MP3 audio records and MPEG video, as well as still images are being increasingly transmitted on the Internet. It is important for vendors to have their copyright protection signatures embedded into these multimedia records. Watermark of multimedia can be embedded into MP3 audio, video, and image records without seriously affecting the presentation quality of the records. Techniques include embedding signature, logo, text and audio information into these entertainment records. A strong watermark technique should defend itself from various types of attacks, as well as prevent the destruction of signatures from data compression. With the watermark embedded, multimedia records are delivered to the end user.

On the other hand, the delivery of multimedia information is time-sensitive. Synchronization solutions to multimedia streams can be divided into two parts: intra-stream synchronization and inter-stream synchronization. For instance, an intra-stream synchronization solution involves the encoding of time stamps or the realization of other techniques to ensure that audio and animation are synchronized in a video clip. The synchronization technique can be extended to include multiple streams (e.g., a video clip is synchronized with a slide show). Moreover, quality of services (QoS) of multimedia data may involve real-time constraints, especially when the service is distributed among workstations linked by networks. Solutions to QoS involve different levels of network protocols, as well as the construction of efficient network infrastructure. These problems will be discussed in Part V.

The discussions in previous chapters contain technical issues, such as specification, design, and algorithms to the research topics of multimedia databases. Part VI includes real applications in image and video databases, VR technique, and media synchronization

solutions. The last part of this book (Part VII) will point out other possible solutions in the future.

The intended audiences of this book are senior or graduate students majoring in computer science, computer engineering, or management information system (MIS), as well as professional instructors, and multimedia product developers. Readers can benefit from this book in searching for state-of-the-art research topics for their research, as well as in the understanding of techniques and applications in content-based information retrieval, watermarking, and distributed multimedia database systems. This book can be used as a textbook in senior research seminars, as well as graduate courses.

Chapter Highlights

Chapter 1: Distributed Multimedia Databases

Distributed Multimedia Database involves network technology, distributed control, security, and multimedia computing. This chapter discusses fundamental concepts and introduces issues of image database and digital libraries, video-on-demand systems, multimedia synchronization, as well as some case studies of distributed multimedia database systems. Requirements of multimedia database management systems and their functions are also presented.

Chapter 2: Bridging the Semantic Gap in Image Retrieval

The emergence of multimedia technology and the rapidly expanding image and video collections on the Internet have attracted significant research efforts in providing tools for effective retrieval and management of visual data. Image retrieval is based on the availability of a representation scheme of image content. Image content descriptors may be visual features such as color, texture, shape, and spatial relationships, or semantic primitives. Conventional information retrieval was based solely on text, and those approaches to textual information retrieval have been transplanted into image retrieval in a variety of ways. However, "a picture is worth a thousand words." Image contents are much more versatile compared with text, and the amount of visual data is already enormous and still expanding very rapidly. Hoping to cope with these special characteristics of visual data, content-based image retrieval methods have been introduced. It has been widely recognized that the family of image retrieval techniques should become an integration of both low-level visual features addressing the more detailed perceptual aspects and high-level semantic features underlying the more general conceptual aspects of visual data. Neither of these two types of features is sufficient to retrieve or manage visual data in an effective or efficient way. Although efforts have been devoted to combining these two aspects of visual data, the gap between them is still a huge barrier in front of researchers. Intuitive and heuristic approaches do not provide us with satisfactory performance. Therefore, there is an urgent need of finding the latent correlation between low-level features and high-level concepts and merging them from a different perspective. How to find this new perspective and bridge the gap between visual features and semantic features has been a major challenge in this research field. This chapter addresses these issues.

Chapter 3: Content-Based Visual Information Retrieval

This chapter provides a survey of the state-of-the-art in the field of Visual Information Retrieval (VIR) systems, particularly Content-Based Visual Information Retrieval (CBVIR)

systems. It presents the main concepts and system design issues, reviews many research prototypes and commercial solutions currently available and points out promising research directions in this area.

Chapter 4: Content-Based Trademark Recognition and Retrieval Based on Discrete Synergetic Neural Network

Synergetic Neural Network (SNN) as proposed by Hermann Haken is a novel top-down self-organized system. In this chapter, its associated discrete SNN is proposed and the recognition stability and the convergence of a generalized discrete SNN is analyzed. We proposed an adaptive algorithm of iterative step length refinement for synergetic recognition, which can ensure fast convergence and network steadily for all kinds of input pattern. Additionally, we apply the SNN to trademark retrieval and study its ability to support affine invariant retrieval of 2D patterns. To this end, we propose an affine invariant input vector in the frequency domain for the SNN to evaluate the retrieval ability of such networks for different types of input queries, for example, query by complete trademark pattern and query by image components. We show experimentally that our proposed SNN method is noise tolerant as well as able to support affine invariant retrieval. This led us to propose a novel paradigm for trademark retrieval based on visual keywords whereby trademark images can be queried in terms of simple geometric components.

Chapter 5: Concepts of Emergence Index in Image Databases

Emergence is a phenomenon where we study the implicit or hidden meaning of an image. We introduce this concept in image database access and retrieval of images using this as an index for retrieval. This would give an entirely different search outcome than ordinary search where emergence is not considered, as consideration of hidden meanings could change the index of search. We discuss emergence, emergence index and approach how to apply this concept in image retrieval in this chapter.

Chapter 6: Distributed Temporal Video DBMS using Vertical Class Partitioning Technique

In this chapter, we describe our work on developing a distributed video DBMS (database management system). The video DBMS provides a temporal modeling framework for describing video data, and it supports data distribution by applying vertical class partitioning techniques. Building on top of our previous work on Four Dimensional Information Space (4DIS) - an object-oriented temporal modeling framework, we apply class partitioning techniques onto a distributed 4DIS video database system as a means for efficient query execution. A detailed cost model for query execution through vertical class partitioning is developed. The effectiveness of our class partitioning approach, in the context of the distributed 4DIS video database system, is demonstrated through the use of a running example.

Chapter 7: Video-on-Demand: Scalability and QoS Control

Video-on-Demand (VoD) systems face scalability problems and Quality of Service (QoS) issues due to the need to satisfy numerous requests for several different videos given the limited bandwidth of the communication links. In order to provide scalable solutions and guarantee given QoS requirements, existing VoD proposals can be roughly divided into two

categories: (a) scheduled multicast and (b) periodic broadcast. In this chapter, we propose (a) a novel scheduled multicast scheme based on a time-dependent bandwidth allocation approach, (b) a Trace-Adaptive Fragmentation (TAF) scheme for periodic broadcast of Variable Bit Rate (VBR) encoded video, and (c) a Loss-Less and Bandwidth-Efficient (LLBE) protocol for periodic broadcast of VBR video. We have designed, simulated and evaluated the proposed schemes, and the simulation results demonstrate the benefits, flexibility and feasibility of the proposals.

Chapter 8: Video Abstraction Techniques for Digital Library

The abstraction of a long video is often useful to a user in determining whether the video is worth viewing or not. In particular, video abstraction provides users of digital libraries with a fast, safe and reliable access of video data. Two approaches, summary sequences and highlights, are possible in video abstraction. The summary sequences are good for documentaries because they give an overview of the contents of the entire video, whereas highlights are good for movie trailers because they contain only the most interesting video segments. The video abstraction can be generated by three steps: analyzing video, selecting video clips, and synthesizing the output. In the analyzing video step, salient features, structures, or patterns in visual information, audio information, and textual information are detected. In the selecting step, meaningful clips are selected from detected features in the previous step. In the output synthesis step, the selected video clips are composed into the final form of the abstract. In this chapter, we will discuss various video abstraction techniques for digital libraries. In addition, we will also discuss a context-based video abstraction method in which contextual information of the video shot is computed. This method is useful in generating highlights because the contextual information of the video shot reflects semantics in video data.

Chapter 9: Video Database Techniques and Video-on-Demand

Generally, a large-scale video server is composed of numerous disk striping groups. The striping policies employed by each disk striping group largely determine the performance of a video server. For storage and transmission efficiency, video data are usually compressed using variable-bit-rate (VBR) encoding algorithms, such as JPEG and MPEG. The amount of data consumed by a VBR video stream varies with time. This property, when coupled with striping, unfortunately, results in load imbalance across disks, degrading the overall server performance significantly. This chapter focuses on VBR video striping. It presents two state-of-the-art VBR striping schemes proposed in the literature: one is designed for homogeneous disks and the other is designed for heterogeneous disks. To gain insights into VBR striping, this chapter also develops performance models for the two striping policies. With these performance models, system designers can predict the maximum service capacity of a server, perform online admission control for clients, and optimize the performance of a server, without performing exhaustive tests on a real system.

Chapter 10: Broadcasting Approaches for VOD Services

A VOD system is typically implemented by a client-server architecture supported by certain transport networks such as telecom, CATV, or satellite networks. The simplest scheme is to dedicate a channel to each client. Many VCR-like functions may be provided (e.g., forward, rewind, pause, search, etc.). Since video is an isochroous medium, the video server has to reserve a sufficient amount of network bandwidth and I/O bandwidth for each video

stream before committing to a client's request. Apparently, such systems may easily run out of channels because the growth of the number of channels can never keep up with the growth of the number of clients. To solve this problem, many schemes had been proposed to resolve the bandwidth problem. In this chapter, we review two kinds of broadcasting schemes. The first one is the batching scheme, in which a set of viewers arriving close in time will be collected and grouped together. Then the video server will serve them altogether with one channel. The second one is the periodic broadcasting approach. The server uses multiple dedicated channels to cooperatively broadcast one video. Each channel is responsible for broadcasting some portion of the video. Each client follows some reception rule to grab data from appropriate channels so as to play the whole video continuously. The server's broadcasting activity is independent of the arrivals of requests. Such an approach is more appropriate for popular or "hot" videos that may interest many viewers during a certain period of time.

Chapter 11: Watermarking on Compressed/Uncompressed Video Using Communications with Side Information Mechanism

Digital watermarking has been proposed as very useful technology in the protection of digital data such as image, audio, video, formatted documents (PDF or PS), and 3D objects. In the literature, most of the existing watermarking approaches are conducted on images. However, video is even more useful and should be protected with higher priority. In particular, video sequences usually contain rich properties that images do not have. On the other hand, the types of attacks applied on a video are much different from those applied on an image. In this chapter, we will focus ourselves specifically on video watermarking. The content of this chapter is divided into two parts. In the first part, the existing video watermarking techniques are briefly reviewed. We have pointed out their advantages and disadvantages to realize what can be done about video watermarking. In the second part, we proposed a compressed domain video watermarking scheme for copyright protection. For the sake of real-time video watermark detection, our method is directly conducted in the MPEG-2 bitstream. More specifically, watermarks are inserted into the VLC domain. We shall discuss how to select proper data in a video bitstream to embed watermarks while preserving perceptual fidelity. In addition, video watermarks are embedded by a new proposed watermarking technique, which is based on the concept of communications with side information. The power of our method is reflected by its robust capability against attacks. Future work will also be pointed out to further improve the current scheme.

Chapter 12: Hiding Images Using Dynamic Bit-Replacement and Human Visual System

A simple image hiding scheme in spatial domain is proposed in this chapter. The main idea is to utilize a threshold mechanism to embed as much information of the secret image into the cover image as possible. The changing of the cover image is difficult to discover by the human eyes because the threshold mechanism is set up especially to fit the human visual system. The experimental results show that the human visual system has improved the quality in terms of perceptibility. On the hiding capacity issue, the proposed method has the capability to embed two times the size of the secret image of previous work. A partial encryption strategy is used for the security of the secret image. In addition, a two-dimensional permutation function, torus automorphism, is also introduced in this chapter.

Chapter 13: Embedding Robust Gray-Level Watermark in an Image Using Discrete Cosine Transformation

Digital watermarking is an effective technique to protect the intellectual property rights of digital images. In general, a gray-level image can provide more perceptual information; moreover, the size of each pixel in the gray-level image is bigger. Commonly, gray-level digital watermarks are more robust. In this chapter, the proposed watermarking scheme adopts a gray-level image as the watermark. In addition, discrete cosine transformation (DCT) technique and quantization method are applied to strengthen the robustness of the watermarking system. Both original image and digital watermark, processed by DCT transformation, can build a quantization table to reduce the information size of the digital watermark. After quantized watermark is embedded into the middle frequency bands of the transformed original image, the quality of the watermarked image is always visually acceptable because of the effectiveness of the quantization technique. The experimental results show that the embedded watermark can resist image cropping, JPEG lossy compression, and destructive processes such as image blurring and sharpening.

Chapter 14: Multicast: Concept, Problems, Routing Protocols, Algorithms and QoS Extensions

The advancement in optical fiber and switching technologies has resulted in new generation high-speed networks that can achieve speeds of up to a few gigabits per second. Also, the progress in audio, video and data storage technologies has given rise to new distributed real-time applications. These applications may involve multimedia, which require low end-to-end delay. The applications' requirements, such as the end-to-end delay, delay jitter, and loss rate, are expressed as QoS parameters that must be guaranteed. In addition, many of these new applications may involve multiple users, and hence the importance of multicast communication. In this chapter we discuss the basics of multicasting, its routing protocols and algorithms, along with different QoS-based multicast routing

Chapter 15: Multimedia and Multi-Stream Synchronization

Synchronization is an important aspect of the design and implementation of distributed multimedia database systems. In this chapter, we first examine the models that have been proposed in the literature to specify multimedia and multi-stream synchronization and the methods to implement synchronization mechanisms in distributed multimedia database systems. Their strengths and limitations have been compared and the issues about time uncertainty and interactivity in multimedia and multi-stream synchronization specification and implementation are discussed. Moreover, we use an example to show how to incorporate a synchronization agent into a distributed multimedia database system. Finally, we discuss the trend of future research and development of this topic.

Chapter 16: A Simple Prediction Method for Progressive Image Transmission

BPM is a simple and intuitive method to implement the progressive image transmission. However, its reconstructed image quality at each of the beginning stages is not good. In this paper, we propose a simple prediction method to improve the quality of the reconstructed image for BPM at each of the beginning stages. By partitioning the input image into smaller blocks, our method transmits an important part of the pixel information of each block to the

receiver in each stage. To reconstruct the whole image, the receiver recovers the missing pixel information in each block by linear prediction based on the transmitted pixel information. The experiment results show that our method can significantly improve the reconstructed image quality at each of the beginning stages compared to the BPM and IBPM proposed previously.

Chapter 17: A Multimedia Database Supports Internet-Based English Learning

This work presents a novel English distance learning system that was developed through multimedia database and Internet technologies called English Multimedia Corpus. The system includes English articles, dialogs, and videos. A student can study English writing and reading as well as view Web browser listings to connect the Corpus server. In the system, semantic query and "Link grammar" are applied to construct the English Multimedia Corpus system. Furthermore, it promotes the query level from keyword-base and content-based query to a semantic level. The main function of this system is to query the English sentence pattern through keywords from the English Multimedia Corpus. The other function is to detect grammatical errors in written English. Thus, the system not only teaches English grammar, but also, due to its database, allows teachers to understand the most frequent mistakes.

Chapter 18: Case Study: Cairo – A Distributed Image Retrieval System for Cluster Architectures

Cairo is a distributed, cluster-based image retrieval system that provides a high-quality, object-based image analysis and search. The state-of-the-art retrieval approach, which compares entire images, is extended by an exhaustive search in all image sections for the occurrence of selected regions of interest. The large computational effort requires the use of parallel architectures in order to maintain reasonable system response times. The goal of this chapter is to give an overview of the different techniques required for distributed image retrieval, such as querying mechanisms and operations for image analysis and comparison, related aspects of parallel processing, scheduling, and data placement. The proposed architecture, Cairo, combines all these aspects and offers one possible design for a user-friendly, flexible, efficient, and distributed image database. The global system structure, the design of the individual modules, the internal communication, and typical use cases are outlined in the main part of this chapter.

Chapter 19: VRML-Based System for a Three-Dimensional Virtual Museum

A Virtual Reality Modeling Language (VRML)-based interactive three-dimensional (3D) museum on the Web is designed and implemented. Users may walk through a 3D representation of the whole museum, viewing its collections and seeing objects in 3D together with information about them. To allow all of these, we have designed a set of interactive dynamic Web pages, using a client-server platform that delivers the information required by the user, both 3D data of the objects and texts and 2D information. All this information is stored in a database to enable easier access and management.

Chapter 20: Organizing Multimedia Objects Using a Class Algebra Database

This chapter considers the question of how to locate appropriate multimedia objects. The traditional use of Web search engines is limited by the fact that there are as of yet no standard XML tags describing the many various types of multimedia objects. The users would much prefer to choose from menus, especially from a classification hierarchy where subclasses are all listed. Traditional SQL relational databases are generally not flexible enough to permit dynamic addition of attribute/value pairs to describe these multimedia objects and to display the resulting schema. The solution suggested in this chapter is the sharing of class and binary relation definitions, including the binary superclass/subclass relation of the IS-A classification hierarchy of object-oriented technology. This object-oriented database implementation is based on class algebra. Similar users can see each other's objects, classes, and relations, and add their own objects into appropriate places in both the class hierarchy and the binary relations. The binary relations look like subdirectories, and the available binary relations and methods can be chosen from a menu. The distributed database system, therefore, looks like a typical network operating system, and it is expected that average users may be able to make database queries and updates without any need of programmers.

Chapter 21: Use of the Frame Synchronization Technique to Improve the Visualization Realism on the Web

Along with the popularity of the World Wide Web, merging the virtual reality technique into the WWW environment is a popular area of research in recent years. The emerging of the VRML standard is the result of such effort. The panoramic image viewer is another achievement of this type of research. However, these two types of browsers have their own merits and faults, and these differences are revealed in the discrepancy of their respective user interfaces. This chapter presents a frame synchronization technique so that the display on the VRML browser can harmonize with the picture of the panoramic image viewer. This chapter first presents a full in depth comparison of these two browsers, and then draws out the issues that are essential to synchronize the frame displays on these two browsers. The equations and mechanisms to enable such synchronization then follow. Finally, the frame synchronization mechanism that was implemented with experiments to demonstrate its effectiveness is also given. The frame synchronization mechanism provides a simple yet effective method to increase visualization realism inside the virtual world without sacrificing the freedom of navigation.

Chapter 22: Future Multimedia Databases and Research Directions

Distributed Multimedia Databases will be a trend for future information repository. With the growing popularity of the Internet and the improvement of its bandwidth, the accessibility, diversity, mobility, and scalability of multimedia information retrieval will accelerate information exchange in our global society. Previous chapters presented in this book cover most of the recent issues in distributed multimedia systems. This final chapter points out some research issues and challenges of the underlying new technologies.