Preface

Multimedia systems and content-based image retrieval are very important areas of research in computer technology. Plenty of research works are being done in these fields at present.

These two areas are changing our life-styles because they together cover creation, maintenance, accessing and retrieval of video, audio, image, textual and graphic data. But still lots of important issues in these areas remain unresolved and further research works are needed to be done for better techniques and applications.

The primary objective of this book is to combine these two areas of research together as one is related to the other and provide an up-to-date account of the works being done. We addressed research issues in these fields where some progresses have been made already. Also we encouraged researchers, academics and industrial technologists to provide new and brilliant ideas on these fields which could be pursued for further research.

Content-based image retrieval (CBIR) is a bottleneck of the multimedia systems. Although attempts have been made to perform CBIR on an efficient basis based on shape, color, texture and spatial relations, it has yet to attain maturity. A major problem in this area is computer perception. It is very difficult to make a computer think like a human being to sense what the image or object within the image is. In other words, there remains a big gap between low-level features like shape, color, texture and spatial relations and high-level features like table, chair, car, etc. This means to find the semantic meanings out of an image are still very difficult. Only to some extent, success has been achieved and only in certain categories of images. No universally accepted software package is available to find semantic or high-level meanings out of an object. Another issue in this area of research is image segmentation. This is also a problem of computer perception. Plenty of research is being done to have meaningful and thorough image segmentation based on color, texture, and shape. Only limited success is achieved in certain specific types of image while no universally accepted software package is available for any kind of image. We addressed these two issues in Section III

and IV. Issues related to video shot transitions and video content-based retrieval techniques are also covered in Section III and IV.

Section I of the book is the introduction. We have given a general introduction of the two areas, namely, multimedia systems and content-based image retrieval, from the very elementary level. We discussed the problems in these areas and some of the works done in the field of content-based image retrieval since the beginning of last decade of the last century.

Section II covers research works done in multimedia structures and security. Digital watermarking techniques are applied for content protection of multimedia data. It is a relatively new field of research. New solutions are provided for music sheets to overcome the drawbacks of raster-oriented watermarking algorithms for images, which is applied for music sheets. Also, this section covers face recognition technology with an emphasis on facial scan or face recognition and a Chinese duplicate image retrieval system.

In Section V, we present the research works to provide home users with effective and efficient tools to organize and access large digital images and video albums. We introduce new concepts of emergence index in image retrieval where we consider hidden or implicit meanings of an image in addition to explicit meaning for indexing purposes.

Section VI deals with various content-based image retrieval techniques and Section VII talks about user interfaces. The area of user interfaces is also a very important area of research in multimedia systems as it is the medium between the system and the end user. Successful content-based image retrieval is not possible without proper user interface through which the search process begins. We provided a state-of-the-art review of research on user interface including researches on face animation.

The audience for this book would be researchers who are working in these two fields. Also, researchers from other areas who could start-up in these fields could find the book useful. It could be a reference guide for researchers from other related areas. Reading this book can also benefit undergraduate and post-graduate students, who are interested in multimedia technology.

CHAPTER HIGHLIGHTS

Chapter 1: Multimedia Systems and Content-Based Image Retrieval

In this chapter, we present a basic introduction of the two very important areas of research in the domain of information technology, namely, multimedia systems and content-based image retrieval. The latter is still a widely unresolved problem. We discuss some of the works done so far in content-based image retrieval in the context of multimedia systems.

Chapter 2: A Duplicate Chinese Document Image Retrieval System Based on Line Segment Feature in Character Image Block

With the rapid progress of digital image technology, the management of duplicate document images is also emphasized widely. As a result, this chapter suggests a duplicate Chinese document image retrieval (DCDIR) system, which uses the ratio of the number of black pixels to that of white pixels on the scanned line segments in a character image block as the feature of the character image block. Experimental results

indicate that the system can indeed effectively and quickly retrieve the desired duplicate Chinese document image from a database.

Chapter 3: Technology of Music Score Watermarking

Content protection for multimedia data is widely recognized especially for data types that are frequently distributed, sold or shared using the Internet. Particularly, the music industry, in dealing with audio files, realized the necessity for content protection. Distribution of music sheets will face the same problems. Digital watermarking techniques provide a certain level of protection for these music sheets. But classical raster-oriented watermarking algorithms for images suffer several drawbacks when directly applied to image representations of music sheets. Therefore, new solutions have been developed which are designed regarding the content of the music sheets. In comparison to other media types, the development for watermarking of music scores is a rather young art. The chapter reviews the evolution of the early approaches and describes the current state of the art in the field.

Chapter 4: Face Recognition Technology: A Biometric Solution to Security Problems

Face recognition technology is one of the most widely used applications in computer vision. It is widely used in the applications related to security and human-computer interfaces. The two reasons for this are the wide range of commercial and law enforcement applications and the availability of feasible technologies. In this chapter, the various biometric systems and the commonly used techniques of face recognition, Feature Based, Eigen Face Based, Line Based Approach and Local Feature Analysis are explained along with the results. A performance comparison of these algorithms is also given.

Chapter 5: Histogram Generation from the HSV Color Space Using Saturation Projection

In this chapter, we make an in-depth analysis of the visual properties of the HSV (Hue, Saturation, Intensity Value) color space and describe a novel histogram generation technique for color feature extraction. In this new approach, we extract a pixel feature by choosing relative weights of hue and intensity based on the saturation value of the pixel. The histogram retains a perceptually smooth color transition between its adjacent components that enables us to do a window-based smoothing of feature vectors for the purpose of effective retrieval of similar images from very large databases. The results have been compared with a standard histogram generated from the RGB color space and also with a histogram similar to that used in the QBIC system (Niblack et al., 1993).

Chapter 6: Video Content-Based Retrieval Techniques

The increasing use of multimedia streams nowadays necessitates the development of efficient and effective methodologies and systems for manipulating databases storing these streams. These systems have various areas of application such as video-on-demand and digital libraries. The importance of video content-based retrieval (CBR) systems motivates us to explain their basic components in this chapter and shed light on their underlying working principles. In general, a content-based retrieval system of video data consists of the following four stages:

- Video shot boundary detection
- Key Frames (KFs) selection
- Features extraction (from selected KFs)
- Retrieval stage (where similarity matching operations are performed)

Each one of these stages will be reviewed and expounded based on our experience on building a Video Content-based Retrieval (VCR) that has been fully implemented from scratch in JAVA Language (2002). Moreover, current research directions and outstanding problems will be discussed for each stage in the context of our VCR system.

Chapter 7: Object-Based Techniques for Image Retrieval

To overcome the drawback of using only low level features for the description of image content and to fill the gap between the perceptual property and semantic meaning, this chapter presents an object-based scheme and some object level techniques for image retrieval. According to a multi-layer description model, images are analyzed in different levels for progressive understanding, and this procedure helps to gain comprehensive representations of the objects in images. The main propulsion of the chapter includes a multi-layer description model that describes the image content with a hierarchical structure, an efficient region-based scheme for meaningful information extraction; a combined feature set to represent the image at a visual perception level; an iterative training-and-testing procedure for object region recognition, a decision function for reflecting common contents in object description and a combined feature and object matching process, as well as a self-adaptive relevance feedback that could work with or without memory. With the proposed techniques, a prototype retrieval system has been implemented. Real retrieval experiments have been conducted, results show that the object-based scheme is quite efficient and the performance of object level techniques have been confirmed.

Chapter 8: Object-Based Video Analysis and Interpretation

In this chapter, we present a novel scheme for object-based video analysis and interpretation based on automatic video object extraction, video object abstraction, and semantic event modeling. In this scheme, video objects (VOs) are first automatically extracted, followed by a video object abstraction algorithm for identifying key frames to reduce data redundancy and provide reliable feature data for the next stage of the algorithm. Semantic feature modeling is based on temporal variation of low-level features of a video object. Dynamic Bayesian networks (DBNs) are then used to characterize the spatial-temporal nature of the video objects. The system states in the proposed DBNs directly correspond to the physical concepts. Thus, the decoding of the DBN system states from observable variables is a natural interpretation of the behavior of the video objects. Since the video objects are generally considered as the dominant semantic features of video clips, the proposed scheme provides a powerful methodology for content description, which is critical for large scale MPEG-7 applications.

Chapter 9: Advances in Digital Home Photo Albums

With rapid advances in sensor, storage, processor, and communication technologies, home users can now afford to create, store, process, and share large digital images

and subsequently, video albums. With more and more digital albums being accumulated, home users need effective and efficient tools to organize and access images and videos in a semantically meaningful way without too much of a manual annotation effort. To address this genuine need, an international collaboration project (2000-2003) among CNRS, France, School of Computing, National University of Singapore, and Laboratories for Information Technology, Singapore was formed to develop the next generation Digital Image/Video Album (DIVA) for home users. The goal of this chapter is to explain the work carried out in the DIVA international project. In particular, we will describe the needs of users in the domain of management and retrieval of still images, before explaining the research work done in semantic indexing and retrieval of still images

Chapter 10: Algorithm for the Retrieval of Image with Emergence Index in Multimedia

In this chapter, we discuss emergence phenomenon where we study the hidden meanings of an image. We present 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 searches where emergence is not considered, as consideration of hidden meanings could change the index of a search. We talk about emergence, emergence index and accessing multimedia databases using emergence index to locate geographic areas in this chapter along with the algorithm.

Chapter 11: An Efficient Content-Based Retrieval System for Large Image Database

Content-based image retrieval has become more desirable for the development of large image database, due to the accessibility of user preferred images in many applications, such as a photo database. This chapter presents an efficient method of retrieving images from an image database. Our system combines color, shape and spatial features to index and measure similarity of images. Several color spaces that are widely used in computer graphics are discussed and compared for color clustering. In addition, this chapter proposes a new automatic indexing scheme of image databases according to our clustering method and color sensation, which could be used to retrieve images efficiently. As a technical contribution, a Seed-Filling-like algorithm that could extract the shape and spatial relationship feature of an image is proposed. Due to the difficulty of determining how far objects are separated, this system uses qualitative spatial relations to analyze object similarity. Also, the system is incorporated with a visual interface and a set of tools, which allows the users to specify or sketch the images conveniently for query purposes. Besides, the feedback learning mechanism enhances the precision of retrieval. The experience shows that the system is able to retrieve image information efficiently by the proposed approaches.

Chapter 12: A Spatial Relationship Method Supports Image Indexing and Similarity Retrieval

The increasing availability of image and multimedia-oriented applications markedly impacts image/multimedia file and database systems. Image data are not well-defined keywords such as traditional text data used in searching and retrieving functions. Consequently, various indexing and retrieving methodologies must be defined based on the characteristics of image data. Spatial relationships represent an important feature of objects (called icons) in an image (or picture). Spatial representation by 2-D String and its variants, in a pictorial spatial database, has been attracting growing

interest. However, most 2-D Strings represent spatial information by cutting the icons out of an image and associating them with many spatial operators. The similarity retrievals by 2-D Strings require massive geometric computation and focus only on those database images that have all the icons and spatial relationships of the query image. This study proposes a new spatial-relationship representation model called "Two Dimension Begin-End boundary string" (2D Bε-string). The 2D Bε-string represents an icon by its MBR boundaries. By applying "dummy objects," the 2D Bε-string can intuitively and naturally represent the pictorial spatial information without any spatial operator. A method of evaluating image similarities, based on the modified "Longest Common Subsequence" (LCS) algorithm, is presented. The proposed evaluation method can not only sift out those images of which all icons and their spatial relationships fully accord with query images, but for those images some of whose icons and/or spatial relationships are similar to those of query images. Problems of uncertainty reflection of images, the query targets and/or spatial relationships thus solved. The representation model and similarity evaluation also simplify the retrieval progress of linear transformations, including rotation and reflection, of images.

Chapter 13: A Stochastic and Content-Based Image Retrieval Mechanism

Multimedia information, typically image information, is growing rapidly across the Internet and elsewhere. To keep pace with the increasing volumes of image information, new techniques need to be investigated to retrieve images intelligently and efficiently. Content-based image retrieval (CBIR) is always a challenging task. In this chapter, a stochastic mechanism, called Markov Model Mediator (MMM), is used to facilitate the searching and retrieval process for content-based image retrieval, which serves as the retrieval engine of the CBIR systems and uses stochastic-based similarity measures. Different from the common methods, our stochastic mechanism carries out the searching and similarity computing process dynamically, taking into consideration not only the image content features but also other characteristics of images such as their access frequencies and access patterns. Our experimental results demonstrate that the MMM mechanism together with the stochastic process can assist in retrieving more accurate results for user queries.

Chapter 14: Mind the Gap: Content-Based Image Retrieval and the User Interface

The user interface provides the medium between the end user and the system, and is instrumental in facilitating effective human-computer interaction. Despite being identified as one of the major research areas of the field, few studies have investigated the design of existing or new user interface metaphors or paradigms for specifying visual queries. Research has demonstrated that the design of the user interface must match the tasks, activities, and goals of the end user population for whom it is being developed. Currently, there is little empirical evidence to support the use of any interface metaphor or paradigm. What type of user interface is required to support human-computer interaction in a content-based image retrieval environment is an open research issue. This chapter presents an overview of user interface research activity and related areas in the field of content-based image retrieval.