Parallel and Distributed Multimedia Databases

S. Geisler
University of Clausthal, Germany

O. Kao
University of Paderborn, Germany

INTRODUCTION

Sensing and processing of multimedia information is one of the basic traits of human beings. The development of digital technologies and applications allows the production of huge amounts of multimedia data. The rapidly decreasing prices for hardware such as digital cameras/camcorders, sound cards and the corresponding displays led to wide distribution of multimedia-capable input and output devices in all fields of the everyday life, from home entertainment to companies and educational organisations. Thus, multimedia information in terms of digital pictures, videos, and music can be created intuitively and is affordable for a broad spectrum of users.

An important question in this context is related to the archiving of the acquired information. The old-fashioned albums with pictures from holidays, children, special occasions, and so forth are replaced by photo-CDs and DVDs. Analogously, digital videos are edited, valorised by including meta-information (occasion, place, date …) and archived on DVDs. If a particular scene, image, or sound file is needed, then one can use its memory to find the corresponding medium. This type of organisation is surely not applicable to large multimedia archives, which often exist in industrial and educational sectors and where Petabytes worth of multimedia data are produced year for year. All this information has to be systematically collected, registered, stored, organised, and classified. Therefore, in many branches professional archives for such multimedia information are established, such as document management systems, digital libraries, photo and video archives used by public authorities, corporations, broadcasting and TV companies, as well as archives for satellite and surveillance photos. The scope and spread of such systems grow day by day and lead to new demands for efficient retrieval of the archived information based on user-specific description of the sought image, video or audio.

The search for a medium similar to the given one is, due to the complexity of multimedia information, a very challenging problem and requires a number of novel mechanisms. Beside the search procedures, also methods to formulate queries, and ways to visualise the results have to be provided. Moreover, the search has to be performed efficiently in order to achieve acceptable response times for the user. Therefore, a combination of modern multimedia archives with powerful parallel and distributed architectures described in this article is mandatory for the integration of multimedia retrieval into real-world applications.

BACKGROUND

The necessity for organisation and retrieval of multimedia data led to development of a large number of prototypes and operational multimedia database management systems, which manage the multimedia data in terms of storage, annotation, and retrieval. In the early years this task was tended to by existing database management systems (DBMS) with multimedia extensions. The basis for representing and modelling multimedia data in such systems is so-called Binary Large Objects (BLOBs), which store images, video and audio sequences without any formatting and analysis done by the system. The media are saved in the current form in the database and their additional information – called meta-information – is inserted into the database tables. Typically, the file name, categories and additional key words entered by the user serve as meta-information. Once the user submits a key word about the sought media, the blocks with meta-information are searched using the existing database functions and compared with the input. In case of a key word match, the corresponding media is presented.

These extensions reflect a certain aspect of multimedia database systems, but this approach does not satisfy the requirements of multimedia archives, as the manual annotation of the media is too time-consuming and not applicable in real-world applications. Furthermore, key words are not sufficient to represent content of images or videos entirely (*An image says more than 1000 words*). Therefore, the media annotation and retrieval has to be content-based; that is, features describing the multimedia content have to be extracted automatically from the media itself and compared to the corresponding features of the sample medium. The functionality of such a multimedia database is well defined by Khoshafian and Baker, (1996):

"A multimedia database system consists of a high performance DBMS and a database with a large storage capacity, which supports and manages, in addition to alphanumerical data types, multimedia objects with respect to storage, querying, and searching."

The DBMS is already provided by traditional databases and therefore will not be discussed in the following sections. Instead, the focus is set on the mechanisms for multimedia retrieval and high-performance implementation.

MULTIMEDIA RETRIEVAL

The content-based annotation of multimedia data requires the integration of additional information, which can be classified into the following categories:

- Technical information describes details of the recording, conversion, and storage. Examples: filename, resolution, compression, frame rate.
- Extracted attributes are features that are deduced by analysing the content of the media directly. Examples: average colour or colour histograms of an image, camera motion in videos, pitch in audio files.
- Knowledge-based information links the objects, people, scenarios, and so forth detected in the media to entities in the real world.
- World-oriented information encompasses information on the producer of the media, the date and location, language, and so forth.

Technical and world-oriented information can be modelled with traditional data structures. The knowledge-based information assumes semantic analyses of the media, which is nowadays still not possible in general. However, many recent research efforts in this direction promise the applicability of semantic information in the future (Zhao & Grosky, 2002).

Most of the existing multimedia retrieval systems are specialised to work on media of a limited domain, for example news (Christel & Hauptman, 2002; Yang & Chairsorn, 2003), American football (Li & Sezan, 2002), or integrate general retrieval algorithms like face, speech or character recognition. They use features extracted from the media content to annotate and retrieve the multimedia objects, which are usually related to the colour, edge, texture, layout properties in case of images or consider object motion in case of videos or specific tone sequences for audios. Table 1 gives an overview of several prominent, specialised systems, which introduced main research retrieval concepts to the scientific community. Meanwhile, many of these systems became a part of commercial products or a part of a general multimedia archive. A survey is provided in Venters and Cooper (2000).

In the following, the retrieval workflow for multimedia data will be depicted by considering images as an example. The user has a specific image in mind and starts a query for this specific image or for similar samples. The user can for example browse the data set or give a suitable key word. However, for content-based similarity search, sophisticated interfaces are necessary:

- Query by pictorial example: the user supplies the system with a complete sample image, which is similar to the sought one.
- Query by painting: the user sketches the looked-for image with a few drawing tools (Rajendran & Chang, 2000).
- Selection from standards: lists of sample instances – called standards – can be offered for individual features.
- Image montage: the image is composed of single parts similar to a mosaic.

Table 1. Prominent examples for image, video and audio databases.

- Image databases
  - Qbic (Flickner, Sawhney et al., 1995)
  - Photobook (Pentland, Picard & Sclaroff, 1994)
  - Surfimage (Nastar, Mitschke et al., 1998)
- Audio databases
  - VARIATIONS (Dunn & Mayer, 1999)
  - MUSART (Birmingham, Dannenberg et al., 2001)
- Video databases
  - VideoQ (Chang, Chen et al., 1998)
  - Virage Video Engine (Hampapur, Gupta et al., 1997)
  - CueVideo (Ponceleon, Srinivasan et al., 1998)