0 Organizational Issues

Recommended literature

- Schmitt: Ähnlichkeitssuche in Multimedia-Datenbanken, Oldenbourg, 2005


Course Web page


- Contains slides, exercises, related papers and a video of the lecture

- Any questions? Just drop us an email…

0 Organizational Issues

- Lecture
  - 09:45-12:15 (3 lecture hours with a break)
  - Exercises, detours, and homeworks

- 4 or 5 Credits depending on the examination rules

- Exams
  - Oral exam
  - Achieving more than 50% in homework points is advised

0 Organizational Issues

- Castelli/Bergman: Image Databases, Wiley, 2002

- Khoshafian/Baker: Multimedia and Imaging Databases, Morgan Kaufmann, 1996

- Sometimes: original papers (on our Web page)

1 Introduction

1.1 What are multimedia databases?

1.2 Multimedia database applications

1.3 Evaluation of retrieval techniques
1.1 Multimedia Databases

- What are multimedia databases (MMDB)?
  - Databases + multimedia = MMDB

- Key words: **databases** and **multimedia**
- We already know databases, so what is multimedia?

1.1 Basic Definitions

- Multimedia
  - The concept of multimedia expresses the integration of different digital media types
  - The integration is usually performed in a document
  - Basic media types are text, image, vector graphics, audio and video

1.1 Data Types

- **Text**
  - Text data, Spreadsheets, E-Mail, …
- **Image**
  - Photos (Bitmaps), Vector graphics, CAD, …
- **Audio**
  - Speech- and music records, annotations, wave files, MIDI, MP3, …
- **Video**
  - Dynamical image record, frame-sequences, MPEG, AVI, …

1.1 Documents

- **Document types**
  - Media objects are documents which are of only one type (not necessarily text)
  - Multimedia objects are general documents which allow an arbitrary combination of different types
- Multimedia data is transferred through the use of a medium

1.1 Basic Definitions

- **Medium**
  - A medium is a carrier of information in a communication connection
  - It is independent of the transported information
  - The used medium can also be changed during information transfer
1.1 Medium Example

- Book
  - Communication between author and reader
  - Independent from content
  - Hierarchically built on text and images
  - Reading out loud represents medium change to sound/audio

1.1 Medium Classification

- Based on receiver type
  - Visual/optical medium
  - Acoustic mediums
  - Haptic medium — through tactile senses
  - Olfactory medium — through smell
  - Gustatory medium — through taste

- Based on time
  - Dynamic
  - Static

1.1 Multimedia Databases

- We now have seen...
  - What multimedia is
  - And how it is transported (through some medium)

- But... why do we need databases?
  - Most important operations of databases are data storage and data retrieval

1.1 Multimedia Databases

- Persistent storage of multimedia data, e.g.:
  - Text documents
  - Vector graphics, CAD
  - Images, audio, video

- Content-based retrieval
  - Efficient content based search
  - Standardization of meta-data (e.g., MPEG-7, MPEG-21)

1.1 Historical Overview

- Stand-alone vs. database storage model?
  - Special retrieval functionality as well as corresponding optimization can be provided in both cases...
  - But in the second case we also get the general advantages of databases
    - Declarative query language
    - Orthogonal combination of the query functionality
    - Query optimization, index structures
    - Transaction management, recovery
    - ...

- Retrieval procedures for text documents (Information Retrieval)
- Relational Databases and SQL
- Presence of multimedia objects intensifies SQL-92 introduces BLOBs
- First Multimedia-Databases
1.1 Commercial Systems

- Relational Databases use the data type **BLOB** (binary large object)
  - Un-interpreted data
  - Retrieval through metadata like e.g., file name, size, author, …
- **Object-relational extensions** feature enhanced retrieval functionality
  - Semantic search
  - IBM DB2 Extenders, Oracle Cartridges, …
  - Integration in DB through UDFs, UDTs, Stored Procedures, …

1.1 Requirements

- Requirements for multimedia databases (Christodoulakis, 1985)
  - Classical **database functionality**
  - **Maintenance** of unformatted data
  - Consideration of special **storage** and **presentation** devices

1.1 Requirements

- To comply with these requirements the following aspects need to be considered
  - **Software architecture** – new or extension of existing databases?
  - **Content addressing** – identification of the objects through content-based features
  - **Performance** – improvements using indexes, optimization, etc.

1.1 Retrieval

- **Retrieval**: choosing between data objects. Based on…
  - a SELECT condition (exact match)
  - or a defined similarity connection (best match)
- Retrieval may also cover the **delivery** of the results to the user

1.1 Retrieval

- Closer look at the **search functionality**
  - „Semantic“ search functionality
  - Orthogonal integration of classical and extended functionality
  - Search does not directly access the media objects
  - Extraction, normalization and indexing of content-based features
  - Meaningful similarity/distance measures
1.1 Content-based Retrieval

- “Retrieve all images showing a sunset!”

- What exactly do these images have in common?

1.1 Schematic View

- Usually 2 main steps
  - Example: image databases

1.1 Detailed View

1.1 More Detailed View

1.2 Applications

- Lots of multimedia content on the Web
  - Social networking e.g., Facebook, MySpace, Hi5, Orkut, etc.
  - Photo sharing e.g., Flickr, Photobucket, Imeem, Picasa, etc.
  - Video sharing e.g., YouTube, Megavideo, Metacafe, blip.tv, Liveleak, etc.

- Cameras are everywhere
  - In London “there are at least 500,000 cameras in the city and one study showed that in a single day a person could expect to be filmed 300 times”
1.2 Applications

• Picasa face recognition

• Picasa, face recognition example

• Picasa example

1.2 Applications

• Picasa, learning phase

1.2 Applications

• Consider a police investigation of a large-scale drug operation
  • Possible generated data:
    – Video data captured by surveillance cameras
    – Audio data captured
    – Image data consisting of still photographs taken by investigators
    – Structured relational data containing background information
    – Geographic information system data

1.2 Sample Scenario

• Possible queries
  – Image query by keywords: police officer wants to examine pictures of “Tony Soprano”
    • Query: “retrieve all images from the image library in which ‘Tony Soprano’ appears”
  – Image query by example: the police officer has a photograph and wants to find the identity of the person in the picture
    • He hopes that someone else has already tagged another photo of this person
    • Query: “retrieve all images from the database in which the person appearing in the (currently displayed) photograph appears”
1.2 Sample Scenario

- **Video Query:** (Murder case)
  - The police assumes that the killer must have interacted with the victim in the near past
  - Query: “Find all video segments from last week in which Jerry appears”

- **Heterogeneous Multimedia Query:**
  - Find all individuals who have been photographed with “Tony Soprano” and who have been convicted of attempted murder in New Jersey and who have recently had electronic fund transfers made into their bank accounts from ABC Corp.

1.2 Characteristics

- ... so there are different types of queries ... what about the MMDB characteristics?
  - **Static:** high number of search queries (read access), few modifications of the data
  - **Dynamic:** often modifications of the data
  - **Active:** the functionality of the database leads to operations at application level
  - **Passive:** database reacts only at requests from outside
  - **Standard search:** queries are answered through the use of metadata e.g., Google-image search
  - **Retrieval functionality:** content based search on the multimedia repository e.g., Picasa face recognition

1.2 Example

- **Passive static retrieval**
  - Art historical use case

- **Active dynamic retrieval**
  - Wetter warning through evaluation of satellite photos

---

1.2 Example

- Possible hit in a multimedia database

- Typhoon Warning for the Philippines
1.2 Example

- **Standard search**
  - Queries are answered through the use of metadata e.g., Google-image search

- **Retrieval functionality**
  - Content based e.g., Picasa face recognition

1.3 Retrieval Evaluation

- **Basic evaluation** of retrieval techniques
  - **Efficiency** of the system
    - Efficient utilization of system resources
    - Scalable also over big collections
  - **Effectivity** of the retrieval process
    - High quality of the result
    - Meaningful usage of the system
  - What is more important? An effective retrieval process or an efficient one? Depends on the application!

1.3 Evaluating Efficiency

- Characteristic values to measure **efficiency** are e.g.:
  - Memory usage
  - CPU-time
  - Number of I/O-Operations
  - Response time
  - Depends on the (Hardware-) environment
  - **Goal:** the system should be efficient enough!

1.3 Evaluating Effectivity

- Measuring **effectivity** is more difficult and always depending on the query
  - We need to define some **query-dependent** evaluation measures!
    - Objective quality metrics
    - Independent from the querying interface and the retrieval procedure
      - Allows for comparing different systems/algorithms

- Effectivity can be measured regarding an **explicit query**
  - Main focus on evaluating the behavior of the system with respect to a query
  - **Relevance** of the result set
  - But effectivity also needs to consider **implicit information needs**
    - Main focus on evaluating the usefulness, usability and user friendliness of the system
    - **Not relevant for this lecture!**
1.3 Relevance

- **Relevance** as a measure for retrieval: each document will be binary classified as **relevant** or **irrelevant** with respect to the query
  - This classification is manually performed by "experts"
  - The response of the system to the query will be compared to this classification
    - Compare the obtained response with the "ideal" result

1.3 Involved Sets

- Then apply the automatic retrieval system:

1.3 False Positives

- False positives: irrelevant documents, classified as relevant by the system
  - False alarms
    - Needlessly increase the result set
    - Usually inevitable (ambiguity)
    - Can be easily eliminated by the user

1.3 False Negatives

- False negatives: relevant documents classified by the system as irrelevant
  - False dismissals
    - Dangerous, since they can’t be detected easily by the user
      - Are there “better” documents in the collection which the system didn’t return?
      - False alarms are usually not as bad as false dismissals

1.3 Remaining Sets

- Correct positives (correct alarms)
  - All documents correctly classified by the system as relevant
- Correct negatives (correct dismissals)
  - All documents correctly classified by the system as irrelevant
- All sets are disjoint and their reunion is the entire document collection

1.3 Overview

- Confusion matrix:
1.3 Interpretation

- Relevant results = \( fd + ca \)
  - Handpicked by experts!
- Retrieved results = \( ca + fa \)
  - Retrieved by the system

1.3 Precision

- Precision measures the ratio of correctly returned documents relative to all returned documents
  - \( P = \frac{ca}{ca + fa} \)
- Value between \([0, 1]\) (1 representing the best value)
- High number of false alarms mean worse results

1.3 Recall

- Recall measures the ratio of correctly returned documents relative to all relevant documents
  - \( R = \frac{ca}{ca + fd} \)
- Value between \([0, 1]\) (1 representing the best value)
- High number of false drops mean worse results

1.3 Precision-Recall Analysis

- Both measures only make sense, if considered at the same time
  - E.g., get perfect recall by simply returning all documents, but then the precision is extremely low…
- Can be balanced by tuning the system
  - E.g., smaller result sets lead to better precision rates at the cost of recall
- Usually the average precision-recall of more queries is considered (macro evaluation)

1.3 Actual Evaluation

- Alarms (returned elements) divided in \( ca \) and \( fa \)
  - Precision is easy to calculate
- Dismissals (not returned elements) are not so trivial to divide in \( cd \) and \( fd \), because the entire collection has to be classified
  - Recall is difficult to calculate
- Standardized Benchmarks
  - Provided connections and queries
  - Annotated result sets

1.3 TREC

- Text REtrieval Conference
- De-Facto-Standard since 1992
- Establish average precision for 11 fixed recall points \((0; 0.1; 0.2; \ldots; 1)\) according to defined procedures (trec_eval)
- Different tracks, extended also for video data, ‘Web retrieval and Question Answering
### 1.3 Example

<table>
<thead>
<tr>
<th>Query</th>
<th>fa</th>
<th>ca</th>
<th>fd</th>
<th>cd</th>
<th>P</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q₁</td>
<td>8</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>0,2</td>
<td>0,25</td>
</tr>
<tr>
<td>Q₂</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>8</td>
<td>0,8</td>
<td>0,8</td>
</tr>
<tr>
<td>Average</td>
<td>0,5</td>
<td>0,525</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 1.3 Representation

- Precision-Recall-Curves

Which system is the best?
What is more important: recall or precision?

![Precision-Recall-Curves](chart.png)

### Next lecture

- Retrieval of images by color
- Introduction to color spaces
- Color histograms
- Matching

![Color Example](image.png)