MultiMedia Retrieval (M(M)R)
Organization + Introduction

2017-2018, term 1

Egon L. van den Broek
News of the day
Part I

Who? You! 😊
Who are you?

- **Educational background**
  - GMT
  - Computing Science
  - Mathematical Sciences
  - Informatics

- **Courses**
  - Beeldverwerking
  - Data-analyse en retrieval
  - Data mining
  - Pattern Recognition
  - Computer vision
Workload

- 10 weeks (incl. this week and exam)
- 7.5 ECTS
- 1 ECTS = 28 hours
- 210 hours workload
- 21 hours/week
- so, a lot …
What’s this?
practice
Part II

Why? Course goals
General science

When the student successfully completed this course, (s)he

• has executable knowledge on how to conduct a structured literature review
• has executable knowledge on empirical computer science’s research methods
• is able to compose an academic research report
MMR, general

When the student successfully completed this course, (s)he will have …

1. basic understanding of audio, video, and biosignals

2. basic understanding of (text-based) Information Retrieval (IR)

3. a general understanding of MMR
MMR, specific

When the student successfully completed this course, (s)he will have executable knowledge on

1. the use of distance metrics (within MMR)
2. how to benchmark MMR systems (incl. performance metrics)
3. understanding of the semantic gap
Experience the complete *real-life* processing pipeline

Use knowledge and tech from
- data mining (incl. pattern set mining)
- Information Retrieval (IR)
- image processing, computer vision, and graphics
- pattern recognition (incl. evolutionary computing)
- probabilistic reasoning

and, if you want, from
- algorithms and networks (incl. geometric algorithms)
- advanced functional programming
- scheduling and timetabling
Part III

How? Examination
Literature


• Recommended:
  – International Journal of Multimedia Information Retrieval
who hasn’t got the handbook?
Exam

• Motivation:
  – MMR requires critical amount of knowledge.
• About
  – the handbook
  – the lectures
• Slides available shortly after the lectures.
• Make your own notes (its an academic skill)
Project

• Motivation:
  – Learning by doing 😊
  – To respect MR’s complexity, you have to experience it.
  – Learn on what we can do easily, computing machinery fails terribly.
  – Includes: feature extraction, pattern matching, performance comparison
Part IV

What? Organization
Lecturers

• Egon (formally: dr.dr. Egon L. van den Broek)
• Miroslav (formally: dr.ir. Miroslav Živković)
• Frans (formally: dr. Frans van der Sluis)
• ?

• Contact:
  – @ lectures and work groups / practicals
  – room BBG-477 (Egon) and BBG-478 (Miroslav)
  – e-mails: e.l.vandenbroek@uu.nl and m.zivkovic@uu.nl
Schedule

• 14 lectures/practicals:
  – Tuesday 13:15 – 17:00
  – Thursday 09:00 – 10:45

• Nothing is compulsory (except the possible guest lecture)

• Locations:
  – Tuesday: BBG-065 (week 41: BBG-161)
  – Thursday: BBG-161 (weeks 43+44: BBG-205)
Grading

• Written exam: 1/3
• Project: 2/3
• Pass if:
  – average $\geq 5.6$ and
  – parts both $\geq 4.0$
• Retake exam:
  – is a written exam or an additional assignment
  – only if original final grade is $\geq 4.0$
Multimedia for Art ReTrieval (M4ART)

Egon L. van den Broek*, Thijs Kok, Theo E. Schouten, Eduard Hoenkamp, Peter J. Petiet and Louis G. Vuurpijl

- Combination of Information Retrieval (IR) and Content-Based Image Retrieval (CBIR)
- Users: both laypersons and experts
- Domain: The National Gallery of the Netherlands (the Rijksmuseum)
- Advanced IR: In-/exclusion of terms, boolean search, and the use of specific fields of annotations
  Advanced CBIR: Color space, quantization, distance measure, and texture selection
- Inspection of CBIR matching process through histograms
- Comparison of IR and CBIR by distance matrices
HUMAN-CENTERED CONTENT-BASED IMAGE RETRIEVAL

EGON L. VAN DEN BROEK*, THUIS KOK, THEO E. SCHOUTEN, AND LOUIS G. VUURPIJL

Content: - Total size: 30,000 images
- National art rental galleries
- National Gallery of The Netherlands (the Rijksmuseum)
- Uploaded images

Users: - Both laypersons and experts

Techniques: - Information Retrieval (IR)
- Cognitive Computer Vision (CCV)
- Content-Based Image Retrieval (CBIR)

- Keyword and Boolean search
- Auto keyword completion
- Assisted Boolean queries
- Tagclouds

- Open Source technology (Linux/Apache/MySQL/PHP/Java)
- Multilanguage support
- Supports subset search
- Asynchronous JavaScript And XML (AJAX)
- Mobile optimized front-end

- 11 color categories model embedded
- Texture analysis
- Result caching
- Benchmarked


* Corresponding author: vandenbroek@acm.org
On distance …
Augmenting games (1)
Augmenting games
Games in perspective
Part V

MMR: The dominant practice
What is Metadata? (1)

• Origin “meta”: A Greek word meaning “after”

• Metadata = data about data

• Metadata for Multimedia:
  – Descriptions
  – Management
  – Usage
What is Metadata? (2)

• Metadata for Multimedia is complex, because:
  – Objects are very large; hence,
    • Expensive to transmit and process
    • Difficult to scan and search
  – Stored in compressed formats; hence,
    • Extraction of subsets is difficult
  – Rich of content but hard to summarize in structured forms
  – Subject to a complex creation process
    • Both from a technical and an economical/rights perspective
  – Consists of multiple components, which are interrelated
Why do we need Metadata?

- **Description**
  - To create indexes and catalogues, facilitating access
- **Querying**
- **Administration**
  - Management and administration (e.g., including access rights)
- **Preservation**
  - Data archival (e.g., including migration)
- **Technical**
  - Data formats, compression ratios, scaling routines, encryption keys, security, etc.
- **Use**
  - Indicate level and type of use (e.g., multiversion and user tracking)
Classification of metadata

• Content-independent; e.g.,
  – Date of creation
  – Location of a document
  – Type of camera

• Content-dependent
  – descriptive
    • Created manually!
    • e.g., annotation
  – non-descriptive
    • Domain-independent (e.g., low-level features)
    • Domain-dependent (e.g., high-level features such as a land cover)
Metadata standards

• Dublin Core Metadata Element Set
  – Contributor
  – Coverage
  – Creator
  – ...

• Resource Description Framework (RDF/RDFS) (W3C)
  – Basic Information Model
  – Structured Property Values: Blank Nodes
  – RDF Schema
  – ...

• Moving Picture Experts Group (MPEG) (ISO)
  – Description Definition Language (DDL)
  – Descriptors (Ds)
  – Description Schemes (DSs)
  – Binary format (BiM)
  – System Tools
Cooperative annotation, in the old days

- Blobworld
- VindIT
Cooperative annotation, today
Part VI

Definitions
What is media?

One of these:

1. middle state or condition; mean.
2. something intermediate in nature or degree.
3. an intervening substance, as air, through which a force acts or an effect is produced.
4. the element that is the natural habitat of an organism.
5. surrounding objects, conditions, or influences; environment.
6. an intervening agency, means, or instrument by which something is conveyed or accomplished: *Words are a medium of expression*.
7. one of the means or channels of general communication, information, or entertainment in society, as newspapers, radio, or television.

or

– a medium of cultivation, conveyance, or expression

sources:
*http://dictionary.reference.com/browse/medium*
*http://www.merriam-webster.com/dictionary/media*
Also media

• Mass media:
  – Link between info producers and consumers: newspapers, TV, radio, internet

• Transmission media:
  – Physical means of transmitting signals: wires, optical cables, microwaves

• Storage media:
  – Physical means of storing data: ticker tape, magnetic tape, floppy/diskette, cd/dvd, usb-stick, hard disk
Multimedia: First use

• The term multimedia seems to be first used in 1962, simply meaning the use of several media
  (Merriam-Webster’s Collegiate Dictionary 10th ed. 1998)

• Used in the 1960s to describe presentations combining photographic slides and audio tapes
Multimedia: A definition

“applications that use multiple modalities to their advantage, including text, images, drawings, graphics, animation, video, sound (including speech), and, most likely, interactivity of some kind.” (p. 4)

Research topics (p.5-6):

• Multimedia processing and coding
  – e.g., audio/image/video processing, compression algorithms, multimedia content analysis, content-based multimedia retrieval, and multimedia security.

• Multimedia system support and networking
  – e.g., network protocols, Internet, wireless networks, OS, servers, clients, and databases.

• Multimedia tools, end systems, and applications
  – e.g., hypermedia systems, Uls, authoring systems, multimodal interaction, and integration (e.g., Aml).
Multimedia aspects: Throughput

- Network bandwidth
- Framebuffer, video card, audio synchronization
- Disk access rate
- Disk storage capacity
- Operating System requirements
Multimedia Retrieval: A definition

• A stipulate definition: *The application of (multimedia) content analysis (e.g., computer vision) to the retrieval problem (i.e., searching for multimedia in large databases / big data).*

• Content:
  – Text
  – Other modalities (e.g., audio, image, video)
Multimedia Retrieval: An example

• Content-Based Image/Video Retrieval (CBIR/CBVR)

• “CBIR/CBVR is the application of computer vision to the image/video retrieval problem.”
• Difference between CBIR and CBVR: temporal dimension
• Smeulders et al. (2000) - CBIR at the end of the early years; 4 problems identified:
  1. “CBIR techniques still yield unacceptable retrieval results,
  2. They are restricted to the domain that is covered,
  3. They lack a suitable user interface, and
  4. Are mainly technology-driven, and, subsequently, require the use of domain knowledge to fulfil their information need”
Fields of application (1)

• WWW
  – 1945: Vannevar Bush’s MEMEX
  – 1966: Douglas Engelbart: Introduced hyperlinks
  – 1970s: ARPA internetworking project
  – fact: between 14 and 22 image per webpage

• Databases
  – Museums
  – Medical
    • e.g., dermatological images, cytological specimens, 3D cellular structures, MRIs, CT brain scans

• Photobook
Fields of application (2): others & specific

- Logo retrieval
- CAD searching
- Product catalogues
- Museum collections
- Photo archives
- Music selection
- Crime investigation, law enforcement
- Video searching
- Encyclopedia search
- Copyright protection
Retrieval aspects

- Acquisition and storage
- Feature extraction
- Feature indexing
- Query formulation
- Feature matching
- Result visualization
- Feedback loop
Retrieval framework

- Media: images, sound, video, 3D scenes
- Features: color, texture, shape
- Indexing: feature space, object space
Query formulation:

Query by Example

query:

target:
Query formulation:

Query by Sketch
Key-word Based Retrieval

Convert content to keywords at insertion? run-time? once the query is given?

Questions
expressiveness
ambiguity in language (synonyms, homonyms, anonyms)
Enabling technologies

- data mining
- signal (incl. image) processing
- Information Retrieval (IR)
- pattern recognition (incl. geometric algorithms)
- Usability eXperience design
Problems MMR faces (p. 4-7)

• polysemy
• gravity of the sample
• incomplete categories and magic values
• curse of dimensionality
• performance optimization
• noise, distortions, and missing data
Where are we now?

1. How **stupid** a computer still is in perceiving
2. The language / picture barrier: **learning to see**
3. The role of **invariance** and what **features** to use
4. **Multimedia** integrated approach
5. Features and similarity have to be become **perceptual**
6. **Interaction** needs are extreme
7. Compute and algorithmic **power** needs for the big databases