

Integrating content-based image retrieval into clinical practice

Henning Müller
Division of Medical Informatics
University Hospitals of Geneva
5.5.2003

Overview

- Content-Based access to visual data
- medGIFT
- Possible benefits
- Potential problems
- Implementation issues
- Our integration strategy
 - Step by step
 - Proof of performance
 - Specialization
- Conclusions

Content-based access to visual data

- Necessary to manage large amount of multimedia information available, i.e.. on the Internet
 - Not structured, not annotated
- Search images by visual content only
 - Automatically extracted features
 - Avoid expensive annotation, error correction
- QBE – Query by example(s)
 - Positive, negative examples, relevance feedback
- Semantic gap
 - Visual features used are low level
 - Concepts that are searched for are often high level

medGIFT

- Open source image retrieval system (www.gnu.org/software/gif/)
- Extensible architecture
 - Features, storage methods, etc.
- Adaptations necessary for the medical domain
 - More grey levels, texture filters, less colours
 - 32-64 grey levels perform best
- Standard communication interface, MRML based on XML
- Integration of medGIFT into the case database system Casimage
 - Currently more than 35000 images

Benefits of an integration into medical practice

- Teaching
 - Lecturers can find good example images
 - Students can browse image databases on their own
 - Web-based teaching
- Research
 - Studies can include visual features
 - Choice of cases can be optimized
- Diagnostics
 - Diagnostic support
 - Similar cases can supply «a second opinion»
 - Case-based reasoning, evidence-based medicine

Potential problems

- Acceptance of new technologies
 - Fear of technology
 - Use of a tool needs to be learned
 - Advantages and problems need to be understood to really profit from the image retrieval technology
- Implementation difficulties
 - Does a completely new application need to be developed
 - Do existing applications offer interfaces for integration
 - Interactivity of the algorithms
 - High load, large databases
- Quality issues
 - Is the system functioning?
 - Are we evaluating the correct thing? Can I obtain images?

Implementation issues

- Based on open source tools
 - GIFT, Linux
 - Use of system with any browser on any platform
- Clear definition of communication interfaces for the integration into various applications (CasImage, PACS, ...)
 - Web services (XML, http, etc.)
 - MRML as retrieval language
- Creation of a toolbox to adapt the system to various application areas
 - Visual features used
 - User profiling
 - Data mining techniques

Integration strategy

- Step-by-step integration
- Start with a case database system
 - Casimage is used by a large number of MDs
 - People know the user interface
 - Additional button for similarity search does not need much learning
 - Functionality will be learned slowly in this way
- Specialization to get into real clinical practice
 - Lung image retrieval
 - Dermatologic image retrieval
- Say what we can do and what we cannot do
 - Evaluation of the system performance

Specialization

- Lung image retrieval (HRCT)
 - Diagnostics depend strongly on texture of the images
 - Diagnostics are difficult
 - First projects do exist
 - Interaction is necessary to get optimal results
 - ROI, choice of slices
- Dermatologic image retrieval
 - Colour and texture features are important
 - Find similar cases to ease classification
 - Avoid invasive actions
- Image groups are very different from each other



Proof of performance

- Evaluation of retrieval results
 - Creation of a representative database
 - Creation of query tasks
 - Creation of ground truth data
 - Development of performance measures
 - Evaluation of our current system
 - Optimization of our system based on such an evaluation
 - Develop better features (segmentation)
- Clinical study is in preparation
 - Ask for user needs
 - Proof the performance in practice
 - Measure acceptance and problems

Conclusions

- We are in the process of integrating content-based image retrieval applications into medical practice
- Clinical integration is a slow process
 - Impossible to change the world quickly
- Acceptance needs to be obtained with clinical studies and a proof of performance
- There is a lot of potential for the use of image processing tools in medical practice