Tutorial on Medical Image Retrieval - medGIFT -

medinfo 2004, 8.9.2004

Henning Müller, Daniel Keysers

Service of Medical Informatics
Geneva University & Hospitals, Switzerland
Aachen Technical University, Germany
Overview

- The Viper project
- The GNU Image Finding Tool
  - Download, use and support
  - Interfaces, MRML
- Casimage
- The medGIFT project
  - Interface
  - Adaptations
- Current work
  - Noise reduction, specialization
- **Visual Information Processing for Enhanced Retrieval**
- Demo: http://viper.unige.ch/
- Uses techniques well known from text retrieval
  - Large feature space with a Zipf distribution of features, similar to words in text
  - Frequency-based feature weights
  - Inverted files
  - Relevance feedback schemes
- Now: work on video retrieval, classification and semantic annotation
Visual features used

- **Global color histogram** (HSV, 18, 3, 3, 4 gray levels)
- **Color blocks** at different scales and locations
- **Histogram of Gabor** filter responses
  - 4 directions, 3 scales, quantized in 10 strengths
- **Gabor blocks** at different scales and locations
- ~85,000 possible features, 1,000-3,000 features per image, distribution similar to words in text collections
• Outcome of the Viper project
• Open **framework** for image retrieval
• Continued at several Universities
  - Germany, Australia, Switzerland, ...
• KMRML as a plugin into Konqueror
• Plugin for Gimp
• User interfaces in Java, PHP, CGI/perl, JSP
• Infrastructure at **GNU**
  - Mailing list for help and error reports
  - ftp server, web pages
The GIFT framework

- Feature extraction
- Image collection
- New images
- Storage method
- Stored index
- Feedback algorithm
- Feature weighting
- MRML Interface
- Query engine
- Wavelets
- Regions
- ...
- Open communication interface
- **Socket**-based communication
- Images are referred to by URLs
- **XML**-based language, Human-readable

```xml
<mrml session-id="1" transaction-id="44">
<query-step session-id="1"
resultsize="30"
algorithm-id="algorithm-default">
<user-relevance-list>
  <user-relevance-element
    image-location="http://viper/1.jpg"
    user-relevance="1"/>
<user-relevance-element
    image-location="http://viper/2.jpg"
    user-relevance="-1"/>
</user-relevance-list>
</query-step>
</mrml>
```
Functionalities of GIFT

- Works under **Linux**, ports underway for MacOS X and Windows via Cygwin
- **GPL** license, open source
- Normally a simple config; make; make install
  - Depending on the version, sometimes perl modules miss

- Gift-add-collection - add all files in a directory/sub dir
- Gift - start server at port number
- Gift-create-inverted-file – creates the index
- FeatureExtraction - can be changed depending on DB
- **Case database for teaching**
- http://www.casimage.com/, interface developed with the proprietary 4D software
- >60,000 images, 9000 images externally accessible, 500 added per week
- **Case descriptions (textual)** available in XML
  - Very varying quality
  - Mix of French and English
- **Interface is compatible to the MIRC (Medical Image Resource Center) standard** of the RSNA
• http://www.sim.hcuge.ch/medgift/ (open source)
• Project for content-based search in medical image databases
• **Goals** of the project
  • Better management of visual medical data (retrieval)
  • **Visual Knowledge Management**
    • Textual and visual data
  • Diagnostic aid
    • Specialized retrieval (lung CTs, dermatological images)
  • Access to PACS data
  • In the short term:
    • Research
    • Teaching
The user interface

Query image

Link to casimage

Diagnosis

Similarity score
Main projects within medGIFT

- **Evaluation** of image retrieval algorithms
  - Reference data sets, imageCLEF
- **Specialized application** domains
  - Lung CT retrieval
  - Dermatology
  - Pathology
- Development of a feature framework
  - Feature selection
- Combination of **textual** and **visual** features
  - Which combination for which task?
  - Systematic testing with the imageCLEF data
Lung CT segmentation
• Many images contain a large background part that is basically **noise**
  • Or even worse if colored or textured
• Teaching file contains **frame** around images
Retrieval with/without background
Conclusions

- medGIFT/GIFT is an open source **framework**, free of charge for medical image retrieval
- Makes it easy to get a project started
- Has an open **infrastructure** and a framework of interchangeable components
- Exchange of components between research groups is one of the goals