(a) Benefits from content-based visual data access in radiology

(b) University and University Hospitals of Geneva

Service of Medical Informatics

Geneva, Switzerland

(c) Henning Müller (PhD), Antoine Rosset (MD), Arnaud Garcia, Jean-Paul Vallée (PhD, MD), Antoine Geissbuhler (MD)

(d)

(e) RSNA infoRAD space # 9508, invited for submission to radiographics

(f) Tel. ++41 22 372 6175

Fax ++41 22 372 8680

henning.mueller@sim.hcuge.ch

Address of corresponding author:

Henning Müller

Service of Medical Informatics

University and University Hospitals of Geneva

24, Rue Micheli-du-Crest

1211 Geneva 14

Switzerland

(a) Benefits from content-based visual data access in radiology

Abstract (max 250 words)

This article describes through concrete application scenarios the possible benefits that applications using content-based visual information retrieval algorithms can provide for radiological practice. Purpose of the article is to make people aware of the technologies available and their limitations. The article is also supposed to motivate radiologists to use image retrieval systems, and to use their medical knowledge and experience to augment retrieval quality and improve system developments. The literature is reviewed citing positive results and example systems as well as common problems of existing systems.

The image retrieval system medGIFT and its integration into a radiology teaching file system are presented. An evaluation of retrieval quality using query topics that represent the teaching database well, using a gold standard generated by a radiologist was performed. Results (on average 14 of the first 20 results are relevant) underline our thesis that the technology is ready for the use in non-critical domains such as teaching, especially when using relevance feedback. For the evaluation of the technology as a diagnostic aid, focused databases will need to be created including gold standards generated by specialists to evaluate and optimize systems. Only with good ground truth data it is possible to learn and include enough knowledge into the visual features representing the images to make image retrieval usable as diagnostic aid.

Content-based image retrieval has the potential to become an important technology for radiological practice but there is still work ahead to make this vision a reality.

Brief, one-sentence summary statement

Content-based image retrieval has the potential to be an important factor in radiology research, teaching and diagnostic aid when used properly and augmented with experience and medical knowledge.

Introduction to visual information retrieval

Content-based visual data access without the use of textual descriptions is a very active research topic in computer vision and image processing. Many applications exist in the research domain as well as commercial systems. Medical applications are often cited as one of the principal areas where content-based visual queries can be beneficial. Still, only very few visual image retrieval systems have been used in clinical routine, most remain prototypes.

More on general image retrieval systems for non-medical applications such as journalists' image archives or on trademark retrieval can be found in [1,2,3]. Most current systems use the approach to formulate queries using example images, called QBE (**Query by example(s)**). This requires having a proper starting image to formulate the query. Other systems allow formulating queries by selecting regions from pre-segmented images [4] or by text and images combined [5]. In general, the images are represented in the databases by automatically extracted visual features that are supposed to correspond to the visual image content or the way we perceive it. Mainly used for image retrieval are:

- grey levels and color descriptors, in a local or global fashion,
- texture descriptors,
- shapes of segmented objects.

Grey levels in the image and their distributions or layout throughout the image are often represented through histograms that can be compared with a simple intersection or a Euclidean distance. Local grey level descriptors can be represented by the most frequent grey in a certain area or by local grey level histograms. T**extures** can be described by wavelet filter responses [6] that measure the changes of the grey levels in various directions and scales throughout the image, or features derived from co-occurrence matrices that count the frequency of neighboring grey levels in various directions and distances to describe a texture. This allows describing the scale of a texture, the principal directions and whether the changes are very quick or rather gradual. Texture descriptors make mainly sense when they are extracted from a region that is homogenous in texture. S**hape** features can be used to characterize identifiable or segmented objects. Used features are mathematical moments of the shape but also features that describe the roundness of the form or the number of changes between convex and concave segments of the contour. Often, the goal is to extract features that are invariant with respect to the size of the object and with respect to rotations. By comparing the features of two images, we can calculate a similarity score between the two. Different distance measures for comparisons exist such as the simple Euclidean or the city block distance.

In general, all the features are on a fairly low semantic level in contrast to text that might come with the images. These features are also often in contrast to the high-level semantic concepts that users are mainly looking for like an object such as a tumor or a certain texture representing a disease such as emphysema. This semantic difference between image representation and the image content is called the **semantic gap**. Another gap or information loss is the **sensory gap** that is already due to the information loss from the original physical structure to the digital image, for example, due to projection where three-dimensional structures are represented by two-dimensional images (e.g. Chest x-rays). The always-limited resolution of digital images is another reason for a sensory gap between images and reality.

Most published articles on content-based medical image retrieval seem to be either written in medical departments where a clear need for image retrieval systems exists and is often defined [7,8,9] or in computer science departments where medical data sets are used but no link with the clinical routine exists [10,11]. Only few active research projects with clear clinical goals and running prototypes exist currently. The ASSERT project (Automatic Search and Selection Engine with Retrieval Tools) of Purdue University is a rather active one. It focuses on the analysis of textures in high-resolution CTs of the lung [12]. A clinical test using the system as diagnostic aid shows an improvement in diagnostic quality, especially for less experienced practitioners [13]. 11 persons had to diagnose the same set of images once with the help of ASSERT and once without any help. The tests were performed with several weeks in between so the persons did not remember the previous results. The percentage of correct diagnoses improved from 29% to 62% with aid from ASSERT. The improvement was higher for general radiologists than for chest specialists. No group decreased its diagnostic performance. The IRMA project is also very active [14]. It works on visual similarity retrieval and automatic image classification. In IRMA, a multi-dimensional code was created to annotate image databases [15] with axes for modality, body orientations, body region examined, and the biological system under control. An image database is currently being annotated containing 10’000 images from clinical routine, mainly conventional radiographies. The medGIFT retrieval system described in this paper differs from IRMA and ASSERT. It uses a very large feature space and is based on techniques well known from text retrieval. Relevance feedback and user interaction are very important components. Another difference is the fact that it does not need classification and a-priori knowledge for retrieval. The features itself are supposed to model the visual similarity of the images. More references and a detailed review of content-based medical image retrieval systems can be found in [16].

Sometimes, medical images are also retrieved with text, only [9,17]. This cannot really be called content-based retrieval but rather context-based retrieval as the text describes the context in which the image was taken or evaluated and rarely the visual content. As text, the radiology report or the text supplied by the teaching file can be taken, if available with the image in digital form. These texts are generally treated to remove very frequent so-called stop words (like “the”, “a”). Then, stemming removes the unimportant end part (“contained”, “containing” both become “contain”) before the remaining words can be indexed and used for retrieval. The imageCLEF competition (http://ir.shef.ac.uk/imageclef2004/) shows that both text and visual features have an important influence on retrieval quality. Best results can be obtained when combining the two. Whereas text has the advantage to cover semantics, it has the disadvantage of being task and user dependent, and even the same person will annotate the same image in slightly different words when performing the same task again. When a new image is produced, there is no text available and the radiologist will need to formulate queries. Automatically extracted visual features are "objective" for one image and can be obtained without additional work.

Materials and Methods used in this article

This article gives an overview of the literature available on applications in medical visual image retrieval. From the available literature, important application fields are identified and scenarios for the potential benefits of image retrieval in radiology are presented with visual retrieval examples using the retrieval system medGIFT.

The image retrieval engine medGIFT is described in more detail. It is based on the open source software tool GIFT (GNU Image Finding Tool). Our teaching file system casimage offers integration into the visual retrieval framework and gives us access to large teaching files. CasImage is an in-house development and medGIFT is equally being developed and specialized for the use with medical images in the medical informatics service, Geneva. medGIFT retrieves images based on local and global grey level and texture similarities. medGIFT was evaluated for medical image retrieval in the context of the imageCLEF competition [18,19]. Some of the evaluation results concern the number of grey levels that deliver best retrieval results. This number is surprisingly low for optimal image retrieval. The system was used with several grey level quantizations and the performance was evaluated against a gold standard generated by a radiologist.

Applications for visual data access

Teaching

Teaching can be the first domain to really profit from content-based access methods [20]. Many teaching files such as casimage [21] or the online system myPACS (http://www.mypacs.net/) exist. The systems are supposed to give a maximum flexibility to the practitioner entering cases, being as much integrated into the clinical workflow as possible. Like this, interesting or typical cases can be exported directly from the PACS or the viewing station without the necessity of complex transformations. Inclusion of images into presentations and texts should equally be easy by drag and drop. Such an easy-to-use system gives flexibility to practitioners but also prohibits a strict control of the entered data for validity. As a consequence the data are often of mediocre quality, containing spelling errors and non-standardized abbreviations. The records stored in the casimage database also contain multi-lingual entries posing more problems. Sometimes, single records are multi-lingual as data was copy/pasted from a French document with a translation for a web demonstration being started but never finished. Content-based search can be an easy option to complement the text-based or hierarchical access methods to the data. This can allow students to browse the available data themselves in an easy and straightforward fashion by clicking “show me similar images”. This can stimulate self-learning and a comparison of similar cases and their particularities. On the other hand, lecturers can find optimal cases for teaching even in parts of the database that they did not generate themselves, which might be annotated differently from their cases or simply in another language. This can also include visually similar cases with a different diagnosis, which can be important for teaching. Good starting images can still be found using text-based or hierarchical search. Especially very large databases with more than 60,000 images (such as casimage) can instantly profit from a new way of browsing with automatically extracted visual features. Evaluation of retrieval based on the exactly same diagnoses is not necessary for navigation in teaching files and thus current quality is sufficient for this kind of research.

The RSNA already created the **MIRC** (Medical Imaging Resource Centre, http://mirc.rsna.org/) standard for sharing image data for teaching. Currently, several large databases can be queried via a web page by textual queries. It would be very interesting to index all these images in a visual form and extend the MIRC standard to allow visual queries with example images. This is a very useful extension of the MIRC protocol.

Research

In research, the situation is very similar to teaching. The quality of retrieval does not always have to be on a diagnosis level and a little time for browsing can be spent to optimize query results. Content-based methods can be used in a variety of applications to complement text. They are an option for the retrieval of images of a certain kind to be included into a study. Visual access can also be used as a quality control to find images that might have been misclassified. Images of newly discovered diseases can attempt to be found in old databases even when it was not clear how exactly they were indexed textually. Visual features can be included directly into medical studies. What are the visual features that patients of certain diseases in a certain stadium of the disease have? Data mining in the visual features can be used to find potentially important visual characteristics of diseases or visual differences between diseases. One goal is a real **visual knowledge management**, where images and associated textual data can be analyzed together. Multimedia data mining can lead to unknown links between visual features and diagnoses or other patient demographics. The implicit information that is stored in an image, the textual case description, treatment and outcome can consequently be used to improve future, similar cases.

Diagnostic aid

Most systems that are currently described in the literature are tools destined for diagnostic aid. Visual features have been used to aid lung diagnostics [12], to classify pathology slides [22], and melanoma images [23], and for many more applications. Figure 1 shows an example for an application of image retrieval as a diagnostic aid for lung CT images with a typical result.

(Figure 1)

An image retrieval system can help in places where the diagnosis depends strongly on direct visual properties of images in the context of evidence-based medicine [24] or case-based reasoning [25]. Main problem is the evaluation of systems for diagnostic aid. Most often, only a very small database is extracted and systems are optimized on this database and then evaluated. This cannot lead to good results as the algorithms need to have a much larger base for the optimization, otherwise algorithms will not work on other images as they are too specialized. Another "problem" is the advancement of medical imaging techniques. New techniques deliver other, often better images. For automatic retrieval and analysis of images this means that the algorithms might not work with the new images in the same way they did on the old ones.

There is a clear need for tools that can easily be adapted for various fields of applications and that can learn the features based on a new group of images and imaging techniques. Systems need to be frameworks of reusable components where each component can easily be replaced. The basis for evaluation is the availability of image databases and ground truth for various tasks. Initiatives for such evaluation databases are underway in various organizations [26]. The need for standard data sets cannot be underestimated. In areas such as text retrieval, standard test sets and databases have lead to a significant improvement in retrieval quality. Problem and importance of reference databases have already been discussed in the 1970s [27]. Research in content-based medical image retrieval can profit largely from such datasets. Ground truth needs to be available to advance research through a comparison of techniques on the same basis.

PACS and electronic patient record

Of course, the goal of image retrieval has to be an integration of content-based access into various clinical applications such as the PACS and the electronic patient record. This has already been proposed several times [10,11,17]. Still, the main problem of integration into PACS is the sheer amount of data that is being produced in hospitals. Without a proper selection algorithm for cases and slices, the indexed data will quickly become unmanageable, especially when using modern multi-slice devices that produce hundreds or thousands of images in a single series. Often, these problems are neglected in the literature.

Integration into the electronic patient record and access to all cases via content-based retrieval as well as textual retrieval would of course be an ultimate solution to be able to use all implicit knowledge being stored in the images and their accompanying textual information. Still, for such a scenario, multiple problems will need to be solved and appropriate retrieval algorithms for all sorts of images need to be implemented. Problems will also include privacy of the patients as their treatment data is used to improve the treatment of new cases.

Limits of automatic visual retrieval

Many possibilities of image retrieval have already been discussed but there are also several limits and problems. Most problems are linked to the low-level visual features being used. The system does not know its limits and search is not semantic but based on broad visual appearance, only.

Figure 2 is an example with limited retrieval quality. Although the first retrieved image is relevant, among the next images shown on screen are several that are not at all scintigraphic images. This is due to the rather unsharp lines and the light grey background color. Other images with the same grey background and similarly unsharp objects are found.

[Figure 2]

Only with relevance feedback, marking several images as relevant or non-relevant, the system can refine the search and find only scintigraphic images The system does not know which part is a-priori the most important part of the image for the user. Only when feeding back several images, the system can adapt to the user needs.

The medGIFT retrieval system

MedGIFT was developed to work together with casimage (http://www.casimage.com/), a radiological teaching file that has been in daily routine use for several years now [21,28]. More than 60,000 images from more than 10,000 medical cases have been indexed, so far. The database is available on the Intranet of the hospital, with a smaller database being publicly available via Internet and MIRC. MedGIFT itself is an image retrieval engine [29]. It is based on the open source system GIFT (http://www.gnu.org/software/gift/), outcome of the viper project of the University of Geneva (http://viper.unige.ch/). This system offers components for content-based indexing and retrieval of images such as feature extraction algorithms, feature indexing structures and a communication interface called MRML (Multimedia Retrieval Mark-up Language, http://www.mrml.net/). The interface allows for an easy integration into various applications such as a teaching file, document management systems or tools for diagnostic aid. GIFT uses techniques from text retrieval such as frequency-based feature weights, inverted file indexing structures and relevance feedback mechanisms [30]. Frequency-based feature weights mean that the importance of visual features is determined by their frequency in the image and by the frequency in the entire collection of all images, similar to the weighting of words in text retrieval engines. Rare words contain more information and are more important than frequent words. The inverted file structure is also commonly used in textual search engines such as google. Inverted file means that the index is not based on documents that refer to features (or words) but on the features that point to the documents in which they appear. This is in analogy to google where an index of all words exists and for each word a list of web pages that contain it. As visual features to represent images, four feature groups are chosen:

- local and global texture features based on responses of Gabor filters;
- color/grey scale characteristics on a global image scale and locally within image regions.

Gabor filters measure the change in the image in a certain direction and scale. This means that it describes a texture with respect to its directions as well as with respect to the size and strength. Small or slow changes can easily be distinguished from quick and large changes. Local features are obtained by successively dividing the image into four regions of the same size and extracting the mode color of each region. This creates a multi-scale representation of the image. Figure 3 shows an example of such an image representation with several blocks at various scales.

[Figure 3]

Local Gabor filters allow determining in which region which shapes or textures occur. The potential feature space is very large (85,000 possible visual features). Each image contains roughly 1’000-2’000 features. Frequencies of visual features are similar to frequencies of words in text. The weighting scheme consequently weights rare features higher than frequent features in analogy with text search. More details on the GIFT technology can be found in [30].

To improve results with medical images that are primarily in grey scale, the number of grey levels was increased from the 4 grey levels of GIFT. For color photographs, the grey levels are shown to be unimportant for retrieval as the human visual system is less sensitive to them than to colors. The number of texture descriptors based on Gabor filter responses was equally raised as textures are expected to be more important in medical images than in color photography. Best overall results in first tests were obtained when using ~4-16 grey levels which is a surprisingly small number. To fine-tune the number of grey levels, we simply index the database in various ways and then evaluated the results for each quantization against a gold standard defined by a radiologist [31]. More tests are needed to define an optimal number of grey levels for each query task. A much larger number of grey levels seems to create too specific queries and miss relevant images. This is far from the 256 gray levels that JPEG offers and even further from the resolution of CR or DR images in DICOM. Still, low-level features for retrieval work better when this information is less specific. A new user interface based on php (http://www.php.org/), a scripting language to create web-based interactive applications, was developed showing the diagnosis of retrieved images under the image thumbnail and is linked with the teaching file as can be seen in Figure 4.

[Figure 4]

The retrieval engine allows submitting an unlimited number of images combined as query and also images as negative examples or negative feedback to focus and refine the research further. On screen, the images are sorted by visual similarity to the query image(s) and the similarity score is displayed under the image. The diagnosis and the level of similarity are also shown. When clicking on one of the images shown, the case database is opened with the corresponding case, including a textual description and further images in full resolution. The system is an interactive tool, which means that response times need to be below 1 second [32]. On a current Pentium 4 computer with 2.8 GHz using a database of 9’000 images the response times for one-image-queries are always below 0.5 seconds.

Discussion

The examples and the description of the casimage/medGIFT combination have shown that content-based image retrieval can be used for the management of medical image data. Still, there are several open questions and problems to solve. One important question is the evaluation of medical image retrieval systems using textual as well as visual retrieval. A benchmarking event for image retrieval system comparison has been established at the CLEF conference (Cross Language

Evaluation Forum). A medical image retrieval task was added in 2004 and 11 research groups from Europe, North America and Asia participated. A first evaluation of our system shows that when using one step of relevance feedback, on average 14 of the first 20 images are relevant which shows that the technology can be used in non-critical domains such as the search of interesting cases for teaching.

Current applications are often extremely specialized for a very small application domain and hard to adapt to new requirements and new types of images, or they are extremely general without the possibility to use them for diagnosis-based (specialized) retrieval. New image retrieval projects will need to be based on common platforms to allow on the one hand the important specialization for clinical domains using as much a-priori information as possible and on the other hand very general retrieval in PACS-like databases or teaching files with a large variety of images. Not only a single research group should share such a platform. Also between several research groups such a sharing of source code should be done so new technologies and features can easily be implemented and also be compared. Reimplementation of basic functionalities needs to be avoided. Specialization is very important to get applications working and into clinical routine for tasks where the radiologists can profit from the help. Extremely important is the evaluation of algorithms on real-life data. Evaluation databases will need to be generated for specialized retrieval, including ground truth for the task being evaluated. The importance of evaluation cannot be underestimated. Projects for the identification of interesting medical imaging problems and on the generation of reference image data sets are underway in the US and Europe [26]. One important factor for image retrieval research is also the evaluation of the user behavior with a retrieval system. It is important to adapt systems to user needs using interfaces that the users accept [33].

With respect to content-based data access it is important to explain the technology, its potential and problems so expectations are realistic and users are not promised perfect retrieval results. System improvements are only possible through several loops of feedback to include as much medical knowledge as possible into the retrieval engine for a certain task. Close cooperation between radiologists and computer scientists will be necessary for successful projects.

Conclusion

Image retrieval systems in the medical domain are in the process of getting into first applications to complement the conventional text-based search. They allow accessing and navigating in extremely large visual archives and extract hidden information without the high cost of manual annotation and codification of databases. Still, visual access to databases will stay a complement to text-based search and will at least in the foreseeable future not replace it. It is important that the two are developed closely together. Still, to get acceptance in the clinical domain, there is a clear need for real clinical applications that use content-based access mechanisms. Only working clinical applications will help to get acceptance in the medical community for more than "playing" with a retrieval system. To achieve this, systems will need to include as much medical knowledge as possible. A very close cooperation between medical practitioners and medical computer scientists will be necessary to achieve this goal. Promising applications will need to be identified and implemented based on a framework of components for image retrieval, so redevelopments of software are avoided and easy adaptation of the software will be possible.

Acknowledgements

Thanks to Julien Vignali for finishing our user interface in time for the RSNA presentation.

References

[1] Smeulders AWM, Worring M, Santini S, Gupta A, Jain R, Content-Based Image Retrieval at the End of the Early Years, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (12):1349-1380, 2000.

[2] Enser PGB, Pictorial information retrieval. (Progress in Documentation) Journal of Documentation 51(2), 1995, 126-170.

[3] Flickner M, Sawhney H, Niblack W et al. Query by Image and Video Content: The QBIC system. IEEE Computer, 28(9):23-32, September 1995.

[4] Carson C, Thomas M, Belongie S, Hellerstein JM, Malik J, Blobworld: A System for Region-Based Image Indexing and Retrieval, Conference on Visual Information Systems, Amsterdam, The Netherlands, pp. 509-516, 1999

[5] Westerveld T, Image Retrieval: Content versus Context, In Content-Based Multimedia Information Access, RIAO 2000, pp 276-284, Paris, France, 2000.

[6] Vetterli M, Kovacevic J, Wavelets and subband coding, Prentice Hall, 2000.

[7] Lowe HJ, Antipov I, Hersh W, Arnott Smith C. Towards knowledge-based retrieval of medical images. the role of semantic indexing, image content representation and knowledge-based retrieval. In Proceedings of the Annual Symposium of the American Society for Medical Informatics (AMIA), pages882-886, Nashville, TN, USA, October 1998.

[8] Tagare HD, Jaffe C, Duncan J. Medical image databases: A content-based retrieval approach. Journal of the American Medical Informatics Association, 4(3):184- 198, 1997.

[9] El-Kwae E, Xu H, and Kabuka MR. Content-based retrieval in picture archiving and communication systems. Journal of Digital Imaging, 13(2):70-81, 2000.

[10] Orphanoudakis S, Chronaki C, Vamvaka D. I2C net: Content-based Similarity Search in Geographically Distributed Repositories of Medical Images. Computerized Medical Imaging and Graphics, 20 (4): 193-207, 1996.

[11] Ogiela MR, Tadeusiewicz R, Semantic-oriented syntactic algorithms for content recognition and understanding of images in medical databases, IEEE international conference on Multimedia and Expo, Tokyo, Japan, 2001.

[12] Shyu CR, Brodley CE, Kak AC, Kosaka A, Aisen AM, Broderick LS. ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases. Computer Vision and Image Understanding, 75(1/2):111-132, July/August 1999.

[13] Aisen AM, Broderick LS, Winer-Muram H, Brodley CE, Kak AC, Pavlopoulou C, Dy J, Shyu CR, and Marchiori A. Automated storage and retrieval of thin-section CT images to assist diagnosis: System description and preliminary assessment. Radiology, 228:265-270, 2003.

[14] Lehmann TM, Güld MO, Thies C, Fischer B, Spitzer K, Keysers D, Ney H, Kohnen M, Schubert H, Wein BB, Content-based image retrieval in medical applications, Methods of Information in Medicine – to appear.

[15] Lehmann TM, Schubert H, Keysers D, Kohnen M, Wein BB. The IRMA code for unique classification of medical images. Proceedings SPIE 5033, 2003.

[16] Müller H, Michoux N, Bandon D, Geissbuhler A, A review of content-based image retrieval applications–clinical benefits and future directions, International Journal of Medical Informatics 73:1-23, 2004.

[17] Le Bozec C, Jaulent MC, Zapletal E, Degoulet P, Unified Modeling Language and Design of a Case-Based Retrieval System in Medical Imaging, AMIA Annual Symposium 1998, Nashville, TN, USA, 1998.

[18] Müller H, Geissbuhler A, Ruch P, Report on the CLEF Experiment: Combining Image and Multi-Lingual Search for Medical Image Retrieval, Proceedings of the 2004 Cross Language Evaluation Forum, Bath, England, 2004.

[19] Clough P, Sanderson M, Müller H, The CLEF Cross Language Image Retrieval Track (ImageCLEF) 2004, Springer Lecture Notes in Computer Science 3115, pages 243-251, 2004.

[20] Bucci G, Cagnoni S, de Dominicis R, Integrating content-based image retrieval in a medical reference database, Computerized Medical Imaging and Graphics, 20(4):231-241, 1996.

[21] Müller H, Rosset A, Vallée JP, Geissbuhler A. Integrating content-based visual access methods into a medical case database. In Proceedings of the Medical Informatics Europe Conference (MIE 2003), St. Malo, France, May 2003.

[22] Tang LH, Hanka R, Ip HHS, Lam R, Extraction of semantic features of histological images for content-based retrieval of images, IEEE Symposium on computer-based medical systems (CBMS), Houston, TX, USA, 2000.

[23] Schmid-Saugeon P, Guillod J, Thiran JP, Towards a computer-aided diagnosis system for pigmented skin lesions, Computerized Medical Imaging and Graphics 27:65-78, 2003.

[24] Bui AAT, Taira RK, Dioniso JD, Aberle DR, El-Saden S, Kangarloo H, Evidence-based radiology – requirements for electronic access, Academic Radiology 9:662-669, 2002.

[25] Abidi SSA, Manickam S, Leveraging XML-based medical records to extract experimental clinical knowledge – an automated approach to generate cases for medical case-based reasoning systems, International Journal of Medical Informatics 68:187-203, 2002.

[26] Horsch A, Prinz M, Schneider S, Sipilä O, Spinnler K, Vallee JP, Verdonck-de Leeuw I, Vogl R, Wittenberg T, Zahlmann G, Establishing an International Reference Image Database for Research and Development in Medical Image Processing, Methods of Information in Medicine 43, pp. 409-415, 2004.

[27] Sparck Jones K, van Rijsbergen CJ, Report on the need for and the provision of an ideal information retrieval test collection, British Library Research and Development Report 5266, Computer Laboratory, University of Cambridge, 1975.

[33] Rosset A, Ratib O, Geissbuhler A, Vallée, JP. Integration of a multimedia teaching and reference database in a PACS environment. RadioGraphics, 22(6):1567-1577, 2002.

[28] Rosset A, Müller H, Martins M, Dfouni N, Vallée JP, Ratib O, Casimage Project–a digital teaching files authoring environment, Journal of Thoracic Imaging 19(2):1-6, 2004.

[30] Squire DMG, Müller H, Müller W, Marchand-Maillet S, Pun T, Design and Evaluation of a Content-based Image Retrieval System, In Syed M Rahman ed., Design and Management of Multimedia Information Systems: Opportunities and Challenges, 7, pp. 125-151, Idea Group Publishing, 2001.

[31] Müller H, Rosset A, Vallée JP, Terrier F, Geissbuhler A, A reference data set for the evaluation of medical image retrieval systems, International Journal on Computerized Medical Imaging and Graphics, 28, pp. 295-305, 2004.

[32] Nielsen J, Usability Engineering, Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 1993.

[33] Kushniruk AW, Patel VL, Cognitive and usability engineering methods for the evaluation of clinical information systems, Journal of Biomedical Informatics 37(1):56-76, 2004.

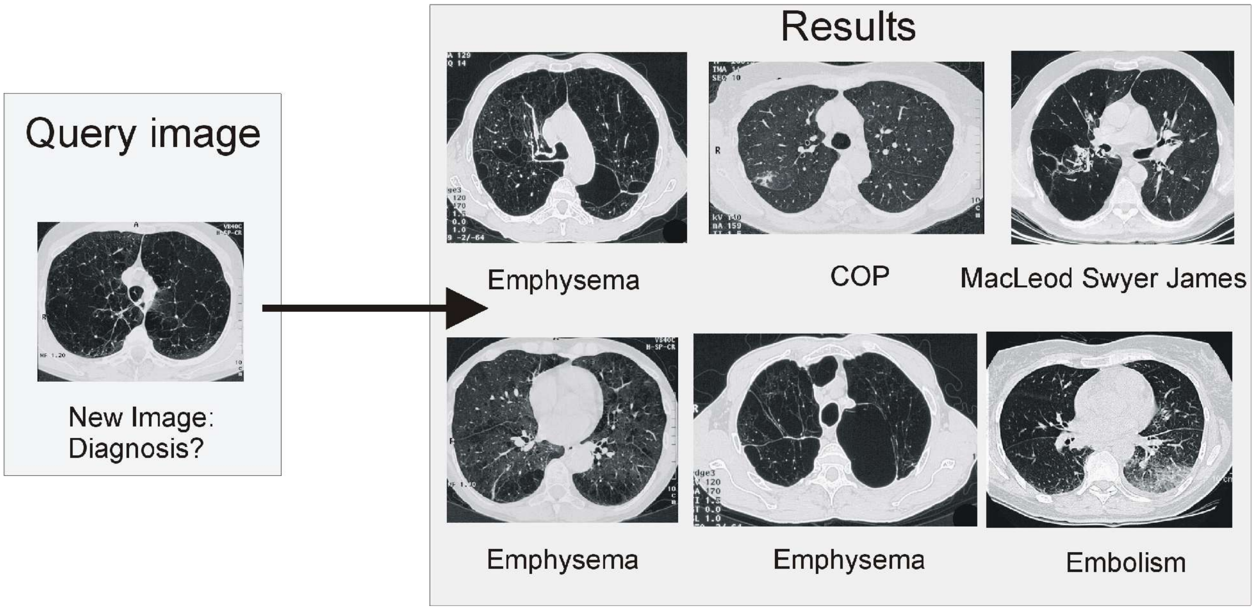

Figure 1: Content-based image retrieval as a diagnostic aid using MedGift and the database Casimage. A new image is presented and the results are shown on the right screen with their diagnoses and a link to the complete case description. The query image on the left shows emphysematous lesions with multiple, confluent, centrilobular and paraseptal areas of low attenuation without visible walls. The resulting search proposes 5 cases of Emphysema including 1 case of an unilateral emphysema (MacLeod Swyer Syndrome) and 2 cases of small area of consolidation in the pulmonary parenchyma (COP and pulmonary embolism). The typical pattern of pulmonary parenchyma destruction observed in the 5 cases of emphysema strongly suggests the diagnosis of emphysema for the query image. COP= ryptogenic Organized Pneumonia.

Figure 2: A retrieval result where the retrieval partly fails because the query does not contain much information with respect to varying grey level changes or strong textures. There is no sharply lined object in the image, which would ease retrieval.

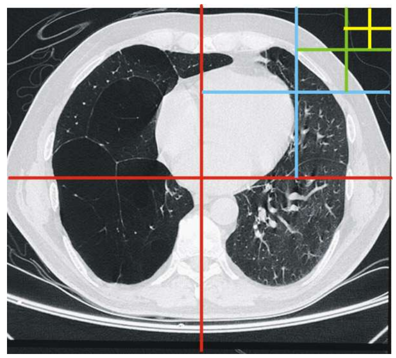

Figure 3: The image is first partitioned into four equally sized regions and this is repeated for each of the sub regions to extract local image characteristics.

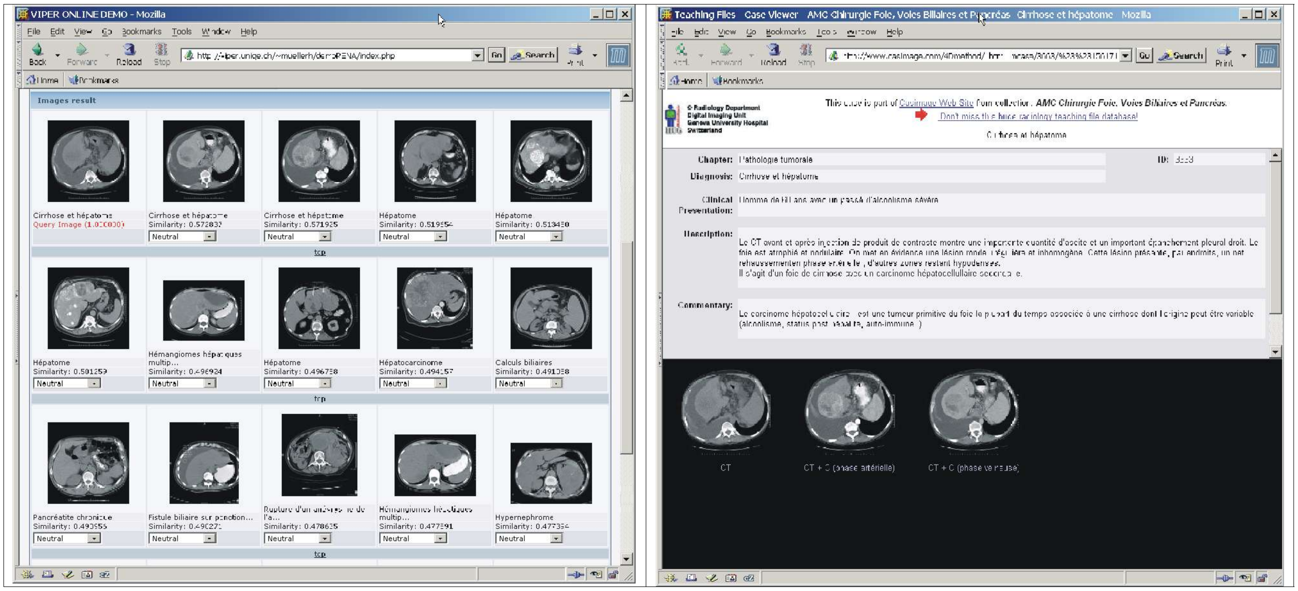

Figure 4: The interface of medGIFT and the corresponding textual CasImage case description when clicking on an image in the medGIFT interface.

Listing of Figure legends:

Figure 1: Content-based image retrieval as a diagnostic aid using MedGift and the database Casimage. A new image is presented and the results are shown on the right screen with their diagnoses and a link to the complete case description. The query image on the left shows emphysematous lesions with multiple, confluent, centrilobular and paraseptal areas of low attenuation without visible walls. The resulting search proposes 5 cases of Emphysema including 1 case of an unilateral emphysema (MacLeod Swyer Syndrome) and 2 cases of small area of consolidation in the pulmonary parenchyma (COP and pulmonary embolism). The typical pattern of pulmonary parenchyma destruction observed in the 5 cases of emphysema strongly suggests the diagnosis of emphysema for the query image. COP = Cryptogenic Organized Pneumonia.

Figure 2: A retrieval result where the retrieval partly fails because the query does not contain much information with respect to varying grey level changes or strong textures. There is no sharply lined object in the image, which would ease retrieval.

Figure 3: The image is first partitioned into four equally sized regions and this is repeated for each of the sub regions to extract local image characteristics.

Figure 4: The interface of medGIFT and the corresponding textual CasImage case description when clicking on an image in the medGIFT interface.