# The *medGIFT* project on medical image retrieval

Henning Müller, Christian Lovis, Antoine Geissbühler
University and Hospitals of Geneva, Service of Medical Informatics, 24 Rue Micheli-du-Crest, Geneva, Switzerland
Email: henning.mueller@sim.hcuge.ch

***Abstract*— Medical images are an essential part of diagnostics and treatment planning. The variety of images produced and the amount are rising constantly. Digital radiology has also brought new possibilities for the use of medical images several contexts. In fields such as evidence–based medicine or case–based reasoning medical image data can play a prominent role if tools are available to ease access to images and the accompanying textual data. Retrieval algorithms need to meet the information need of the users at a certain time. The right information needs to be accessible to the right persons at the right time.**

**The *medGIFT* project described in this paper includes several axes around the retrieval of medical images from a variety of databases and image kinds as well as for several applications. The framework is based around the open source image retrieval tool *GIFT* (GNU Image Finding Tool) and adds tools to this environment to create a system adapted for the domain–specific needs in medical image retrieval. These tools include the preprocessing of images for better retrieval, through the extraction of the main object or even through segmentation in specialised fields such as lung image retrieval. The combination and integration of *GIFT* with tools for text retrieval such as Lucene and EasyIR are other applications. Another strong point of *GIFT* is the creation of an infrastructure for image retrieval evaluation. The *ImageCLEFmed* benchmark is a result of the project and the outcome does not only help locally but is accessible for many research groups on all continents. These axes and the goals behind current developments are described in this paper.**

## I. Introduction

Production and availability of digital images is rising in all domains and as a consequence the retrieval of images by visual means has been one of the most active research areas in the fields of image processing and information retrieval over the past ten to fifteen years [1–3]. Goal is most often to retrieve images based on the visual content, only, to allow navigation even in poorly or non–annotated databases. Most systems use simple low–level features such as the image layout, shape, color, and texture features [4]. Newer systems add segmentation in often limited domains and try to match visual features and keywords to attach semantic meaning to images [5]. Still, it becomes clear that visual features can only satisfy part of the information need of users. Text is still the method of choice for most queries, especially as a starting point, whereas visual browsing can be important to refine the first results found or specific needs ("Show me chest x-rays looking similar to tuberculosis but have a different diagnosis").

In the medical domain, the need to index and retrieve images has been defined early [6–9] and a variety of applications has been developed for general image classification [10] as well as for aiding diagnostics [11]. Unfortunately, most of the projects are rather distant from clinical routine [12] and unrealistic assumptions are taken into account such as the indexation of an entire PACS [13] (Note: the Geneva radiology currently produces over 30.000 images a day and has millions stored in the PACS). Overviews of applications in the medical image retrieval domain can be found in [14, 15].

Many of the problems of image retrieval in the medical domain are linked to a distance between medical divisions and the computer sciences departments that most systems are developed in. Thus, little is often known about the use of images in a clinical settings and few of the applications work on realistically sized databases or are integrated and usable in a clinical context. Another problem is the lack of evaluation of research prototypes. Often extremely small datasets are used and settings are made to fit the system rather than the other way around. Evaluation of several systems on the same datasets has not been performed before the *ImageCLEF* initiative. The *medGIFT* project is trying to attack these problems and develop an open source framework of reusable components for a variety of medical applications to foster resource sharing and avoid costly redevelopment. A survey has been done to find out real user needs and an evaluation resource has been created in the framework of the *ImageCLEF* retrieval campaign, so research groups can compare their algorithms based on the same datasets and on realistic topics. The different axes for these developments will be described in the following chapters.

## II. An image retrieval framework

*MedGIFT* is strongly based on the GNU Image Finding tool (*GIFT*[1]) as its central piece. Main developments are on the integration of various new components around *GIFT* to create a domain–specific search and navigation tool.

### A. GIFT/MRML

*GIFT* is the outcome of the Viper[2] project of the University of Geneva [16]. It is a retrieval engine and encompassing framework for the retrieval of images by their visual content only. Several simple scripts allow to index entire directory trees, execute queries by a command line tool and generate inverted files. The visual features used are meant for color photography and include a simple color histogram as well as color blocks in various areas of the images and at several scales. Most interesting part of *GIFT* is the use of techniques well known from text retrieval. The features are quantised into bins so their distribution corresponds almost to the distribution

[1] `http://www.gnu.org/software/gift/`
[2] `http://viper.unige.ch/`

of words in texts. Then, frequency–based weights similar to typical tf/idf weightings are used [17]. To allow for an efficient feature access, an inverted file structure is used and pruning methods are implemented [18]. This allows interactive querying with response times under one second on normal Pentium IV desktops even if the database is larger than 50.000 images. This means that the feature space is extremely large with over 80.000 possible features. Usually, an image contains between 1000 and 2000 features. As *GIFT* uses *ImageMagick* to convert images, also medical DICOM images can be used for indexing without any changes to the code.

To separate the actual query engine from a user interface, the Multimedia Retrieval Markup Language (*MRML*[3]) was developed. This query language is based on direct communication of search engine and interface via sockets and eases a variety of applications such as meta–search engines and also the integration of a retrieval tool into a variety of environments and applications. The entire communication is based on the XML standard, which allows for quick development of tools. *MRML* also serves as a language to store log files of user interaction. This information can be used to improve the query performance by long–term learning from user behaviour [19]. Main goal of the framework is to avoid the redevelopment of an entire system by being able to use the base components and just work on parts that changes are needed for.

### B. *User interfaces*

As *medGIFT* is a domain–specific search tool, the user interface has different requirements from other domains. One important part is the display of not only thumbnail images for the browsing but also the text of the diagnosis. Whereas a holiday picture might bear enough information without text, for medical images this text is absolutely necessary. For further analysis much more than just a few keywords are needed because the images themselves out of the context do not seem to be extremely useful. Thus, our interface is integrated with a medical case database developed at the University Hospitals of Geneva called *Casimage*[4] [20]. Most teaching files such as *Casimage* or *myPACS*[5] have similar simple interfaces. This means that a number of images are stored together with an ordered description of a case. On the other hand, not much control is being performed on the quality of the text entered which results in records of extremely varying quality with several being empty and other containing spelling errors and non–standard abbreviations.

Figure 1 shows a typical web interface after a query was executed. The query results are displayed ordered by their visual similarity to the query, with a similarity score shown underneath the images. The diagnosis is also shown underneath the images. A click on the image links with the case database system and allows to access the full-size images.

Images are addressed via URL and it is thus possible to submit any accessible URL directly as query. Images will be

[3]`http://www.mrml.net/`
[4]`http://www.casimage.com/`
[5]`http://www.mypacs.net/`

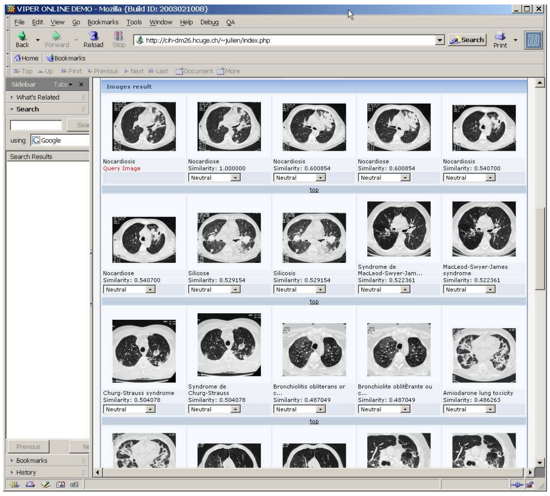

Fig. 1. A screen shot of a typical web interface for medical image retrieval system allowing query by example(s) with the diagnosis underneath the image.

downloaded, features extracted for the query, and a thumbnail will be stored locally for display in the interface. The same thing occurs for images from a local disk that can be submitted directly. This system allows for an easy access to a closed image database for basically all applications in the hospital.

### C. *Features, weightings, mix of visual and textual retrieval*

For the ease of processing all images are first converted to 256x256 pixels. Then, *GIFT* relies on four main groups of features for retrieval:

- global color features in the form of a color histogram in HSV space (Hue=18, Saturation=3, Value=3, Gray=4);
- local color features in the form of the mode color of blocks in various sizes and various regions by successively dividing the image into four equally–sized regions;
- global texture features in the form of a Gabor filter histogram using four directions, and three scales. The filter responses are quantised into 10 bins;
- local Gabor filter responses

Gabor filter responses have often shown their good performance for texture characterisation [21]. Equally the HSV color space has proven to be closer to human perception that spaces such as RGB and it still is easy to calculate [22]. For the medical domain grey levels and textures are more important than the color features that perform best on stock–photography. Thus, the *medGIFT* system uses several configurations of Gabor filters and a higher number of grey levels. Surprisingly small numbers of grey (8-16) lead to best retrieval results.

Two different weightings are used for the four feature groups. The two global histogram features are weighted according to a simple histogram intersection [23]. The two block feature groups that represent around 80% of the features are weighted according to a simple tf/idf weighting:

$$feature\ weight_j = \frac{1}{N}\sum_{i=1}^{N}\left(tf_{ij}\cdot R_i\right)\cdot log^2\left(\frac{1}{cf_j}\right), \quad (1)$$

where *tf* is the *term frequency* of a feature, *cf* the *collection frequency* of a feature, $j$ a *feature number*, $q$ corresponds to a query with $i = 1..N$ input images, and $R_i$ is the *relevance* of an input image $i$ within the range $[-1; 1]$.

Then a $score$ is assigned to a possible result image $k$ with query $q$ containing features $1..j$:

$$score_{kq} = \sum_{j} \left( feature\ weight_j \right), \qquad (2)$$

Scores are calculated for all four feature groups separately and then added in a normalised way, which leads to better results than a simple addition [18].

In connection with easyIR [6] the combination of visual and textual features was attempted in the *ImageCLEF* 2004 competition [24]. Results were the best in the competition with relevance feedback and second best for automatic retrieval. The results are simply normalised separately and then added.

### D. Image pre–treatment

Low–level image features have their problems in effective retrieval of images but other problems seem to be even more important for medical images. Normally, a medical image contains one distinct entity as the images are taken with a very specific goal in mind and under always similar conditions. Problems are the varying machines and settings used and the background information that sometimes contains information on the image taken but can be regarded as noise with respect to visual retrieval.

*1) lung segmentation:* High–resolution lung CT retrieval is one of the few domains that have been applied in a real clinical setting with success [25]. Still, all current solutions require the medical doctor to annotate the image before a classification of the tissue is made and concentrate on a very restricted number of pathologic tissue textures, only. The first and most important question is actually whether the tissue is normal (healthy) or not. For this, it is important to concentrate retrieval on the lung tissue itself, which is a problem with existing solutions [26, 27]. We implemented and optimised the algorithm to work on JPEG as well as DICOM images [28]. The results are satisfying (see Figure 2, 80% in classes 1,2) and we could well index the resulting lung parts for further retrieval.

*2) Object extraction:* As many sorts of medical images are taken with the precise goal to represent a single object, the goal is to extract the object and remove all background unnecessary for retrieval [29]. Some typical images from our database are shown in Figure 3. The removal is mainly done through a removal of specific structures followed by a low pass filter (median) followed by thresholding and a removal of small unconnected objects. After the object extraction phase much of the background is removed and only very few images had too much being removed. Figure 4 shows the results of three images. Some background structures were too big to be removed but the goal was clearly to have as few images as possible with too much removed and this was reached.

[6]http://lithwww.epfl.ch/~ruch/softs/softs.html

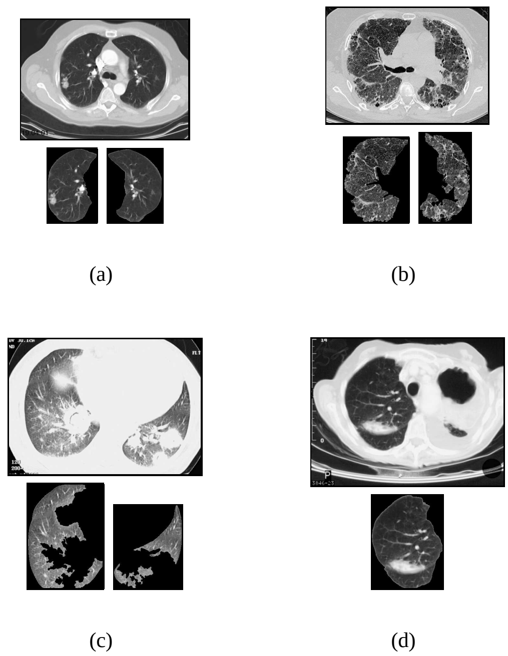


Fig. 2. Four classes of lung segmentation: (a) good segmentation, (b) small parts missing, (c) large parts missing or fractured, (d) segmentation failed (right lung missing in this case).

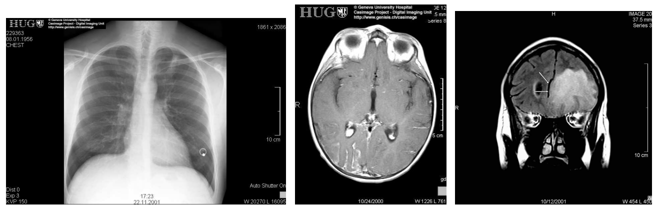

Fig. 3. Images before the removal of logos and text.

The retrieval stage shows that subjectively the results get much better and much more focused, especially with the use of relevance feedback. Still, on the *ImageCLEF* 2004 dataset, the results were actually slightly worse. Part of this can be related to the fact that the system was not part of the pooling before the ground truthing and the technique brings up unjudged but relevant images, which can influence results [30]. Another reason is the missing outline between the object and the background that can well be detected by the Gabor filters. Adding a few lines of background might improve results.

## III. Image classification/feature classification

Image classification is strongly related to image retrieval but takes into account learning data to classify images into several well-defined classes based on usually visual features [31]. In the *ImageCLEF* 2005 competition, a visual classification based on the IRMA[7] dataset was started. The dataset contains 9000 training images representing 57 classes. Then, 1000

[7]http://www.irma-project.org/

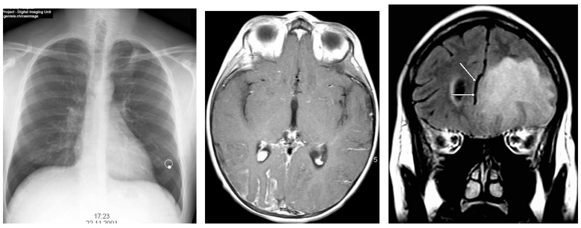

Fig. 4. Images after the removal of logos and text.

images had to be classified correctly into these 57 classes. Due to considerable time constraints, no learning could be performed on the data for our submission. A simple nearest neighbour (NN) algorithm was performed based on simple retrieval results with *GIFT*, adding the scores of the first N=1,5,10 images and taken the class with the highest score as result. Despite the fact that no training data was used, the classification rate of the best configuration was 79.4%, using 8 grey levels and 8 directions of the Gabor filters. Taking into account learning information on these classes in a way explained in [19] can strongly improve these results. Without learning the *GIFT* system had the 6th best performance with only 3 of 12 groups obtaining better results.

Another classification project has been started on the classification of lung CT textures into classes of visual observations [32]. In this project, the lung tissue is fist segmented from the rest of the CT scan. Then, the tissue is separated into smaller blocks and each of the blocks is classified into one class of visual observation such as healthy tissue, micro nodules, macro nodules, emphysema, etc. The system works completely automatic and goal is to highlight abnormal regions in a lung CT automatically. The current system is based on a small set of learning data using 12 CT series and 112 regions annotated by a radiologist. The classification between healthy and pathologic tissue has an accuracy of over 80% with a nearest–neighbour strategy and over 90% using Support Vector Machines (SVM). Part of the errors can be explained with tissue not being annotated exactly. This means that blocks that are annotated as emphysema are actually right next to an emphysema but are in fact healthy tissue (see Figure 5).

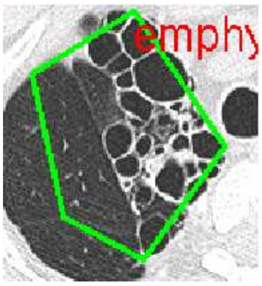


Fig. 5. An example for an annotated region of the Emphysema class, where healthy tissue is marked as well.

For the classification into several classes of visual observations, another problem becomes apparent, the extremely unbalanced training data set. Only healthy tissue, emphysema and micro nodules have a sufficiently large percentage in the training data, and these classes perform much better than classes with only one or two example blocks. An overall classification quality of around 83% has been achieved.

Some of the open source tools used for this include Weka[8], the insight toolkit itk[9] and svmlib.

## IV. Image retrieval evaluation

Much has been written on retrieval evaluation [33, 34] but most of the efforts such as the Benchathlon[10] did not result in systems being compared. There are only few articles on the use of images in a particular domain and how users would like to access and search for them.

### A. *Survey on image use in the medical field*

Images are used in many domains in increasing quantities. Digital images have started to offer new search and usage paradigms as they are accessible directly to the user and search can be performed without the necessity for perfect keywords through visual search. A few research groups have actually conducted surveys on the use of images for journalists [35] and in other domains such as libraries or cultural heritage institutions [1]. In the medical domain, to our knowledge no study on image use and searching habits has been performed as of yet, only studies on general information retrieval [36]. Thus we initiated a survey of medical image users at two institutions, the Oregon Health and Science University (OHSU) and the Geneva University Hospitals including over 30 persons. Clinicians as well as researchers, lecturers, students and librarians were asked on their habits using images, the sorts of images concerned, and the tasks they search for. Another question includes the search methods that they wanted to access images to support their particular tasks. The results are planned to be published when the surveys are finished. First results suggest that there is a strong need to search for images from certified resources. Several people stated to use google to search for images for teaching or to illustrate articles but have sometimes problems figuring out the copyright or the validity for images. To support clinical use, the support of similar cases was suggested to be extremely important as well as the search for pathologies in the electronic patient record. The need to classify images by anatomic region and modality were also given as examples for a need. First results of this study have been used to create topics for the 2005 *ImageCLEFmed* competition.

### *B.* ImageCLEFmed

*ImageCLEF* is part of the Cross Language Evaluation Forum (*CLEF*[11]) that evaluates the retrieval of documents in

[8] `http://www.cs.waikato.ac.nz/~ml/weka/`
[9] `http://www.itk.org/`
[10] `http://www.benchathlon.net/`
[11] `http://www.clef-campaign.org/`

multilingual contexts. This means that the collections can be multilingual, or the query and the document collection are in different languages. In 2003, an image retrieval task was added called *ImageCLEF*[12] using mainly grey scale images and English annotation with query topics being in several languages. In 2004, a visual task from the medical domain was added [37, 38] and participation increased from 4 groups in 2003 to 18 groups in 2004. The database of the medical task is a freely available database of the *Casimage* project and the task was organised by the *medGIFT* group. The query consists of an image only but text was available through automatic query expansion. Outcome is that visual features can enhance the quality of retrieval if used in combination with text.

In 2005, two medical[13] tasks were organised within *ImageCLEF*, one image classification task (see section III) and an image retrieval task based on a larger database containing over 50.000 images. Part of the database are the *Casimage* dataset that contains almost 9.000 images of 2.000 cases [20, 37]. Images present in the data set include mostly radiology, but also photographs, powerpoint slides and illustrations. Cases are mainly in French, with around 20% being in English. We were also allowed to use *PEIR*[14] (Pathology Education Instructional Resource) with annotation from the *HEAL*[15] project (Health Education Assets Library, mainly pathology images [39]). This dataset contains over 33.000 images with English annotation in XML per image and not per case as *Casimage*. The nuclear medicine database of *MIR*, the Mallinkrodt Institute of Radiology[16] [40], was also made available to us for *ImageCLEF*. This dataset contains over 2.000 images mainly from nuclear medicine with annotations per case in English. Finally, the *PathoPic*[17] collection (Pathology images [41]) was included. It contains 9.000 images with an extensive annotation per image in German. Part of the German annotation is translated into English. The topics are based on the survey conducted and are closer to clinical reality than the topics of the 2004 task.

In 2005, over 30 groups registered for *ImageCLEF* and over 20 groups submitted results to one of the four tasks. The evaluation of the submissions is currently being performed.

## V. Conclusions and future ideas

In conclusion it can be said that *medGIFT* is not a project on a single subject but it is rather a project encompassing a variety of subjects around medical image retrieval trying to develop a better understanding of medical imaging tasks and image use in the medical domain. Many of the sub–projects have just started and further results are expected. Goal is to use wherever possible existing open source software and solutions to keep development costs low. An integral part of the project is the creation of an evaluation framework for medical retrieval that is anchored in the *ImageCLEFmed* tasks. This project gives research groups without the contact to a medical institution the possibility to work on real medical data and realistic tasks with the goal to create applications usable in a clinical environment. To develop these tasks, surveys and the contact to medical practitioners are extremely important

Although many of the currently developed prototypes are not usable in a clinical setting, much of the knowledge can be reused for these applications and many ideas are actually evolving while developing these prototypes. One of the ideas for an easy integration of the image retrieval interface into existing clinical applications is the use of a harvesting algorithm of images from a calling web page. This means that a simple box is added to a web page to call an interface that automatically copies the images from the calling page to a local directory connects to a *GIFT* server and then allows to choose among the harvested images those relevant for a query task.

Another automatic application is the use of a DICOM header control program. As DICOM headers often show a large number of errors [42], all images from the PACS can be controlled against a reference dataset and images where problems are suspected can be sorted out for a manual correction of a proposed new header information.

For the classification of lung tissue several extensions are foreseen. Lung tissue in the same area can contain several diseases or visual observations so this needs to be included into the classification, creating the need for classifiers for each diagnosis against all other diagnoses. This leads to SVM, that also perform well on unbalanced datasets, which is also the case of the lung CT blocks.

Much still needs to be done in the field of visual medical information management and much needs to be learned about the need of medical practitioners. Only if applications are useful and applicable in a real setting, they will be used.

[12] `http://ir.shef.ac.uk/imageclef/`
[13] `http://ir.ohsu.edu/image/`
[14] `http://peir.path.uab.edu/`
[15] `http://www.healcentral.com/`
[16] `http://gamma.wustl.edu/home.html`
[17] `http://alf3.urz.unibas.ch/pathopic/intro.htm`


## Acknowledgment

Part of this research was supported by the Swiss National Science Foundation with grant 632-066041.


## References


[1] P. G. B. Enser, "Pictorial information retrieval," *Journal of Documentation*, vol. 51, no. 2, pp. 126–170, 1995.

[2] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content–based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22 No 12, pp. 1349–1380, 2000.

[3] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, "Relevance feedback: A power tool for interactive content–based image retrieval," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 5, pp. 644–655, September 1998, (Special Issue on Segmentation, Description, and Retrieval of Video Content). [Online]. Available: http://www-db.ics.uci.edu/pages/publications/1998/TR-MARS-98-10.ps

[4] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by Image and Video Content: The QBIC system," *IEEE Computer*, vol. 28, no. 9, pp. 23–32, September 1995.

[5] J. Jeon, V. Lavrenko, and R. Manmatha, "Automatic image annotation and retrieval using cross–media relevance models," in *International Conference of the Special Interest Group on Information Retrieval (SIGIR 2003)*, Toronto, Canada, August 2003.

[6] H. D. Tagare, C. Jaffe, and J. Duncan, "Medical image databases: A content–based retrieval approach," *Journal of the American Medical Informatics Association*, vol. 4, no. 3, pp. 184–198, 1997.

[7] H. J. Lowe, I. Antipov, W. Hersh, and C. Arnott Smith, "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)*, Nashville, TN, USA, October 1998, pp. 882–886.

[8] C. Traina Jr, J. M. Traina, Agma, R. R. dos Santos, and E. Y. Senzako, "A support system for content–based medical image retrieval in object oriented databases," *Journal of Medical Systems*, vol. 21, no. 6, pp. 339–352, 1997.

[9] G. Bucci, S. Cagnoni, and R. De Domicinis, "Integrating content–based retrieval in a medical image reference database," *Computerized Medical Imaging and Graphics*, vol. 20, no. 4, pp. 231–241, 1996.

[10] T. Lehmann, M. O. G¨uld, C. Thies, K. Spitzer, D. Keysers, H. Ney, M. Kohnen, H. Schubert, and B. B. Wein, "Content–based image retrieval in medical applications," *Methods of Information in Medicine*, vol. 43, pp. 354–361, 2004.

[11] C.-R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. M. Aisen, and L. S. Broderick, "ASSERT: A physician–in–the–loop content–based retrieval system for HRCT image databases," *Computer Vision and Image Understanding (special issue on content–based access for image and video libraries)*, vol. 75, no. 1/2, pp. 111–132, July/August 1999.

[12] M. R. Ogiela and R. Tadeusiewicz, "Semantic–oriented syntactic algorithms for content recognition and understanding of images in medical databases," in *Proceedings of the second International Conference on Multimedia and Exposition (ICME'2001)*, IEEE Computer Society. Tokyo, Japan: IEEE Computer Society, August 2001, pp. 621–624.

[13] H. Qi and W. E. Snyder, "Content–based image retrieval in PACS," *Journal of Digital Imaging*, vol. 12, no. 2, pp. 81–83, 1999.

[14] L. H. Y. Tang, R. Hanka, and H. H. S. Ip, "A review of intelligent content–based indexing and browsing of medical images," *Health Informatics Journal*, vol. 5, pp. 40–49, 1999.

[15] H. M¨uller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content–based image retrieval systems in medicine – clinical benefits and future directions," *International Journal of Medical Informatics*, vol. 73, pp. 1–23, 2004.

[16] D. M. Squire, W. M¨uller, H. M¨uller, and T. Pun, "Content–based query of image databases: inspirations from text retrieval," *Pattern Recognition Letters (Selected Papers from The 11th Scandinavian Conference on Image Analysis SCIA '99)*, vol. 21, no. 13-14, pp. 1193–1198, 2000, B.K. Ersboll, P. Johansen, Eds.

[17] G. Salton and C. Buckley, "Term weighting approaches in automatic text retrieval," Department of Computer Science, Cornell University, Ithaca, New York 14853-7501, Tech. Rep. 87-881, November 1987.

[18] H. M¨uller, D. M. Squire, W. M¨uller, and T. Pun, "Efficient access methods for content–based image retrieval with inverted files," in *Multimedia Storage and Archiving Systems IV (VV02)*, ser. SPIE Proceedings, S. Panchanathan, S.-F. Chang, and C.-C. J. Kuo, Eds., vol. 3846, Boston, Massachusetts, USA, September 20–22 1999, pp. 461–472.

[19] H. M¨uller, D. M. Squire, and T. Pun, "Learning from user behavior in image retrieval: Application of the market basket analysis," *International Journal of Computer Vision*, vol. 56(1–2), pp. 65–77, 2004, (Special Issue on Content–Based Image Retrieval).

[20] A. Rosset, H. M¨uller, M. Martins, N. Dfouni, J.-P. Vallée, and O. Ratib, "Casimage project – a digital teaching files authoring environment," *Journal of Thoracic Imaging*, vol. 19, no. 2, pp. 1–6, 2004.

[21] A. Jain and G. Healey, "A multiscale representation including opponent color features for texture recognition," *IEEE Transactions on Image Processing*, vol. 7, no. 1, pp. 124–128, January 1998.

[22] J.-M. Geusebroek, R. van den Boogaard, A. W. M. Smeulders, and H. Geerts, "Color invariance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 12, pp. 1338–1350, 2001.

[23] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.

[24] H. M¨uller, A. Geissbuhler, and P. Ruch, "ImageCLEF 2004: Combining image and multi–lingual search for medical image retrieval," in *Cross Language Evaluation Forum (CLEF 2004)*, ser. Springer Lecture Notes in Computer Science (LNCS), Bath, England, 2005.

[25] A. M. Aisen, L. S. Broderick, H. Winer-Muram, C. E. Brodley, A. C. Kak, C. Pavlopoulou, J. Dy, C.-R. Shyu, and A. Marchiori, "Automated storage and retrieval of thin–section CT images to assist diagnosis: System description and preliminary assessment," *Radiology*, vol. 228, pp. 265–270, 2003.

[26] S. Hu, E. A. Hoffman, and J. M. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X–ray CT images," *IEEE Transactions on Medical Imaging*, vol. 20, no. 6, pp. 490–498, 2001.

[27] G. J. Kemerink, R. J. S. Lamers, B. J. Pellis, K. H. H., and J. M. A. van Engelshoven, "On segmentation of lung parenchyma in quantitative computed tomography of the lung," *Medical Physics*, vol. 25, no. 12, pp. 2432–2439, 1998.

[28] J. Heuberger, A. Geissbuhler, and H. M¨uller, "Lung CT segmentation for image retrieval," in *Medical Imaging and Telemedicine (MIT 2005)*, Wuyi Mountain, China, 2005.

[29] H. M¨uller, J. Heuberger, and A. Geissbuhler, "Logo and text removal for medical image retrieval," in *Springer Informatik aktuell: Proceedings of the Workshop Bildverarbeitung für die Medizin*, Heidelberg, Germany, March 2005.

[30] J. Zobel, "How reliable are the results of large–scale information retrieval experiments?" in *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, W. B. Croft, A. Moffat, C. J. van Rijsbergen, R. Wilkinson, and J. Zobel, Eds. Melbourne, Australia: ACM Press, New York, August 1998, pp. 307–314.

[31] T. M. Lehmann, M. O. G¨uld, T. Deselaers, H. Schubert, K. Spitzer, H. Ney, and B. B. Wein, "Automatic categorization of medical images for content–based retrieval and data mining," *Computerized Medical Imaging and Graphics*, vol. 29, pp. 143–155, 2005.

[32] H. M¨uller, S. Marquis, G. Cohen, and A. Geissbuhler, "Lung CT analysis and retrieval as a diagnostic aid," in *Medical Informatics Europe (MIE 2005)*, Geneva, Switzerland, 2005 – submitted.

[33] H. M¨uller, W. M¨uller, D. M. Squire, S. Marchand-Maillet, and T. Pun, "Performance evaluation in content–based image retrieval: Overview and proposals," *Pattern Recognition Letters*, vol. 22, no. 5, pp. 593–601, April 2001.

[34] N. J. Gunther and G. Beretta, "A benchmark for image retrieval using distributed systems over the Internet: BIRDS–I," HP Labs, Palo Alto, Technical Report HPL–2000–162, San Jose, Tech. Rep., 2001.

[35] M. Markkula and E. Sormunen, "Searching for photos – journalists' practices in pictorial IR," in *The Challenge of Image Retrieval, A Workshop and Symposium on Image Retrieval*, ser. Electronic Workshops in Computing, J. P. Eakins, D. J. Harper, and J. Jose, Eds. Newcastle upon Tyne: The British Computer Society, 5–6 February 1998.

[36] W. R. Hersh and D. H. Hickam, "How well do physicians use electronic information retrieval systems?" *Journal of the American Medical Association*, vol. 280, no. 15, pp. 1347–1352, 1998.

[37] H. M¨uller, A. Rosset, A. Geissbuhler, and F. Terrier, "A reference data set for the evaluation of medical image retrieval systems," *Computerized Medical Imaging and Graphics*, 2004 (to appear).

[38] P. Clough, M. Sanderson, and H. M¨uller, "A proposal for the CLEF cross language image retrieval track (ImageCLEF) 2004," in *The Challenge of Image and Video Retrieval (CIVR 2004)*. Dublin, Ireland: Springer LNCS, July 2004.

[39] C. S. Candler, S. H. Uijtdehaage, and S. E. Dennis, "Introducing HEAL: The health education assets library," *Academic Medicine*, vol. 78, no. 3, pp. 249–253, 2003.

[40] J. W. Wallis, M. M. Miller, T. R. Miller, and T. H. Vreeland, "An internet–based nuclear medicine teaching file," *Journal of Nuclear Medicine*, vol. 36, no. 8, pp. 1520–1527, 1995.

[41] K. Glatz-Krieger, D. Glatz, M. Gysel, M. Dittler, and M. J. Mihatsch, "Webbasierte lernwerkzeuge fr die pathologie – web–based learning tools for pathology," *Pathologe*, vol. 24, pp. 394–399, 2003.

[42] M. O. G¨uld, M. Kohnen, D. Keysers, H. Schubert, B. B. Wein, J. Bredno, and T. M. Lehmann, "Quality of DICOM header information for image categorization," in *International Symposium on Medical Imaging*, ser. SPIE Proceedings, vol. 4685, San Diego, CA, USA, February 2002, pp. 280–287.