

MEDICAL VISUAL INFORMATION RETRIEVAL: STATE OF THE ART AND CHALLENGES AHEAD

Henning Müller, Xin Zhou, Adrien Depeursinge, Mikko Pitkanen, Jimison Iavindrasana, Antoine Geissbuhler

Medical Informatics Service, University and Hospitals of Geneva,
24 Rue Micheli-du-Crest, 1211 Geneva 14, Switzerland
henning.mueller@sim.hcuge.ch

ABSTRACT

Today's medical institutions produce enormous amounts of data on patients, including multimedia data, which is increasingly produced in digital form. These data in their clinical context contain much information and experience that is currently not being used up to its full potential. Through the digital form the data has become accessible for automatic analysis and treatment for a variety of applications. At the same time, the variety of images produced can be confusing even for trained specialists causing an information overload exists for many medical doctors. This suggests that content-based image retrieval can be a valuable tool for helping manage these data and access the right information at the right time.

This article gives a short state of the art of content-based medical image retrieval followed by a description of the *medGIFT* project on image retrieval with its main components. Then, several challenges are used to illustrate areas where much more work is currently needed to advance biomedical image retrieval. This shows that we have now progressed beyond the phase, where medical doctors transfer a database to computer scientists to only evaluate their algorithms. We conclude that visual information retrieval can have a real impact in the medical field if the techniques can adapt to this rapidly changing field and get integrated into the workflow in radiology and other medical fields.

1. MOTIVATION

Visual information is being produced in ever-increasing quantities in many fields. The digital availability of the images and videos makes new ways of managing them or extracting information from them possible. Content-based image retrieval (CBIR) [1] is one of the techniques that can help to manage even extremely large archives with often a limited textual annotation as it allows to navigate by visual content. Particular attention in recent years has been put on the automatic annotation of image collections with keywords [2] as queries with images instead of text have shown to be difficult for many users, whereas we are used to formulate an information need with words. Another idea to optimise retrieval results is through intensive user interaction or relevance feedback [3].

In the medical field, visual navigation and image retrieval have been proposed as an extremely powerful tool in several articles [4, 5, 6], but real applications and tests in clinical practice are very scarce. One notable exception is the ASSERT project [7] that has shown an improvement of diagnostic quality with the use of an image analysis system, particularly for less experienced radiologists.

This work was partly supported by the Swiss national science foundation through grant 205321-109304 and by the European Union (EU) 6th Framework program with the KnowARC project (Contract number 032691)

Currently, the information stored electronically in patient records and image archives is exclusively used in the context of a single patient. With the creation of medical teaching files and the identified need to share images and image archives [8] more data becomes available for research purposes. Legal regulations are still fairly strict for reusing any health-related data. Teaching files have been created as secondary use of personalised image data such as *CasImage*¹ and *Pathopic*². A standard format for sharing medical teaching files is MIRC³ (Medical Imaging Resource Center) and several teaching files are available through this interface. An even more valuable resource is made available through *Goldminer*⁴, images of peer-reviewed cases from the American Roentgen Ray Societies journals. Although a content-based interface is missing, it may help to get access to high quality images and associated meta data. A content-based interface to such an important image resource can help to make CBIR a valuable tool for the medical community and increase acceptance in this domain.

2. STATE OF THE ART

Two overview articles on image retrieval in the medical fields exist for a more comprehensive review [9, 10]. Other than this, most articles can be separated roughly into visionary articles from the medical field promoting the use of image retrieval [4, 5] and technical articles in computer science departments with a limited connection with the medical reality [11, 12]. Image retrieval has had very little practical impact in the medical field and most medical professionals are not aware of the possibilities it can offer.

With respect to techniques, two approaches can be separated: based on *classification* and based on *information retrieval* techniques. Whereas the *IRMA* (Image Retrieval in Medical Applications, [13]) and *ASSERT* projects are based on classification, the *medGIFT* project is rather based on information retrieval techniques. Classification-based approaches can be used for databases where a clear separation into classes is possible and ground truth is available. Typical application scenarios are the pre-classification of images by anatomic region and modality for storage or the correction of DICOM headers that have shown to be error-prone in the anatomic region field. Information retrieval techniques are useful when very large databases are used for which little ground truth is available and a class separation depends strongly on the task that a user is performing. Typical application scenarios are the navigation in large teaching files to find images even when only a limited annotation is available.

¹<http://pubimage.hcuge.ch/>

²<http://alf3.urz.unibas.ch/pathopic/>

³<http://mirc.rsna.org/>

⁴<http://goldminer.arrs.org/>

Most of the techniques applied in medical image retrieval are similar to those applied in the non-medical domain. Relevance feedback and interaction with the user is generally regarded as extremely important to find out more about the particular information needs of a user [14]. The region of interest in medical images is often fairly small and localised in a subregion of the image. On the other hand, images are usually taken under very standardised conditions and from well-defined view angles, so invariances are often different than in stock photography where invariance to particular lighting conditions is of interest. Still, some invariances and particularly the use of local interest points can prove to be successful [15]. Problems with the handling of medical images (not only for retrieval) occur at many levels: DICOM CT (Computed Tomography) or MR (Magnet Resonance Tomography) images are 12–14 Bit grey scale and thus more than what many image processing systems cope with and also more than can be shown screen. Some systems [14] simply transform the images to JPEG and then treat the usual 256 grey scales. This creates an information loss and removes the possibility to modify the images from the clinicians as they are used to in their usual viewing station environment.

To make image retrieval a success in a clinical setting a very close cooperation and much knowledge communication on particular needs are necessary. Many feedback loops are required to make such a technology accepted and limitations need to be explained.

3. THE MEDGIFT PROJECT

This section introduces the *medGIFT* project on medical image retrieval of the University of Geneva; more information can be found in [16]. The full application domain of medical information analysis is attempted to be covered, from information creation and access, to computational power, and towards real applications in contact with clinicians and finally, the evaluation of applications and practical impact.

3.1. Data access — *AneurIST*

*AneurIST*⁵ is a research project financed by the EU to help with the treatment and detection of cerebral aneurisms with over 30 partners worldwide. One important aspect in the project is the data acquisition from several clinical partners distributed among currently 5 European countries. To do so, a reusable decentralised architecture is being developed with the goal to provide direct access to data stored in a distributed way at several clinical centres. Within the Geneva hospitals we are developing an access model that is based on a data model for the disease and reuse a maximum of the data available in the patient record, including images, structured data, and free text. The goal is that for any new disease only the data model needs to be changed and mapped to the patient record. Then, an anonymised access is technically possible without having to recreate a dedicated data acquisition infrastructure for each new research project. Access to the data in several medical centres can help to get a sufficient number of cases even for diseases that occur rarely such as aneurisms (around 80 cases per year in Geneva). Of course, it is necessary to have an approval of the ethics committee of the local hospitals as every data collection process has to be acknowledged by this instance but it makes the acquisition of high quality data less burdensome and significantly reduces costs.

⁵<http://www.aneurist.org/>

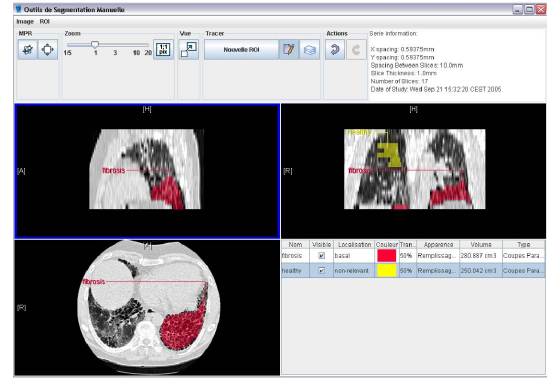


Fig. 1. Screenshot of a tool for the annotation of image regions.

3.2. Computation — *KnowARC*

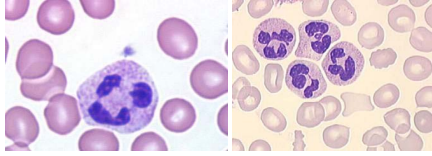
Computational aspects of image retrieval have recently played a limited role as quick database access is often possible through good database indices. Image analysis can often be accepted being slow as this is performed offline and only the actual query process needs to be accelerated to deliver quick responses to the user. Still, grid technologies have the potential to help think one step further: what can be done if nearly unlimited computing power is available? Grid technologies have been proposed for medical image retrieval [11] but most often the goal is to test an existing middleware and optimise the resource use but not to think about completely new solutions that would be impossible without grid technologies.

The *KnowARC*⁶ project creates a contact with the community of grid middleware developers and should help to influence middleware development with the goal of computationally extremely extensive solutions that are envisioned to work on large image databases such as the hospital PACS (Picture Archival and Communication system). One goal from the institutional point of view is to better use the current computer infrastructure. Whereas no research computer infrastructure is available apart from little servers financed by projects, the Geneva hospitals have just like many other large hospitals almost 6 000 desktop computers installed. Using part of this workforce for information analysis tasks can open us new possibilities and none of the confidential data needs to leave the hospitals.

3.3. Applications — *Talisman*

The major advantage of being inside a medical institution is the access to workflow knowledge and the direct communication with the medical experts. Trust needs to be built before a successful application can be developed and the processes and vocabulary need to be understood. The *Talisman* project (Texture Analysis of Lung ImageS for Medical diagnostic AssistaNce) works on the analysis of lung CTs for diagnostic aid with interstitial lung diseases [17] that are often regarded as a hard problem for the non-specialist, comprising a total of over 150 diverse diseases with unspecific symptoms. Particularly in emergency radiology the diversity of modalities and anatomic regions is large and decisions need to be made quickly. Currently, the emergency radiologists annotate image regions in the 3D lung volumes (see also Figure 1) to give us a maximum of information on the commonly observed patterns. In addition to the visual data, a total of 99 parameters are acquired for each patient such as

⁶<http://www.knowarc.eu/>



Show me blood smears that include polymorphonuclear neutrophils.
 Zeige mir Blutabstriche mit polymorphonuklearer Neutrophils.
 Montre-moi des échantillons de sang incluant des neutrophiles polymorphonucléaires.

Fig. 2. Examples for an ImageCLEFmed topic.

smoking history, age, sex, weight, size, or the use of particular medications. All these criteria are in connection with some of the diseases and they need to be acquired in high quality, often from the free text patient anamnesis as not always the information is available in structured form. The goal of the project is to supply similar cases to the emergency radiologist treating a new case and highlighting at the same time regions of the lung tissue that seem abnormal depending on these criteria. As an example, the healthy lung of a 75-year-old is clearly different from that of a 25-year-old. The availability of all the data can help to perform data mining to find unknown connections between characteristics and diseases and can further help to propose important questions to the patient that the clinician might not have asked. Detection of abnormal regions within the lungs can be regarded as a very hard task, much more difficult than distinguishing various marked regions by disease.

3.4. Evaluation — ImageCLEFmed

*ImageCLEFmed*⁷ is part of the Cross Language Evaluation Forum (CLEF⁸) that has started an image retrieval track in 2003. Overviews of the last evaluation campaigns can be found in [18, 19]. Similar to TREC (Text REtrieval Conference) and TRECVID [20], a yearly circle of events is followed: creation of large datasets, distribution of topics with realistic information needs to participants, results submission, ground truthing, evaluation, and finally a workshop to compare results. In the last years, an average of 40 research groups inscribed for the proposed tasks and around 15–20 finally submitted results. An example information need (topic) is given in Figure 2

The medical retrieval task has topics divided into visual, mixed and semantic. A purely visual image classification task based on the IRMA dataset exists as well with a strong community participation. In 2007, a hierarchical classification of images is planned, where visual systems need to decide up to which level they trust their judgement. This will give more insight into the quality of current image classification techniques.

4. CHALLENGES AHEAD

This section presents several challenges in medical image retrieval that are ahead of us and can help increase the acceptance of CBIR in the medical field.

4.1. High-quality data acquisition

One of the main problems with medical CBIR is that the images for retrieval are often taken out of their original context and much information is lost in this process. Even a trained medical specialist could

⁷<http://ir.ohsu.edu/image/>

⁸<http://www.clef-campaign.org/>

often say little about the lung CTs of a patient if nothing is known on the patient itself (age, smoking history, ...). Giving this task to a computer does not make it easier, although some hidden connections can be made if much ground truth data exists. It is really important that for the creation of medical image databases care is taken about the annotation so all important information is acquired. The same holds true for regions of interest within the images that experts can annotate. Only an exact annotation can really help evaluating research applications. This needs to be done by experienced specialists and comes at a fairly high financial cost. Still, this budget has to be included into research projects. By making these databases available to the research community many overlapping efforts can be avoided and it is mainly the responsibility of the funding institutions to force the data availability (as do the National Institutes of Health in the United States).

4.2. Multimodal data integration — case instead of image

Also for non-medical image retrieval it has become clear that using only visual information has many limits for retrieval. For a medical doctor (MD) the unit of search in a clinical setting is also not the image but rather the case. This means that all information on a case needs to be taken into account for retrieval. This can include images, structure data, and free text, and particularly several sorts of images, with varying views, or a CT and an x-ray. Combining this information to find similar cases is not trivial, as missing information has to be taken into account. It is not always standardised which exams are taken in which situation. Tackling this problem is likely to have the biggest impact in a real clinical environment.

4.3. Dissimilarity retrieval

The first question for an MD when regarding images is most often whether it is abnormal or not. The variety of healthy images is actually enormous and the borders of the two are not always clear. The healthy lung of a 60 year old smoker will be very different from the healthy lung of a 25 person in good physical shape. Thus the first step is not necessarily to find similar cases but rather find out how far away (dissimilar) from a healthy model the shown images are to judge the probability of the images being abnormal. Similarity retrieval can then be used to find similar abnormal cases.

4.4. Changing images

One big challenge are definitely the rapidly evolving imaging modalities. New modalities are developed regularly such as PET/CT combinations and existing modalities change their parameters with thinner slices and smaller inter slice distances. Varying contrast agents can also significantly change the results shown in the images. With respect to image analysis this means that segmentation might not work anymore after a change in the imaging equipment, or that varying algorithms need to be used depending on the machine the image was taken with. This requires algorithms for retrieval that can quickly be adapted to a particular setting. It also means that databases are not static but need to be adapted to current settings and require a process of integration into a medical institution that is currently only rarely the case.

Completely new fields will also be evolving with new medical departments becoming digital such as Pathology and Dermatology. Pathology images can for example be extremely large (often 100 000 by 100 000 pixels) and again a large part of the images might not be interesting whereas a few abnormal cells require further attention.

Also here, the variety of image production is large. Many staining products produce very different images and images that are not even stable over time (with stains glowing only a short period of time).

5. CONCLUSIONS

Image retrieval has the potential to have real impact in the medical field but to do so, much domain knowledge needs to be transferred and the image retrieval systems need to be adapted to this quickly changing field. Toolboxes are required that allow for a quick integration of new settings and adaptations to new images. An close cooperation of computer science specialists with medical experts needs to be attempted. A mere exchange of data is not enough and real co-operations are required. Much more is still required to find out user needs and what the real requirements in the medical field are [21]. Most often, the amount of data to be treated is underestimated. The Geneva radiology alone produced 50 000 images per day in 2006, and these numbers are rising. A treatment of all these data might not be possible at the moment but current infrastructures can cope with large volumes of data and make information inherently stored in medical multimedia data available to the medical specialists.

6. REFERENCES

- [1] Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Armarnath Gupta, and Ramesh Jain, "Content-based image retrieval at the end of the early years," vol. 22 No 12, pp. 1349–1380, 2000.
- [2] 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, Aug. 2003.
- [3] Yong Rui, Thomas S. Huang, Michael Ortega, and Sharad Mehrotra, "Relevance feedback: A power tool for interactive content-based image retrieval," vol. 8, no. 5, pp. 644–655, Sept. 1998, (Special Issue on Segmentation, Description, and Retrieval of Video Content).
- [4] Hemant D. Tagare, C. Jaffe, and James Duncan, "Medical image databases: A content-based retrieval approach," vol. 4, no. 3, pp. 184–198, 1997.
- [5] Henry J. Lowe, Ilya Antipov, William Hersh, and Catherine 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, Oct. 1998, pp. 882–886.
- [6] Hairong Qi and Wesley E. Snyder, "Content-based image retrieval in PACS," vol. 12, no. 2, pp. 81–83, 1999.
- [7] Alex M. Aisen, Lynn S. Broderick, Helen Winer-Muram, Carla E. Brodley, Avinash C. Kak, Christina Pavlopoulou, Jennifer Dy, Chi-Ren Shyu, and Alan Marchiori, "Automated storage and retrieval of thin-section CT images to assist diagnosis: System description and preliminary assessment," vol. 228, pp. 265–270, 2003.
- [8] Miachel W. Vannier and Ronald M. Summers, "Sharing images," *Radiology*, vol. 228, pp. 23–25, 2003.
- [9] Henning Müller, Nicolas Michoux, David Bandon, and Antoine Geissbuhler, "A review of content-based image retrieval systems in medicine – clinical benefits and future directions," vol. 73, pp. 1–23, 2004.
- [10] Lilian H. Y. Tang, R. Hanka, and H. H. S. Ip, "A review of intelligent content-based indexing and browsing of medical images," vol. 5, pp. 40–49, 1999.
- [11] Marcelo Costa Oliveira, Walfredo Cirne, and Paulo M. de Azevedo Marques, "Towards applying content-based image retrieval in clinical routine," *Future Generation Computer Systems*, vol. 23, pp. 466–474, 2007.
- [12] Philip Korn, Nicholas Sidiropoulos, Christos Faloutsos, Eliot Siegel, and Zenon Protopapas, "Fast and effective retrieval of medical tumor shapes," vol. 10, no. 6, pp. 889–904, 1998.
- [13] Thomas M. Lehmann, Marc Oliver Güld, Christian Thies, Benedikt Fischer, Klaus Spitzer, Daniel Keyzers, Hermann Ney, Michael Kohonen, Henning Schubert, and Berthold B. Wein, "Content-based image retrieval in medical applications," vol. 43, pp. 354–361, 2004.
- [14] Chi-Ren Shyu, Carla E. Brodley, Avinash C. Kak, Akio Kosaka, Alex M. Aisen, and Lynn S. Broderick, "ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases," (*special issue on content-based access for image and video libraries*), vol. 75, no. 1/2, pp. 111–132, July/Aug. 1999.
- [15] Lokesh Setia, Alexandra Teynor, Alaa Halawani, and Hans Burkhardt, "Image classification using cluster-cooccurrence matrices of local relational features," in *Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval (MIR 2006)*, Santa Barbara, CA, USA, Oct 2006.
- [16] Henning Müller, Christian Lovis, and Antoine Geissbuhler, "The medGIFT project on medical image retrieval," in *Medical Imaging and Telemedicine*, Wujishan, China, Aug. 2005.
- [17] Adrien Depeursinge, Henning Müller, Asmâa Hidki, Pierre-Alexandre Poletti, Alexandra Platon, and Antoine Geissbuhler, "Image-based diagnostic aid for interstitial lung disease with secondary data integration," in *SPIE Medical Imaging*, San Diego, CA, USA, Feb. 2007.
- [18] William Hersh, Henning Müller, Jeffery Jensen, Jianji Yang, Paul Gorman, and Patrick Ruch, "Imageclefmed: A text collection to advance biomedical image retrieval," *Journal of the American Medical Informatics Association*, vol. September/October, pp. 488–496, 2006.
- [19] Henning Müller, Thomas Deselaers, Thomas M. Lehmann, Paul Clough, Kim Eugene, and William Hersh, "Overview of the imageclefmed 2006 medical retrieval and medical annotation tasks," in *CLEF 2006 Proceedings, 2007 – to appear*, Lecture Notes in Computer Science.
- [20] Alan F. Smeaton, Paul Over, and W. Kraaij, "TRECVID: Evaluating the effectiveness of information retrieval tasks on digital video," in *Proceedings of the international ACM conference on Multimedia 2004 (ACM MM 2004)*, New York City, NY, USA, Oct. 2004, pp. 652–655.
- [21] Henning Müller, Christelle Despont-Gros, William Hersh, Jeffery Jensen, Christian Lovis, and Antoine Geissbuhler, "Health care professionals' image use and search behaviour," in *Proceedings of the Medical Informatics Europe Conference (MIE 2006)*, Maastricht, The Netherlands, Aug. 2006.