MEDICAL VISUAL INFORMATION RETRIEVAL: FROM TECHNIQUES TO APPLICATIONS AND EVALUATION

H. Müller[1,2], A. Depeursinge[1], J. Iavindrasana[1], X. Zhou[1], M. Niinimaki[1]
[1] Medical Informatics Service University Hospitals and University of Geneva, Switzerland;
[2] Business Information Systems, University of Applied Sciences Western Switzerland (HES SO), Sierre, Switzerland

Images play an important role in medical diagnosis and treatment planning. They are also produced in increasing varieties and exponentially rising quantities. Medical visual information retrieval is trying to help manage this rising amount of data. Unfortunately, many of the current image retrieval prototypes are fairly disconnected from clinical reality and only tested on small databases with often outdated images. In the MedGIFT project of the University Hospitals and the University of Geneva the goal is to cover the entire spectrum of visual information retrieval applications, from rather technical issues such as new visual features over infrastructures supplying computational power via grid networks and real clinical applications towards the assessment of the technology in realistic scenarios via the ImageCLEF retrieval benchmark. Medical visual information retrieval has a large potential to complement text retrieval in better using clinical data in many medical institutions but much research is still necessary to have a sufficiently high retrieval quality for clinical routine and aiding diagnosis.

1. Introduction

Content–based visual information retrieval has started already in the 1980s [1, 12] but created a very large research interest mainly in the 1990s [20]. The goal was to solve problems related to manual annotation of images with keywords such as subjectivity and high cost through the use of automatically extracted objective visual descriptions. Problems were manifold and there is still no general breakthrough in content–based image retrieval although areas such as object detection and automatic annotation of images with keywords [24] have created a few success stories. Another reason for content–based image retrieval was also the exponentially rising amount of visual data being produced through the availability of cheap digital cameras. Still, most searches for images are still performed via keywords and not via example images that are often hard to find.

In the medical field some of the problems are similar as the amount of visual data being created is exploding exponentially, particularly through the wide availability and use of tomographic examinations creating sometimes thousands of images for a single examination. Image retrieval was proposed in the medical domain already fairly early [13, 22] but without any real applications. Several specialized applications are described in two overview articles [16, 23]. In recent years, several retrieval projects such as SPIRE, IRMA[3] and MedGIFT have created a large number of publications [3, 8]. This article mainly describes the work performed in the context of the MedGIFT research project.

2. Techniques

This chapter describes technical issues related to content–based medical image retrieval in the widest sense, ranging from data access, computational issues to retrieval techniques.

2.1. Data access

One of the most important aspects of most research projects is the access to the actual data. Much time and money can often be spent on data access. One goal of our participation in the EU funded AneurIST[4] project is the creation of a reusable database infrastructure accessing

[3] http://www.irma-project.org/
[4] http://www.aneurist.org/

a life clinical patient record [11]. This means that data is used directly from the clinical record avoiding the double storage and inconsistencies of data silos often created for research projects. In this context, the simple exchange of the data model allows to change access for a new research project limiting the amount of work to access the data drastically. Data anonymisation happens on the fly in this scenario and a restrictive security architecture is in place to avoid any misuse.

2.2. Computational power

Another critical issue in imaging research is the access to computational power. Within the hospitals no central computing infrastructure exists and research projects buy their own machines for their use without any resources sharing. The KnowARC[5] brought us in connection with middle ware developers of the ARC (Advanced Resource Connector) middleware developers [9]. This allowed us to gain experiences with grid technologies and also feed back our experiences to the middleware developers for inclusion in new releases[17].

With this gained experience a grid was installed with standard desktop PC inside the Geneva hospitals. The small testbed of 20 PCS allowed to test the infrastructure with the hospital software distribution system hospital grid [18].

2.3. Retrieval techniques

Most of the retrieval techniques originally developed within the MedGIFT project were based on the GIFT[6] [21]. Particularly for new projects such as Talisman new features we necessary for the description of the lung texture. In this context Wavelet frames were tested in various configurations with very good results [6]. More recent approaches such as patch histograms [7] and salient point based analyses are currently being performed in a project on fracture images. These features currently promise the best results.

3. Applications

Besides purely technical issues, the most important goal of biomedial research is the usefulness and applications of the research results. In general, applications were employed, where medical doctors approached us with a concrete need and proposition.

3.1. Talisman

The TALISMAN (Texture Analysis of Lung ImageS for Medical diagnostic AssistaNce) project focuses on computer–aided diagnosis for interstitial lung diseases (ILDs). ILDs are a heterogeneous group of around 150 diseases, many of them being rare and presenting unspecific symptoms. The image holds a central role in the diagnosis process and high–resolution computed tomography (HRCT) is used when the chest x–ray does not carry enough information. The computer tools developed in TALISMAN include the automated detection of healthy and pathological lung tissue associated with ILDs in three–dimensional HRCT data as well as content–based retrieval of similar image series and cases. The visual information in HRCT data is extracted using Wavelet analysis with desirable properties such as isotropy, fine scale resolution and frames along with grey level histograms in Hounsfield Units. Moreover, in accordance with the interpretation approach of the radiologists with HRCT, the clinical context is used to enrich detection and retrieval. It is represented by 99 clinical parameters associated with the 15 most frequent ILDs. The values are extracted manually from the electronic health record at the University Hospitals of Geneva (HUG) and stored in a multimedia database containing the

[5] http://www.knowarc.eu/
[6] http://www.gnu.org/software/gift/

corresponding annotated HRCT image series of currently 105 cases. Using appropriate combination rules of visual and clinical features along with optimal classifier settings, the multimodal detection of lung tissue patterns allows for a detection rate of over 83% correct matches over 5 lung tissue patterns found in most prevailing ILDs [4, 5].

3.2. Fracture retrieval

The project on fracture retrieval is currently still at a very early stage [15]. The idea for the project came from surgeons at the Geneva hospitals that have a database of fractures accumulated over several years and containing around 20'000 radiographs of fractures. Goal is to find similar cases to a case being prepared for an operation to optimize the choice of the technique to be used. By having access to the past case before the operation but also directly after the operation and best even in later stages to be able to observe the long term results of an intervention. Such a process can help to improve the quality of the intervention significantly in the long term and with very good base data.

4. Evaluation

Technology assessment is the essential step of all research to identify promising techniques and separate them from poor approaches.

4.1. Benchmarks

In many domains benchmarks and standardized datasets have become available over the past years, whenever a research area starts to leave an extremely experimental stage and real applications start appearing. Within the information retrieval field, TREC[7] (Text Retrieval Conference, [25]) was definitely the earliest benchmark and has been run for many years now, with a large number of initiatives starting here in many areas. TrecVID [19] is such an initiative that has been run for over six years, now with a steadily increased number of registered participants, over 100 in 2008. Benefits are clearly visible as at conferences researchers publish with the same datasets making the quality of systems easily comparable. For many researchers such initiatives also avoid having to create expensive datasets, and particularly to have datasets of an important size.

4.2. ImageCLEF

ImageCLEF (part of CLEF — Cross Language Evaluation Forum) has started in 2003 with only four participants [2] but with an increasing number of subtasks on image annotation, medical retrieval and structured data retrieval the number of registrations has reached 63 in 2008 [14]. Main goal is the multilingual and multimodal retrieval of information. The competition showed well that combining visual and textual features for retrieval can lead to best results, but also that the combination process is very sensitive to small changes.

The medical retrieval task of ImageCLEF has over the five years of its existence created three large databases, two of them with over 66'000 medical images. One databases combined six medical teaching files in several languages [10]. The most recent dataset includes images from the medical journals Radiology and Radiographics made available in cooperation with the Goldminer[8] medical image search engine. This database is bringing the retrieval even closer to clinical routine where case from the literature related to a current case can be searched.

[7] http://trec.nist.gov/

[8] http://goldminer.arrs.org/

5. Conclusions

Content–based medical visual information retrieval has a large potential to complement text–based retrieval approaches in the medical field. Particularly in the context of evidence–based medicine and case–based reasoning the image information is currently only very little represented. Stable visual retrieval can help in many specialized domains where visual information plays an important role.

Main areas where important work is still required is the change from image–based retrieval towards a more global case–based retrieval strategy that takes into account images of a patient as well as clinical data and potentially free text such as the anamnesis. Such approaches still need much research but can bring important improvements that no visual approach could obtain. No medical doctor would diagnose on an image alone as the clinical data such as age, gender, sex do paly an extremely important role. Within the ImageCLEF benchmark it is foresee to introduce case–based retrieval for 2009.

6. Acknowledgements

This work was supported by the Swiss National Science Foundation (FNS) with grant 200020–118638/1, the European Union 6th Framework program in the context of the KnowARC (contract 032691) research project, and the University of Applied Sciences Western Switzerland (HES SO) with the BeMeVIS project.

References

1. N.-S. Chang and K.-S. Fu. Query–by–pictorial–example. IEEE Transactions on Software Engineering, SE 6 No 6:519–524, 1980.

2. P. Clough, H. Müller, and M. Sanderson. The CLEF cross–language image retrieval track (ImageCLEF) 2004. In C. Peters, P. Clough, J. Gonzalo, M. Jones, Gareth J. F.and Kluck, and B. Magnini, editors, Multilingual Information Access for Text, Speech and Images: Result of the fifth CLEF evaluation campaign, volume 3491 of Lecture Notes in Computer Science (LNCS), pages 597–613, Bath, UK, 2005. Springer.

3. J. Dahmen, T. Theiner, D. Keysers, H. Ney, T. Lehmann, and B. Wein. Classification of radiographs in the 'image retrieval in medical applications' – system (irma)". In 6th International RIAO Conference on Content-Based Multimedia Information Access, pages 551–566, Paris, France, April 2000.

4. A. Depeursinge, J. Iavindrasana, G. Cohen, A. Platon, P.-A. Poletti, and H. Müller. Lung tissue classification in HRCT data integrating the clinical context. In 21th IEEE Symposium on Computer-Based Medical Systems (CBMS), Jyväskylä, Finland, June 2008.

5. A. Depeursinge, J. Iavindrasana, A. Hidki, G. Cohen, A. Geissbühler, A. Platon, P.-A. Poletti, and H. Müller. A classification framework for lung tissue categorization. In SPIE Medical Imaging, San Diego, CA, USA, February 2008.

6. A. Depeursinge, D. Sage, A. Hidki, A. Platon, P.-A. Poletti, M. Unser, and H. Müller. Lung tissue classification using Wavelet frames. In Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Lyon, France, August 2007.

7. T. Deselaers, A. Hegerath, D. Keysers, and H. Ney.

8. T. M. Deserno, S. Antani, and R. Long. Exploring access to scientific literature using content-based image retrieval. In SPIE Medical Imaging, volume 6516, San Diego, CA, USA, February 2007.

9. M. Ellert, M. Grønager, A. Konstantinov, B. Kónya, J. Lindemann, I. Livenson, J. Langgaard Nielsen, M. Niinimäki, O. Smirnova, and A. Wäänänen. Advanced resource connector middleware for lightweight computational Grids. Future Generation computer systems, 23(2):219–240, 2007.

10. W. Hersh, H. Müller, J. Jensen, J. Yang, P. Gorman, and P. Ruch. Advancing biomedical image retrieval: Development and analysis of a test collection. Journal of the American Medical Informatics Association, September/October:488–496, 2006.

11. J. Iavindrasana, A. Depeursinge, P. Ruch, S. Spahni, A. Geissbuhler, and H. Müller. Design of a decentralised reusable database architecture. In MedInfo 2007, volume 12 of IOS Press, Studies in Health Technology and Informatics, pages 325–329, Brisbane, Australia, 2007.

12. T. Kato. Database architecture for content–based image retrieval. In A. A. Jamberdino and W. Niblack, editors, Image Storage and Retrieval Systems, volume 1662 of SPIE Proceedings, pages 112–123, San Jose, California, February 1992.

13. 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), pages 882–886, Nashville, TN, USA, October 1998.

14. H. Müller, T. Deselaers, E. Kim, J. Kalpathy-Cramer, T. M. Deserno, P. Clough, and W. Hersh. Overview of the ImageCLEFmed 2007 medical retrieval and annotation tasks. In CLEF 2007 Proceedings, volume 5152 of Lecture Notes in Computer Science (LNCS), Budapest, Hungary, 2008. Springer.

15. H. Müller, P. A. Do Huang, A. Depeursinge, P. Hoffmeyer, R. Stern, C. Lovis, and A. Geissbuhler. Content–based image retrieval from a database of fracture images. In S. C. Horii and K. P. Andriole, editors, Medical Imaging 2007: PACS and Imaging Informatics, volume 6516 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, page 65160H, March 2007.

16. H. Müller, 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, 73:1–23, 2004.

17. H. Müller, M. Pitkanen, X. Zhou, A. Depeursinge, J. Iavindrasana, and A. Geissbuhler. Knowarc: Enabling Grid networks for the biomedical research community. In Healthgrid 2007, pages 261–268, Geneva, Switzerland, April 2007.

18. M. Niinimaki, X. Zhou, A. Depeursinge, A. Geissbuhler, and H. Müller. Building a community grid for medical image analysis inside a hospital, a case study. In Workshop on Grids for medical imaging at MICCAI 2008, September 2008.

19. A. F. Smeaton, W. Kraaij, and P. Over. TRECVID 2003: An overview. In Proceedings of the TRECVID 2003 conference, December 2003.

20. 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, 22 No 12:1349–1380, 2000.

21. D. M. Squire, W. Müller, H. Müller, 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), 21(13–14):1193–1198, 2000. B.K. Ersboll, P. Johansen, Eds.

22. H. D. Tagare, C. Jaffe, and J. Duncan. Medical image databases: A content–based retrieval approach. Journal of the American Medical Informatics Association, 4(3):184–198, 1997.

23. 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, 5:40–49, 1999.

24. C.-F. Tsai, K. McGarry, and J. Tait. Qualitative evaluation of automatic assignment of keywords to images. Information Processing and Management, 42(1):136–154, 2006.

25. E. M. Voorhees and D. Harmann. Overview of the ninth Text REtrieval Conference (TREC–9). In The Ninth Text Retrieval Conference, pages 1–13, Gaithersburg, MD, USA, November 2000.