Report on the development of an IR system for medical image documents

Sebastian Dungs
University of Duisburg-Essen
D-47048, Duisburg, Germany
dungs@is.inf.uni-due.de

René Donner
Technical University of Vienna
A-1090, Vienna, Austria
rene.donner@meduniwien.ac.at

Abstract

A prototype of a content based information retrieval system for clinical images is presented. It is targeted at the user group of radiologists working on diagnosing new cases amongst other scenarios. While still being in an early state, the system features sophisticated image retrieval mechanisms and a comprehensive and versatile user interface. CBIR is performed using several types of visual features aggregated over supervoxels and state-of-the art indexing regimes. The user interface uses an agent based framework infrastructure which is easily extensible and suited for complex tasks with difficult information needs [Beckers *et al.*, 2012b]. The system is one main part of the EU-project *Khresmoi* finishing in August 2014.

1 Introduction

The following paper describes the current stage of development of a information retrieval system for medical experts. It is tailored at radiologists and their special demands when it comes to diagnosing diseases by looking at medical images. These images may be taken through means of CT, PET, fMRI or X-ray and are accessible within a hospital network.

Diagnosis by visual analysis requires recognition of patterns and structures in the images that may be an indication of a specific condition the patient is in. Today radiologists often rely on text books as a reference for unknown visual structures. Asking more experienced colleagues is also often the only handy option for doctors in their first years of medical practice.

The system under development—being part of the larger Khresmoi project (see Section 3)—is aiming to overcome these problems. Users will be able to perform searches based on the images of the case at hand. The system performs an image similarity search and returns images containing the same structures as the query case. Moreover, the diagnosis is returned alongside. This is believed to give the experts a starting point in their diagnosis process. The overall goal is to speed up diagnosis of tough cases with rare or unknown diseases by reducing the need to consult external resources like textbooks or human advice.

This leads to interesting research questions, like how a useful and usable user interface is going to look like or which visual characteristics best distinguish certain diseases. An important question is also how the incorporation of both visual and semantic characteristics into the employed machine learning methods allows to improve the performance over simple visual retrieval, only.

This paper is structured as follows: Related work will be discussed in the next section. Section 3 introduces the Khresmoi project. In Section 4 the system is described from the user point of view. The interaction work-flow we envision users to follow is also presented. Furthermore, Section 5 covers the retrieval algorithm used for this prototype. Section 6 sums up and Section 7 discusses our work.

2 Related Work

Clinical image data in a hospital environment is currently typically organized in a *Picture archiving and communication system* (PACS). There are various different systems available distributed by large international companies. While these systems can handle large amounts of data generated in the hospital and are directly connected to the data sources they all have a major drawback. Content based image retrieval is not supported. Our proposed system tries to overcome this deficit.

User interface related developments described in this work are mainly based on the ezDL[1] software framework (compare [Beckers *et al.*, 2012a]). The authors describe an agent based retrieval system to access heterogeneous distributed digital libraries. While the original system is not used in the medical domain, it can be easily adapted and extended by adding further data sources. It also has built in logging functionality on user interaction level which can be activated if user experiments are about to be conducted. How this system was adapted to the Khresmoi project radiology use case is described in the following section.

3 Khresmoi Project

Khresmoi[2] is a project funded by the EU and currently is its third of four years. It aims at developing a multilingual, multi-modal IR system for biomedical information. It advances the current state of the art in several domains. These include the automated extraction of information from biomedical documents, semantic search features, linking information extracted from different sources, automated analysis and indexing of numerous medical images like X-rays as well as 3-dimensional data and supports cross-language IR. Moreover, a flexible user interface framework tailored at supporting a variety of different tasks and user interaction styles is under development. It is a versatile system supporting various platforms. This includes full featured desktop clients for PC, Mac and Linux,

[1] `http://www.ezdl.de`
[2] `http://www.khresmoi.eu`

a browser based version for high flexibility and an Android app for mobile use. All clients share a common brand identity allowing for an easy transition between them.

The project has different use cases each targeted at a special user group. All user groups have different needs which are also addressed by the aforementioned variety of user interface versions, each featuring different tools suited for the most common tasks of the user group. One of these user groups in the field of medical professionals consists of radiologists. A adapted version of the desktop client interface (see Figure 1) is under development. A prototype of this interface as a well as the underlying retrieval mechanism—also referred to as *Khresmoi for radiologists*—will be explained in the following section.

4 Work-flow description

As mentioned in Section 1 one distinct use case for the Khresmoi project is the radiology department in a hospital environment. There it will be used in the process of writing diagnoses based on visual analytics of images of various modalities. Therefore, Khresmoi for radiologists will be connected to a hospital database like a PACS system. This database also provides the basis for case retrieval. The work-flow we envision is as follows:

First a radiologist selects the case in question by means of a tabular *index perspective*. It gives basic information about all cases, like imaging technique used or patient demographics. The table can be sorted and filtered to allow a quick search. Users can request additional data about any case by clicking on it. This will load the report about that case, if any exists, as well as the actual image data. Because the system will be deployed within the hospital network, the transfer time of these potentially large files is expected to remain reasonable. If a user decides to query the system with a specific case she is able to narrow down the search by marking one or more regions in the image data as relevant. How this information is used in the retrieval process is explained in Section 5. In the prototype's current state of development marking regions of interest is based on rectangular areas that can be selected in the two dimensional images.

Results will be presented in a list of result items. Every item consists of a representative preview thumbnail of the image as well as basic meta data. These include patient age and gender, the image acquisition date and a short version of the report along with pathology tags. The default sorting is relevance based, while the results can also be sorted by acquisition date in both orders. Users are able to group the result set according to case meta data values. For instance one might be interested in cases that correspond to a certain disease. All matching cases can be highlighted within the result set or it can be filtered showing only matching cases. Radiologists can retrieve the actual image data and the report for any case by clicking on a result item in the list. It will then be shown in the central part of the interface or in a separate window. All cases can be stored in a tray or basket and will on demand be saved permanently for easy access in the future.

Further information on the actual retrieval process and the novel technological aspects our system are described in the following section.

5 CBIR

The content based image retrieval part of the described system is developed and located at the Medical University Vienna, with the CIR lab. The system includes the capability for the data management and storage of very large medical volume datasets and employs state of the art computer vision techniques to analyze and index the data. The indexed data as well and the indexing service is made available to the Khresmoi framework through a private API. The data set used in the index is a collection of 3876 3D-CTs and MRs extracted from the PACS of the General Hospital Vienna / MUW. The CTs and MRs originate from all different scanners present at the department of radiology.

Processing data flow The data is transferred from the hospital's PACS to an internal data base system, after detailed anonymization of all data and meta-data. Subsequently each volume is processed, its visual features computed, and these features are added to the index. The entire data analysis is formulated as a map-reduce graph, wherein each node in the graph can store its results on disc. The inter-dependencies of the nodes are automatically exploited such that adding a new volume to the data store only triggers the computation of the minimally required subset of nodes in the graph to ensure a valid index. This approach also implicitly provides the ability to run the computations in parallel on a compute cluster, and implicit robustness to errors in the node's computations or machine failures.

Computer vision methodology The main components of the computer vision processing pipeline include the correct orientation of the volume in regards to a reference atlas, the registration to this atlas, the computation of super-voxels and several visual features per super-voxel and finally the computation of the index. The orientation of the volumes as delivered from the PACS can be arbitrary, but is defined by the volume's headers. A simple but important aligning step ensures that the volume has the same orientation as the atlas. The atlas itself consists of one whole-body CT scan. After the orientation, alignment and affine registration is performed between the volume and the atlas. This estimates the translation, rotation and scaling parameters necessary to best align the volume to the atlas. Subsequently, a non-rigid registration estimates the complex non-linear deformation necessary to obtain a correct voxel-by-voxel alignment between the volume and the atlas. The voxels of the aligned volume are divided into non-overlapping parts, i.e. so-called super-voxels. This commonly employed computer vision technique aims to extract regions which are maximally homogeneous within each region and maximally distinct between adjacent regions. The number of super-voxels is set to be three to four orders of magnitude lower than the number of voxels, greatly speeding up the feature computation and indexing, while loosing very little information which would be relevant to the task of retrieving similar regions in the volumes. We employ an adapted variant of SLIC super-voxels incorporating the monogenic signal, which allow for smoother and more regular super-voxels. For each super-voxel a set of well-established visual features is computed, namely gray-level co-occurrence matrices with Haralick-features (as for example employed in [Valentinitsch *et al.*, 2013]) and Haar-like wavelets as described in [Donner *et al.*, 2010], as well as Bags-of-Visual words of Local Binary Patters (LPB) and gray-level histograms. The framework is not limited to these visual features, and the best set of features for each retrieval scenario can be found through a cross-validation

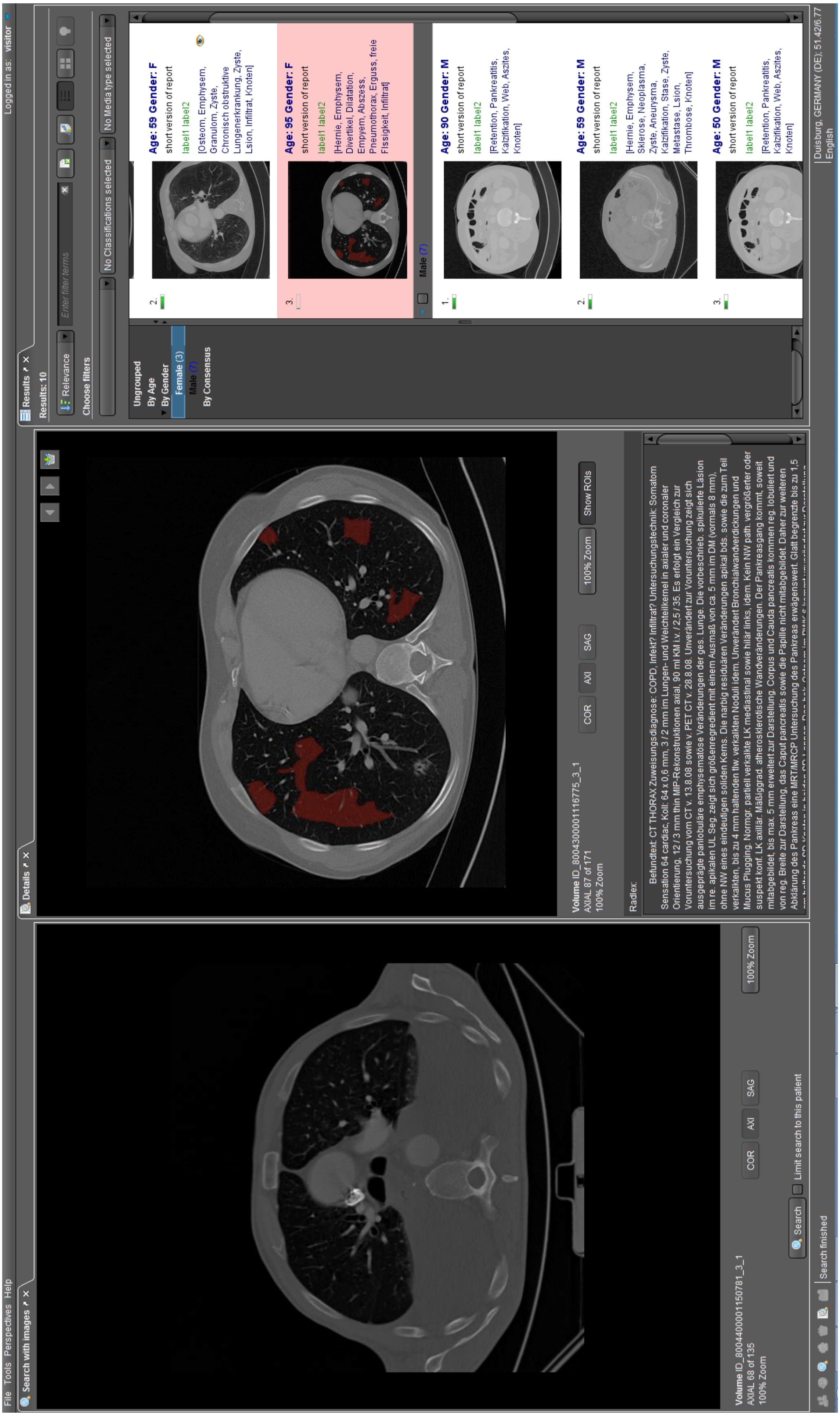

Figure 1: Screen shot of the current user interface prototype

approach. The evaluation of different visual features is the current focus of our research.

These features are concatenated, yielding a $nFeatures \times nSupervoxels$ matrix per volume. The actual indexing of these features across the data set is performed using different methods, which are currently under evaluation. ProductQuantizers [Jégou *et al.*, 2011] are used to quickly retrieve the most similar super-voxels, given a query super-voxel, in the nFeatures-dimensional feature space. The evaluation of such a system is performed on two fronts: one is concerned with measuring the effect and improvement of using this tool in clinical practice. This is mainly assessed through controlled experiments with medical experts using the system. A detailed analysis of the usage, supported by video monitoring and eye-tracking allows to measure improvements in GUI design and overall retrieval performance. Detailed interviews with the study participants are also employed in each evaluation round. The second metric is the numerical evaluation of the details of the retrieval pipeline using a large set of cases with existing diagnostic reports as ground truth. An automated semantic analysis of this corpus yields a distance metric between the cases, and the performance of the retrieval system can be evaluated against it.

6 Conclusion

We presented a system for content based image retrieval planned to be deployed in a hospital environment. Current systems like the common PACS work-stations do not support the retrieval of cases by means of image data. This leads to an inefficient and overly complex work-flow for radiologists while diagnosing cases. Our system can be used to quickly find similar cases to the one at hand without the need to rely on external sources. Therefore, data is at first anonymized and transferred to our system, where it is mapped and aligned to an atlas. By using super-voxels for computation the process is sped up significantly. The volumes are then indexed by their visuals features over super-voxels, using map-reduce graphs to structure the computational data-flow. The actual indexing can be performed using different methods. These as well as the user interface will be further evaluated in the future.

7 Future development

Also in the near future the system's prototype will be extended by additional data sources. It will include 2D image documents taken from the Pubmed[3] database as well as textual documents. This data is provided within the Khresmoi project and part of the other use case prototypes. The radiology system will benefit from this addition because users can access more information relating to the current case. Based on the image retrieval result the system will extract textual keywords which will be used for an initial query to the 2D and textual document sources. Result are presented directly in the user interface without any user interaction while making sure the work-flow is not disrupted. The user may then choose to alter the automatically generated query and reissue it while the initial 3D query and result set remain unmodified. We think this system will improve the accessibility of medical resources needed for diagnosing difficult cases as well as reduce the time between the taking of an image and the making of a diagnosis. Nevertheless, the system will have to undergo user evaluations in the future to test both the user interface and the quality of the actual retrieval. Those evaluations are foreseen to take place in the forth project year.

[3] `http://www.ncbi.nlm.nih.gov/pubmed`

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n 257528 (KHRESMOI).

References

[Beckers *et al.*, 2012a] Thomas Beckers, Sebastian Dungs, Norbert Fuhr, Matthias Jordan, and Sascha Kriewel. ezdl: An interactive search and evaluation system. In *SIGIR 2012 Workshop on Open Source Information Retrieval (OSIR 2012)*, August 2012.

[Beckers *et al.*, 2012b] Thomas Beckers, Sebastian Dungs, Norbert Fuhr, Matthias Jordan, Sascha Kriewel, and Vu Tran. ezdl: A framework for interactive and professional search systems. Technical report, Universität Duisburg-Essen, Thessaloniki, Greece, 2012. Position Paper at the MUMIA Workshop about Integrating IR/MT Technologies for Professional Search.

[Donner *et al.*, 2010] René Donner, Erich Birngruber, Helmut Steiner, Horst Bischof, and Georg Langs. Localization of 3D Anatomical Structures Using Random Forests and Discrete Optimization. In *Proc. MICCAI Workshop on Medical Computer Vision*, pages 86–95, 2010.

[Jégou *et al.*, 2011] Hervé Jégou, Romain Tavenard, Matthijs Douze, and Laurent Amsaleg. Searching in one billion vectors: re-rank with source coding. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 861–864. IEEE, 2011.

[Valentinitsch *et al.*, 2013] Alexander Valentinitsch, Janina M Patsch, Andrew J Burghardt, Thomas M Link, Sharmila Majumdar, Lukas Fischer, Claudia Schuellerweidekamm, Heinrich Resch, Franz Kainberger, and Georg Langs. Computational identification and quantification of trabecular microarchitecture classes by 3-d texture analysis-based clustering. *Bone*, 54(1):133–140, 2013.