Integration of a Research CBIR System with RIS and PACS for Radiological Routine

Benedikt Fischer[*a], Thomas M. Deserno[b], Bastian Ott[c], Rolf W. Günther[c]
[a]Dept. of Medical Informatics, Health Information Systems Lab,
[b]Dept. of Medical Informatics, Image Processing Lab,
[c]Dept. of Diagnostic Radiology,
[a,b,c]RWTH Aachen University of Technology, Pauwelsstr. 30, D-52057 Aachen, Germany

ABSTRACT

In this work, a concept for coupling a system for content-based image retrieval in medical applications (IRMA) with hospital information systems is presented. We aim at improving the work flow of radiologists and evaluating the recognition performance of the IRMA system in clinical routine. The integration is designed such that a failure of IRMA does not affect the routine operation of the other systems. The coupling is realized by generic communication modules with the radiology information system, and the picture archiving and communication system (PACS) over the standard protocols Digital Imaging and Communications in Medicine (DICOM) and Health Layer 7 (HL7). An optional plug-in for the radiological viewing station further enhances the usability. Based on this concept, the pre-fetching of relevant images for recurrent examinations is improved. When an examination is scheduled, all previous images of the patient are read by the IRMA system with DICOM query/retrieve. If the images were not present before in our database, features are extracted, stored, and indexed. After the acquisition of new images from the imaging modality, the new images are automatically retrieved by the IRMA system with DICOM query/retrieve and similar images are selected based on the stored global signatures. These images are then loaded into the online storage of the PACS and are available for diagnostic purposes together with those images already pre-selected by the PACS. Thus the radiologist can avoid further delays resulting from manually fetching further images from archives which have not been automatically selected by alphanumerical meta data. In addition, he is able to sort all fetched images by the computed IRMA-similarity. Furthermore, the hanging of images in the viewing software is planned to be organized by IRMA suggestions automatically, further shortening the time for the examination and reducing manual interactions. Based on the generality of our integration concept, a CBIR-based second opinion to support the diagnostics, and computer-based training of radiologists will be established in near future.

Keywords: PACS, RIS, HIS, Interface, Data Exchange, CBIR, SOA, HL7, DICOM

1. INTRODUCTION

Digital imaging is widely used in medical practice for diagnostic purposes. The large amounts of data emerging from imaging processes have to be managed and organized to allow an efficient retrieval[1]. For this purpose, several levels of storage are used in PACS. The most recent images are kept in an online-storage, allowing immediate access. Depending on scheduling strategies, the images are then moved first to slower intermediate storage systems (e.g. DVD juke boxes) until they are archived in long-term storage systems (e.g. tape racks) requiring the largest access time.

At present, only meta information like body part labels and patient data is used for image description and retrieval in medical routine. The image content itself is not analyzed. However, the stored meta data is frequently incorrect leading to less than optimal retrieval results[2]. A possible solution exists in content-based image retrieval (CBIR) systems which analyze image similarity regardless of annotations by analyzing the image content itself[3]. CBIR assists the physician in the finding process by selecting images which are relevant for the interpretation[4] and ensure that these are available in the online-storage. During the past years, the development concerning CBIR systems and in particular medical CBIR

[*]Corresponding author: Benedikt Fischer, Department of Medical Informatics, Aachen University of Technology (RWTH), Pauwelsstr. 30, D-52057 Aachen, Germany, email: bfischer@mi.rwth-aachen.de; phone +49 241 80 85174; fax +49 241 80 33 85174; web: http://irma-project.org

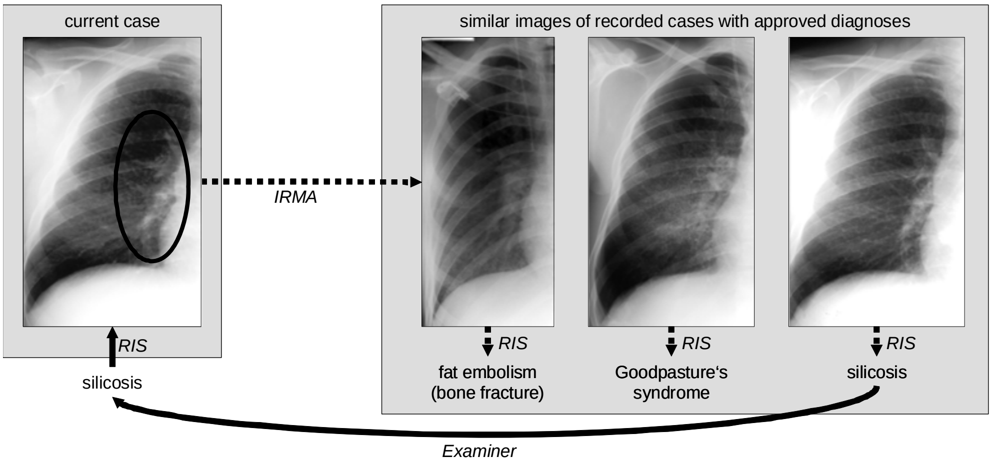

Figure 1. Coupling of IRMA with RIS and PACS: The IRMA system retrieves images similar to the image under analysis from the PACS. The RIS then provides the corresponding findings to support the examiner. The examiner then records his diagnosis via RIS.

systems has grown considerably[5,6]. Nowadays international evaluation contests have emerged to compare different strategies on diverse image sets[1,7]. Still missing, however, is the integration of the research systems into daily medical routine.

In this work, a concept for the integration of the content-based image retrieval in medical applications (IRMA)[8] system into the radiological routine at the University Hospital, Aachen, Germany, is presented. IRMA has been developed at the Department of Medical Informatics, RWTH Aachen University, and provides a framework for image retrieval of general medical image data. It has been extensively tested with thousands of medical images acquired from clinical routine[9,10], and has even been transferred to other fields[11,12]. Until now, the application of the framework took place outside of the diagnostic workflow under laboratory conditions, such that the practical use for the daily work of medical experts remained as theory.

The proposed integration concept has been designed as a service oriented architecture (SOA) making use of common communication protocols to ensure applicability to different information systems. Vendor-specific plug-ins, e.g. to the viewing software are optional and only used for ease of use and to increase the user acceptance. The general setup is transferable to arbitrary system constellations.

2. PURPOSE

The long-term goal of this work is to make the complete functionality of IRMA accessible from the workstations of the radiologists and to employ CBIR technology for the improvement of physicians' work flow. To accomplish this, the integration into medical routine is planned in pilot applications with increasing complexity.

- **Low-Level Applications:** The complexity of the image retrieval part is kept at a low level, with robust solutions which have proven successful for similar tasks in "offline" application. This allows to set the focus on the rather technical requirements needed also for all later stages in the integration process. This includes the definition of interfaces and the communication between the IRMA system for content-based image retrieval and the other information systems in the hospital and the implementation of the corresponding software modules, such as information brokers, schedulers, and user interfaces.
- **Mid-Level Applications**: While in the low level applications, it is sufficient to restrict the image analysis to rather generic global attributes describing the complete image[6], the target applications at this level involve the analysis of local image information, introducing more complex methods, usually specialized for a given application. For example, when case-based reasoning is applied in radiological diagnostics, this requires access to images showing a comparable (pathological) pattern, as well as the clinical finding (Fig. 1). A first introduction to routine should

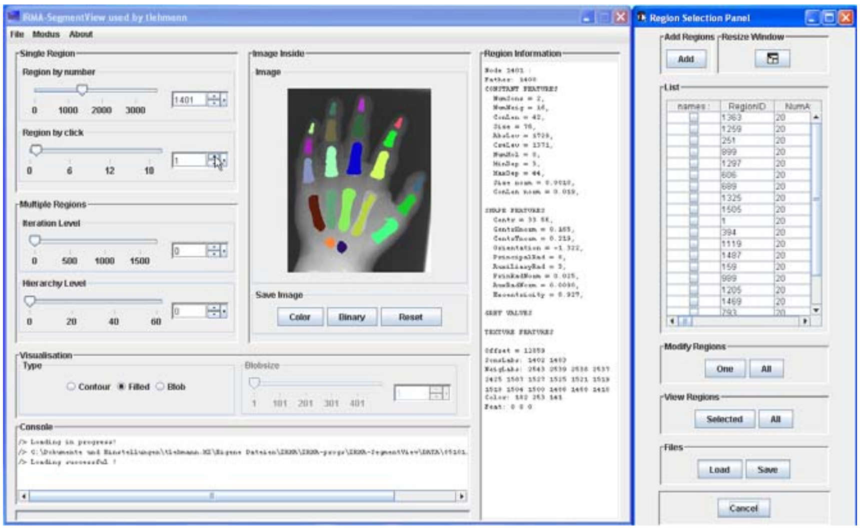

Figure 2. High complexity application: The single hand bones can be displayed for verification after having been identified automatically by prototype graph matching[13]. Afterwards, the computation of bone-age can run in the background, e.g. by the TW3-method[15].

reflect how frequent such an image comparison query is desired by the physicians and how satisfactory the results of the IRMA retrieval system are. In later phases of the integration process, the examiner will select his/her specific region of interest (ROI) in the image, and application-specific feature sets and pattern recognition methods will be made available, e.g. for the detection and classification of lung nodules in chest x-ray (Fig. 1).

- **High-Level Applications:** In mid-level, local ROIs are analyzed. In the approached support for applications of higher complexity, *multiple* single regions of interest are analyzed, e.g. hand bones, and their constellation is compared to a *prototypical structure* or scene inside the image[13]. Imaging systems for computer-aided diagnosis frequently fall into this category. They are developed for specific medical tasks and extract reproducible parameters from standardized image material to support the physician in the diagnostic process and/or to avoid tiring manual procedures. As an example, the overview radiographs of the left hand are analyzed for the determination of bone-age. Single bones of the carpus develop individually and regarded together with the ossification, can be analyzed by predefined schemes[14]. This examination takes around 20 minutes for an experienced radiologist. Hence, integration of IRMA aims at supporting the bone maturity calculation. For the examiner, this means, that after a short verification of the bone identification (Fig. 2), the computation of the bone-age can be done in the background without further manual interactions. The result is then displayed suggesting a second opinion to the medical expert[15].
- **Multi-Level Applications**: The applications at this level make simultaneous use of applications on all previous levels. By means of continuous education of radiological experts, computer-based training offers an excellent opportunity to ensure a high quality of education. Concerning the above mechanisms for image interpretation and the connection to the RIS with diagnostic information, interactive training units can be easily assembled. These can be arranged by specialization, anatomy, pathology or displayed unsorted to the person in training. The HL7-interface to the RIS can provide the (anonymized) corresponding findings and correlate them to the diagnosis of the trainee. Besides the general applicability this may be especially helpful for the training of findings with similar images with distinct diagnoses.

2.1 Support of Pre-Fetching

While the long-term goals have been defined above, the following sections focus on the first step of the overall integration scheme. The purpose therefore is defined as the establishment of the generic service modules allowing the communication between the IRMA framework and the PACS, RIS, and HIS as well as the combination of these services in a pilot application, the support of pre-fetching and hanging that can both be solved on global signatures.

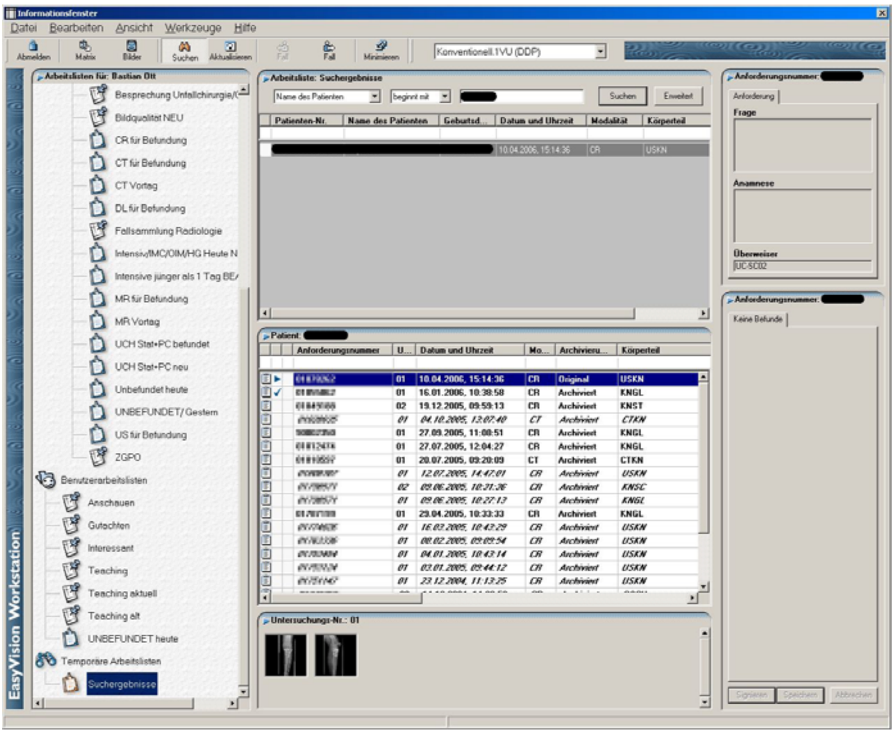

Figure 3. Radiological information center for selecting examinations to be compared. Images contained in the online-storage (straight font) can be displayed immediately. All images shown in italics are stored in an archive and have to be fetched. By automatic pre-fetching (not shown) only images annotated with identical body region-code would be available without further delay.

In recurrent radiological examinations, previous images of the same patient are often consulted for comparison to assess the treatment progress. For scheduled examinations, earlier cases of the patient are retrieved from the information systems beforehand, for instance, during the night before the scheduled date. Due to processing time and memory space limitations, not all previous data of a patient will be retrieved. Instead, existing images in the PACS are first considered for relevance to the current case based on alphanumerical patient data and examination code. If the examination codes of the current and previous cases match, the PACS can easily deliver the corresponding images to display the development over time (Fig. 3).

Frequently, images from different categories must be downloaded manually by the radiologist, mainly as a result of inconsistent labeling. Images of the knee, for example, could be coded in the Sectra PACS as

- USKN (lower leg with knee joint),
- KNGL (knee joint, left),
- KNST (knee joint, standing),
- CTKN (CT knee joint),
- USGA (lower leg, overview), or
- KNSC (kneecap, patella).

Even if these codes may be appropriate for the distinct case, they would not very likely result in a continuous coding and thus lead to incomplete information, requiring manual and time-consuming interaction with the long-term archive during the finding process. With the help of content-based methods for image access, additional images can be selected for pre-fetching, based on their visual similarity. Thus, the radiologist does not have to wait while the PACS is downloading missing images, and the examination can be performed significantly faster.

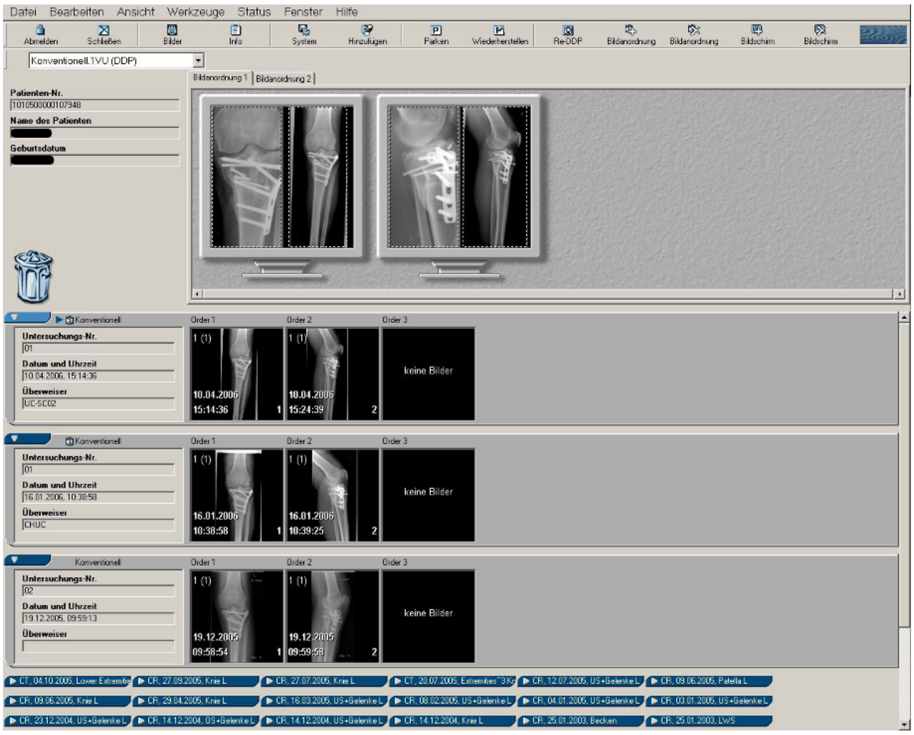

Figure 4. Example layout for planned examinations: The two monitors featured on top can be set-up with user-defined layouts. Shown are two examinations (one on each monitor) in split-screen mode with an overview and a zoomed view. The user first chooses the layout and can then “hang” the images on the monitors.

2.2 Supporting the automatic hanging

As a further application, we want to present the radiologist layout suggestions for the radiograph hanging. Usually, radiologists have established best practices for image grouping on the displays, e.g. image from the current examination is on the left, image from last examination of the same region on the right (Fig. 4). Even if all relevant images are retrieved, their similarity to a new case are at first unknown and the radiologist needs to pick the ones showing the corresponding region from the list of all images. By the content analysis of the IRMA system, the similarity can be computed and ranked to offer the most similar pairs first. By providing the radiologist with such layout suggestions, the amount of time used for manual grouping of the images can be greatly reduced. However, this step requires a tighter coupling of IRMA with the viewing station software.

3. METHODOLOGY

The coupling is implemented as a new module of the IRMA system that interacts with the database and the hospital information system over standard protocols such as digital imaging and communications in medicine (DICOM), health level 7 (HL7) and Hypertext Transfer Protocol Secure (HTTPS).

An important prerequisite of the integration is that the routine operation must not be affected by errors of the IRMA system, which is used solely as a source of additional information. If the provided information is faulty, it can be safely ignored by the radiologist.

The system integration can be subdivided into four levels[16,17]:

1. **data integration**, which requires that data does not need to be entered more than once;
2. **functional integration**, which is ensured when services provided by a module can be used where they are needed;

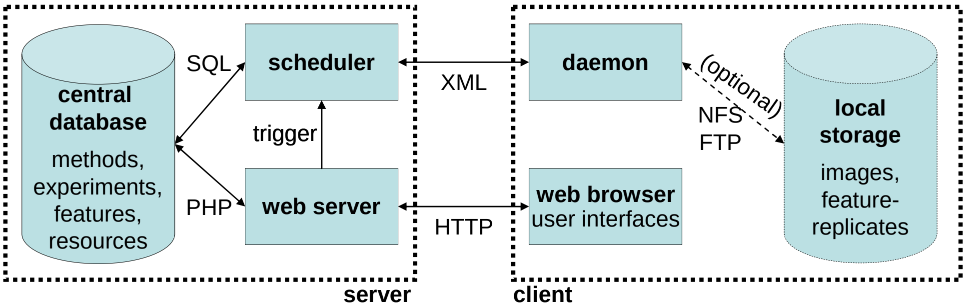

Figure 5. Communication within the IRMA system, split into server- and client-based components.

3. **presentation integration**, which is given when different modules present their data in a unified way; and
4. **context integration**, meaning that settings such as the selection of a certain patient or image, which are done in one module, are passed automatically to another module when this is called.

In our concept, we aim to achieve best possible integration according to all of the above criteria. The IRMA system itself is organized in server- and client-based modules (Fig. 5). On the server side, a central database is used to host methods (algorithms), experiments (algorithms and parameter settings), features (results of computations, images), and resources, e.g. computers or storage devices. It is accessed by a scheduler and a web server that can trigger the scheduler to start computations. On the client side, daemons await requests from the scheduler to perform computations, and a web browser is used to provide the user interfaces for all IRMA applications, which therefore can be accessed simply over the internet[18].

3.1 Communication with the Radiology Information System

The RIS is used for report management and is informed by the hospital information system (HIS) about the admission of new examination dates. When a new examination is scheduled, an HL7 order message (ORM) is sent to the PACS broker. The same message is then dispatched to the IRMA system. The IRMA system is then able to start the necessary preparations for the targeted application, e.g. the pre-fetching. No additional communication between the RIS and the IRMA system is required.

3.2 Communication with the Picture Archives

The communication with PACS is realized with the help of DICOM query/retrieve service and SQL. When an examination is scheduled, all previous images of the patient are read by IRMA with DICOM query/retrieve. If the images have not been present in the IRMA database, their features are extracted and stored.

On the scheduled date, currently acquired images are loaded into the IRMA system with DICOM query/retrieve and their features are extracted. On account of the features, similar images can be selected which are then loaded into the online storage of the PACS and are available for diagnostic purposes together with the images selected by the PACS.

As a further enhancement, the layout suggestions are communicated to the PACS broker via SQL and saved into fields reserved for custom extensions. These may then be used by the viewing station software.

3.3 Communication with the Viewing-Station Software

At the University Hospital of RWTH Aachen University, EasyVision from Philips/Sectra is used as viewing station software. It provides a proprietary application programming interface (API), which is called clinical applications interface (CAI™) and allows the integration of user-defined applications. Therefore, all IRMA applications can be integrated via a corresponding start-button. The IRMA system will then receive the current context, e.g. patient id, image, and the IRMA application to be started. This allows a minimum part of vendor-specific adaptation, guaranteeing a sufficient level of context integration and a maximum portability to other RIS/PACS systems. The feedback of the IRMA core is opened automatically in a separate browser window. This holds for both: display of CBIR results and interactive applications of query refinement and relevance feedback[18]. Any final decision is made by the medical expert and not by computer software, but the software will provide valuable support for decision making.

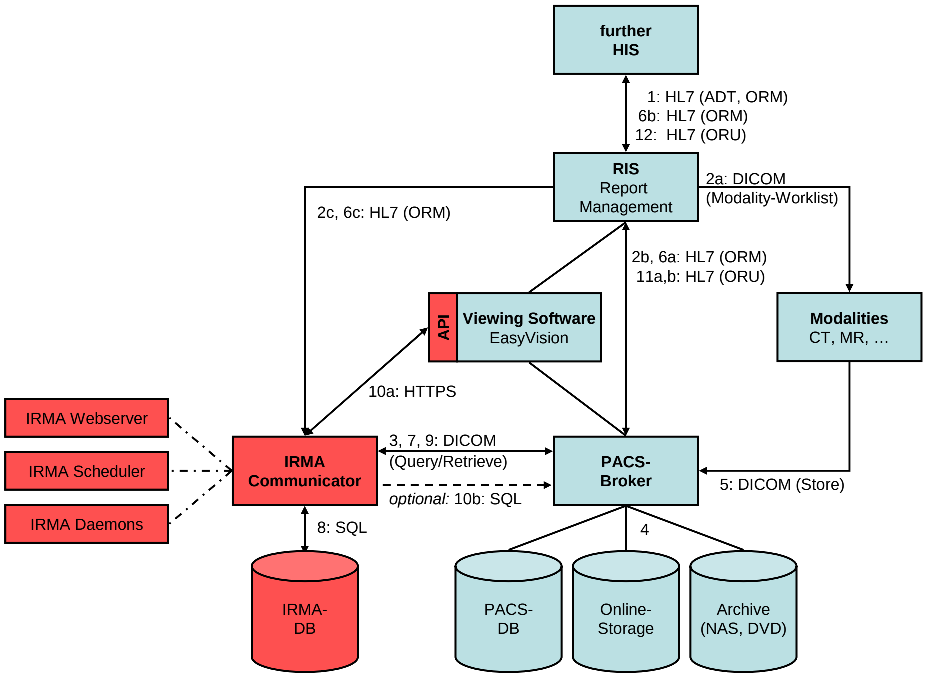

Figure 6. Components and information flow during radiological analysis. Blue: established system for pre-fetching. Red: Required add-ons for the pre-fetching support. The dotted lines at the left indicate internal communications within the IRMA system (see Fig. 5).

4. RESULTS

In this section, we present the implementation of our concept for improving the quality of pre-fetching. For the radiologist, it is necessary to have all relevant preliminary examinations from the PACS archives available for the analysis of the current one. The first task is the pre-fetching of all relevant images by examination code or visual similarity. The second task is the hanging of the images on the diagnostic screens.

4.1 Task 1: Pre-Fetching Support

The information flow during a medical follow-up examination is depicted in Figure 6. Already existent components are visualized blue (light grey), new components necessary for IRMA integration are shown in red (dark grey). The information flow can be grouped by three distinctive points in time:

1. Fixing of the examination date
2. Night before the examination
3. Date of examination

In the following, we describe the interactions at each of these dates. Transactions necessary for the IRMA integration are denoted in *italic*. The communication is numbered according to their execution time. Parallel communication is denoted by adding letters, see for example, 2a, 2b, 2c in the next paragraph.

Fixing of Examination Date

1: The RIS receives an ADT[i] message about patient admission and an examination request (ORM[ii]) from the HIS via the HL7 protocol.

---

[i] ADT: HL7-message for admission, discharge or transfer of a patient

2a: The RIS sends a DICOM modality worklist to the modality which shall be used for the examination. Thus it is ensured that the required context information is available at the examination date.

2b: Simultaneously, the HL7 order message is forwarded to the PACS broker.

2c: This HL7 order message about the examination request is also sent to the IRMA system.

3: IRMA requests all available images of the patient via DICOM query/retrieve from the PACS broker and computes global signatures, which are stored in the IRMA database.

Night before Examination

4: In the night before the examination, the PACS loads images from all preceding examinations with identical body part code into the online storage.

Date of Examination

5: The new images are transmitted from the modality to the PACS via DICOM store.

6a: The PACS Broker informs the RIS about the availability of new images by the means of a status change message via HL7 (ORM).

6b: The RIS forwards this status change to the rest of the HIS.

6c: In addition, the same status change message is passed to the IRMA communicator.

7: The new images are retrieved from the PACS by the IRMA system via DICOM query/retrieve.

8: Subsequently, the signatures for the new images are computed and stored in the database.

9: Based on this signature and the similarity to all signatures already stored in the database, IRMA creates a list of images visually similar to the new and loads them into the online storage via DICOM query/retrieve. Now instant access is granted to images selected by PACS based on the body part code as well as to visually similar images selected by the IRMA system and manual time-consuming loading of non-available images is avoided.

10a: The list is also provided by the IRMA system via a standard webbrowser. It can be retrieved by the physician from every analysis workstation.

11: The radiologist is notified that the task is ready for diagnosis. He chooses a pre-defined layout for the hanging and selects the images to be displayed. After the thorough analysis by the radiologist, the results are transmitted via a HL7 ORU to the RIS.

12: The RIS forwards the HL7 result message to the rest of the HIS.

4.2 Task 2: Hanging Support

The flow of information here is almost identical to the one in specified in the above task. For a maximum integration on all four levels (data, function, presentation, and context) the viewing software is slightly modified via CAI™. The flow of information is concerned only in two steps:

10b: The similarity list can be passed vendor-specific to the PACS broker by SQL, which stores it in database fields reserved for such future applications and allows the easy access for the viewing station software. This can additionally include ROI per image to be used for zooming.

11b: The viewing software checks if a similarity list is available, and if present, automatically selects an appropriate hanging. This is supportive in two ways:

- *The list of available images (bottom right in Figures 3 and 4) does not necessarily have to be sorted by time anymore, but can also be sorted either only by visual similarity, or by time with a sub-ordering by visual similarity.*

[ii] ORM: HL7-message for placing an order

- *For a selected new image, the IRMA system can automatically suggest a hanging with the visually most similar image, for example offering a similar hanging as depicted in Figure 4 with a new image along with zoomed region of interest on the first monitor and a previous image with the corresponding zoom on the second one.*

After the thorough analysis by the radiologist, the results are transmitted via a HL7 ORU to the RIS.

5. DISCUSSION

The presented concept allows us to include CBIR functionality into radiological daily practice without disturbing the routine operations. The required communication modules and interfaces between all involved information systems have been identified and can serve from applications of low CBIR complexity up to high or even multi-level applications. Target applications are outlined for each level of complexity, and the implementation of a first pilot application, the support of pre-fetching, is described in more detail, split into two tasks: The primary task and goal is the pre-fetching of all relevant images, involving all communication modules also needed for the next complexity level. The second task is the automatic hanging of images to further unburden the radiologist from manual interactions.

Specific hanging layouts will have to be discussed with the radiological experts according to their current preferences for the different types of images and applications. A trade-off will also have to be found on the allowed age of the most similar image.

Besides the complexity levels, all four levels of integration are taken into account by the concept: *Data integration* is ensured in our context, as we have access to all available patient images and the system is notified of new examinations via HL7 ORM. *Functional integration* is ensured for the application of pre-fetching. If vendor-specific application interfaces are used, even a full function integration is available, but the concept does not rely on this. *Presentation integration* is given for the task of pre-fetching, as no direct interaction between the user and the IRMA system is required. If the web interface of our system is used, presentation integration is still provided, as the IRMA web interfaces are build from modules that recognize common standards of usability[18]. *Context integration* is given for the application of pre-fetching, as the (anonymous) patient data is available to the IRMA system. If the IRMA web interface is used, the context integration is ensured through the transmission of the patient id using the proprietary API of the viewing station software.

6. CONCLUSIONS

We have presented a concept for integration of a research CBIR system into the radiological routine based on standard protocols. Main prerequisite was that the routine operations are not disturbed by our additional components. The concept complies with this requirement, as the only write transaction to the PACS system is the additional load of relevant images into the online storage. If the IRMA system should fail, the routine operations would not be affected. The concept comprises a step-by-step approach for the integration of CBIR tasks of increasing complexity and offers integration concerning data, functionality, presentation, and context. The support of pre-fetching as a pilot application is presented in detail, promising early success by coupling proven low-level complexity CBIR methods with the established information flow in daily clinical routine. With this work, the foundation is laid for future work which will include applications in computer-based teaching, computer-aided diagnosis and evidence-based medicine[5].

7. ACKNOWLDEGEMENTS

The IRMA project and the related project Structural Prototypes in Radiological Routines (SPIRR) are funded by the German Research Foundation (Deutsche Forschungsgemeinschaft DFG) grants: Le 1108/4-1, Le 1108/4-2, -3, Le 1108/6, and Le 1008/9. We also acknowledge the contributions of Ilja Bezrukov in the integration process. We are grateful to Dr. Andreas Keizers and Frank Vogelsang from Sectra GmbH, Aachen, Germany, as valuable cooperation partners.

8. REFERENCES

[1] Müller H, Deselaers T, Deserno TM, Clough P, Kim E, Hersch W. Overview of the ImageCLEFmed 2006 medical retrieval and annotation tasks. Lecture Notes in Computer Science 2007; 4730: 595-608.

[2] Güld MO, Kohnen M, Keysers D, Schubert H, Wein BB, Bredno J, Lehmann TM: Quality of DICOM header information for image categorization. Proceedings SPIE 2002; 4685: 280-287.

[3] van Bemmel JH. Medical informatics – Art or science? Methods of Information in Medicine. 1996; 35 (3): 157-72.

[4] Smeulders AWM, Worring M, Santini S, Gupta A, Jain R. Content-based image retrieval at the end of the early years. IEEE Trans Pattern Anal Mach Intell 2000; 22(12): 1349-80.

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

[6] Deserno TM, Antani S, Long RL: Ontology of gaps in content-based image retrieval. *Journal of Digital Imaging* 2007, accepted.

[7] Deselaers T, Müller H, Clough P, Ney H, Lehmann TM. The CLEF 2005 automatic medical image annotation task. International Journal of Computer Vision 2007; 74(1): 51-58.

[8] Lehmann TM, Güld MO, Thies C, Fischer B, Spitzer K, Keysers D, et al. Content-based image retrieval in medical applications. Methods of Information in Medicine. 2004; 43 (4): 354-361.

[9] Güld MO, Deserno TM. Baseline Results for the CLEF 2007 Medical Automatic Annotation Task. In: Working notes of the CLEF 2007 Workshop, Budapest, Hungary, September, 2007.

[10] Lehmann TM, Molander B, Güld MO, Thies C, Gröndahl HG. Content-based access to oral and maxillofacial radiographs. Dentomaxillofacial Radiology 2007; 36(6): 328-35.

[11] Deserno TM, Antani S, Long LR. Exploring access to scientific literature using content-based image retrieval. Proceedings SPIE 2007; 6516: OL1-OL8.

[12] Müller KM, Fischer B. Objective condition assessment of sewer systems. Proceedings LESAM 2007, Portugal. In press.

[13] Fischer B, Sauren M, Güld MO, Bezrukov I, Deserno TM. Matching of structural prototypes for content-based image retrieval in medicine. Proceedings SPIE 2008; 6914. In press.

[14] Vogelsang F, Kohnen M, Schneider H, Weiler F, Kilbinger MW, Wein BB, Günther RW: Skeletal maturity determination from hand radiograph by model based analysis. Proceedings SPIE 2000; 3979: 294-305.

[15] Aja-Fernandez S, de Luis-Garcia R, Martin-Fernandez MA, Alberola-Lopez C. A computational TW3 classifier for skeletal maturity assessment: A Computing with Words approach. Journal of Biomedical Informatics 2004;37(2):99-107.

[16] Leisch E, Sartzetakis S, Tsiknakis M, Orphanoudakis SC. A framework for the integration of distributed autonomous healthcare information systems. Medical Informatics, Special Issue. 1997; 22 (4): 325-335.

[17] Winter A, Ammenwerth E, Brigl B, Haux R. Krankenhausinformationssysteme. Handbuch der Medizinischen Informatik. München: Hanser. 2005; 552-625

[18] Deserno TM, Güld MO, Plodowski B, Spitzer K, Wein BB, Schubert H, Ney H, Seidl T. Extended query refinement for medical image retrieval. Journal of Digital Imaging 2007; online-first, DOI 10.1007/s10278-007-9037-4.