A Generic Framework for Semantic Medical Image Retrieval

Manuel Möller, Michael Sintek

manuel.moeller@dfki.de, michael.sintek@dfki.de
German Research Center for Artificial Intelligence (DFKI) GmbH
Kaiserslautern, Germany

Abstract. Performing simple keyword-based search has long been the only way to access information. But for a truly comprehensive search on multimedia data, this approach is no longer sufficient. Therefore semantic annotation is a key concern for an improvement of the relevance in image retrieval applications. In this paper[1] we propose a system architecture for an automatic large-scale medical image understanding which aims at a formal fusion of feature extraction techniques operating on the bit-level representation of images (and time series data) with formal background knowledge represented in ontologies. We put forward a hierarchical framework of ontologies to formulate a precise and at the same time generic representation of the existing high level knowledge in the medical domain. We present a system architecture which aims at an integration of high- and low-level features on various abstraction levels allowing cross-modal as well as cross-lingual retrieval through Content Based Image Retrieval (CBIR), query by keyword/text, and query by concept. Experiences from implementing a framework supporting these features are reported as well as our efforts to select and acquire relevant domain knowledge.

1 Introduction

As more and more hospitals switch to a complete IT-based management of patient data, *e. g.*, by using Electronic Health Records (EHR), huge amounts of medical information has to be stored and made accessible using computer technology. The nature of this data is diverse in more than one aspect. Textual descriptions contain all sorts of information. The range goes from patients' names, age, gender across descriptions of particular diagnosis reports to complete medical cases which can span across long periods of time.

But medical information is not limited to text documents. Rapid advances in imaging technology have dramatically increased the amount of medical image

[1] This research has been supported in part by the THESEUS Program in the MEDICO Project, which is funded by the German Federal Ministry of Economics and Technology under the grant number 01MQ07016. The responsibility for this publication lies with the authors.

data generated daily by hospitals, pharmaceutical companies, and academic medical research.[2] Technologies such as 4D 64-slice Computer Tomography (CT), whole-body Magnetic Resonance Imaging (MRI), 4D Ultrasound, and the fusion of Positron Emission Tomography and CT (PET/CT) can provide incredible detail and a wealth of information with respect to the human body anatomy, function, and disease associations. Therefore, one has to take into account data from various modalities (text documents, various imaging modalities etc.) to obtain comprehensive information. Meanwhile, search and retrieval should be independent of the concrete modality. We will use the term *cross-modal* to address the modality independence throughout this document. Another aspect is the independence of particular languages like English or German; we will use the term *cross-lingual* to refer to this.

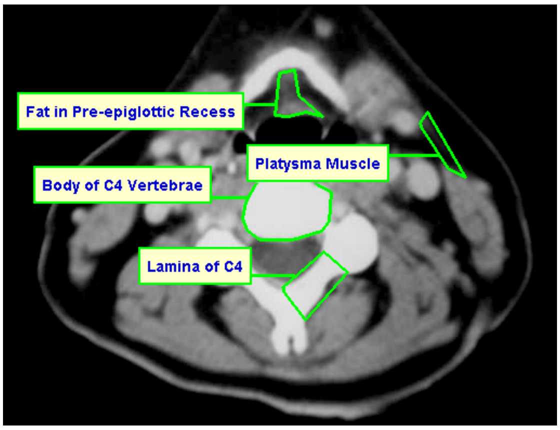

Fig. 1. Annotated Neck Region

The increase in the volume of data has brought about significant advances in techniques for analyzing such data. The precision and sophistication of different image understanding methods, such as object recognition and image segmentation, have also improved to cope with the increasing complexity of the data. However, these improvements in analysis have not resulted in more flexible or generic image understanding techniques. Instead, the analysis methods are very object specific and difficult to scale across different applications. Consequently, current image search techniques, whether for Web sources or for medical Picture Archiving and Communications System (PACS), are still dependent on the manual and subjective association of keywords to images for retrieval.

[2] For example, University Hospital of Erlangen, Germany, has a total of about 50 TB of medical images. Currently they have approx. 150,000 medical examinations producing 13 TB per year.

Manually annotating the vast numbers of images which are generated and archived in the medical practice is not an option and also unnecessary since several object recognition algorithms already exist. The problem of current systems is that it still needs human intelligence to decide which object recognition algorithm to apply because knowledge from various dimensions has to be taken into account (see Sect. 3.2). Our goal is to eventually allow a fully automatic image segmentation and abstract annotation as shown in Fig. 1 (image taken from Radiologic Anatomy Browser[3]).

Section 2 reviews related work in this field of research and foundations that we base our system on. In Sect. 3 we describe in detail the design considerations for modeling the background knowledge across various abstraction levels starting with very generic concepts like time and space down to specific concepts from the medical practice. In Sect. 4 we give an overview about our proposed system architecture which integrates techniques from both the symbolic and sub-symbolic world of AI. Then, in Sect. 5, we give an overview of the software framework that we have started to implement. Finally, in Sect. 6 we present a conclusion of our experiences and define our next steps.

2 Related Work

[Buitelaar et al., 2006] investigate methods that integrate the annotation of multimedia data of all forms within one single retrieval framework. Once our proposed system is running, their work could be an extension. [Romanelli et al., 2007] show how MPEG7[4] can be used as a generalized formalism for segmentation of arbitrary document formats and annotation of segments. This technique was already applied successfully in the project SmartWeb.[5]

There are numerous advanced object recognition algorithms for the detection of particular objects on medical images: [Hong et al., 2006] at the anatomical level, [Tu et al., 2006] at the disease level, and [Comaniciu et al., 2004] at the functional level. But the specificity of these algorithms is also their limitation: Existing object recognition algorithms are not at all generic. One challenge in automatic medical image annotation that we want to address is to implement a framework which is able to decide automatically which object identifier to apply for certain images and image regions.

In the recent past, applications like FIRE [Deselaers et al., 2004] showed that image retrieval only based on sub-symbolic interpretation of the images works reasonably well for a number of applications. In the IRMA project[6] [Lehmann et al., 2003], such technology was applied to the medical domain.

A number of research publications in the area of ontology-based image retrieval emphasize the necessity to fuse sub-symbolic object recognition and abstract domain knowledge. [Vompras, 2005] proposes an integration of spatial

[3] http://rad.usuhs.mil/rad/iong/homepage.html
[4] MPEG Homepage: http://www.chiariglione.org/mpeg/
[5] http://www.smartweb-project.de
[6] IRMA: Image Retrieval in Medical Applications, http://irma-project.org/

context and semantic concepts into the feature extraction and retrieval process using a relevance feedback procedure. [Papadopoulosa et al., 2006] combine sub-symbolic machine learning algorithms with spatial information from a domain ontology. [Su et al., 2002] present a system that aims at applying a knowledge-based approach to interpret X-ray images of bones and to identify the fractured regions. [Mechouche et al., 2007] present a hybrid method which combines symbolic and sub-symbolic techniques for the annotation of brain Magnetic Resonance images. While it focuses only on one modality and body region, their approach shares the use of OWL DL [McGuinness and van Harmelen, 2004], SWRL rules [Horrocks et al., 2004] and DL reasoning with our proposal. The BOEMIE EU project[7] also focuses on knowledge acquisition independent from modalities. But while we obtain the formal medical domain knowledge from existing large-scale ontologies they use a bootstrapping approach to evolve ontologies to also cover missing concepts [Castano et al., 2006].

3 Ontological Modeling

As for the term *ontology* in this document we follow the definition by [Gruber, 1995]: "An ontology is a formal specification of a (shared) conceptualization." Ontologies are usually structured in various layers or levels, with the rationale that those at higher levels are more stable, shared among more people, and thus change less often than those at lower levels. Usually, one distinguishes representational ontologies, upper-level ontologies, mid-level ontologies, and low-level or domain ontologies.

3.1 Proposed Ontology Hierarchy

In the following paragraphs we will describe the ontology hierarchy that we designed for the presented problem in the medical domain. Fig. 2 illustrates this hierarchy.

Representational Ontologies define the vocabulary with which the other ontologies are represented; examples are RDF/S [Brickley and Guha, 2004] and OWL. The used Representational Ontology may vary for the ontology we want to include. We do not need Open World Assumption as in OWL and reasoning for the descriptions of documents and images but we need it for concepts in the Medical Ontologies.

The *Upper Ontology* is a high-level, domain-independent ontology, providing a framework by which disparate systems may utilize a common knowledge base and from which more domain-specific ontologies may be derived [Kiryakov et al., 2001]. It describes very general concepts like time, space, organization, person, and event which are the same across all domains.

The *Information Element Ontology* belongs to the mid-level ontologies which serve as a bridge between abstract concepts defined in the upper ontology and

[7] `http://www.boemie.org/`

domain specific concepts specified in the domain ontologies. While ontologies may be mapped to one another at any level, the mid-level and upper ontologies are intended to provide a mechanism to make this mapping of concepts across domains easier. Mid-level ontologies may provide more concrete representations of abstract concepts found in the upper ontology [Semy et al., 2004].

The Information Element Ontology contains the information elements that we want to annotate (images, text documents, videos, ...). Building upon the aforementioned generalized segmentation provided by MPEG7 segments of images and documents are treated the same way as complete documents.

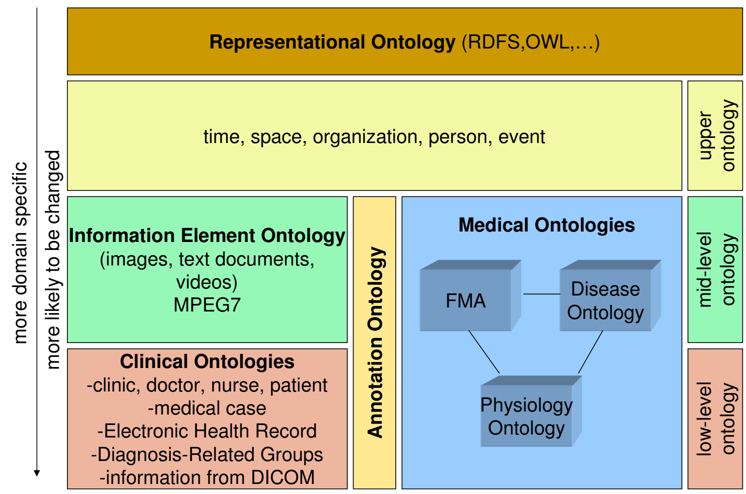

Fig. 2. Proposed Ontology Hierarchy

Clinical Ontologies belong to the low-level domain ontologies. These ontologies "specif[y] concepts particular to a domain of interest and represent concepts and their relationships from a domain specific perspective. While the same concept may exist in multiple domains, the representations may widely vary due to the differing domain contexts and assumptions. [...] Reusing well established ontologies in the development of a domain ontology allows one to take advantage of the semantic richness of the relevant concepts and logic already built into the reused ontology." [Semy et al., 2004]. We use the clinical ontologies to specify roles and domain specific abstract data aggregations (like Electronic Health Records) from the clinical practice. For example, the concepts nurse, doctor, patient, and medical case belong to these ontologies.

For the *Medical Ontologies* a clear separation into mid- and low-level ontologies is not possible since they have to cover a broad spectrum of concepts. It ranges from particular disease symptoms to abstract descriptions of human anatomy and abstract relations between anatomical entities.

In the medical domain, huge amounts of knowledge are already formulated in abstract ontologies like, *e. g.*, within the Foundational Model of Anatomy (FMA) Ontology [Rosse and Mejino, 2003]. To be able to map concepts from this ontology to other ontologies, we rely on work by [Noy and Rubin., 2007] which translates the frame-based FMA to OWL. Due to the intractability of the FMA with current reasoners caused by its high complexity, we only use fragments which are relevant for object detection on medical images. To cover not only anatomical concepts but also diseases and functional relations between anatomical entities we will integrate pathological and physiological ontologies as well.

The *Annotation Ontology* includes concepts which are used to annotate elements from the ontologies on the left-hand side of Fig. 2 with concepts from the Medical Ontologies that were detected during object recognition. Using an ontology (instead of just a simple relation) allows us to express that, *e. g.*, an image *partially* deals with a specific concept from the anatomical ontology, because only parts of it are on the picture. Since we use automatically computed annotations, these annotations come only with a certain likelihood. At least in some cases we want to annotate the relations with a probability. The Annotation Ontology allows to express such qualifications as properties of attributes.

3.2 Examples of Use for Fusion of Syntax and Semantics

To explain the benefits of an integration of existing low-level object recognition techniques with declarative domain knowledge from an ontology we will give four scenarios. For the proposed system we can separate two different tasks: *analysis* while new documents are added to the system and *search and retrieval* when queries are answered. The first two scenarios apply for analysis as well as for search and retrieval. The remaining two only apply for search and retrieval.

Search Space Reduction One of the key challenges in semantic image annotation is the huge search space of possible objects that can be depicted. In the medical domain several highly specific object recognition algorithms exist. Each of them is very accurate at detecting a particular type of object in medical images from a specific modality (CT, PET, ultrasound, *etc.*).

Given an arbitrary image it still needs human intelligence to select the right object recognizers to apply to an image. Aiming to gain a pseudo-general object recognition one can try to apply the whole spectrum of available object recognition algorithms. But it turns out that in generic scenarios even with state-of-the-art object recognition tools the accuracy is below 50 percent [Chan et al., 2006, Müller et al., 2006].

Therefore we propose to implement an iterative object recognition process. In the first step, generic algorithms try to detect landmarks like bones. The formal background knowledge about anatomy can then be applied to decide by means of a reasoner, which objects are likely to be around the detected landmark and the image regions to search. Possible dimensions for limiting the search space using background are body region, image modality (certain anatomical entities can be made visible only by certain imaging modalities), anatomical knowledge (if there is a heart on an image, it does not make sense to look for a knee joint), *etc.*Thus the search space can be limited drastically making the object detection faster: only specific image regions are searched for particular objects which are likely to appear in these regions.

Semantic Relevance Test Applying purely low-level object recognition algorithms on a given image will end up in a list of objects that were detected. By using the formal knowledge about anatomy, these results can be checked whether their combination makes sense from an abstract anatomical point of view. If, *e. g.*, the left ventricle is found on an image, it makes sense to look for the right ventricle as well. But it does not make sense at all to look for a knee joint in the immediate neighborhood. This information can be fed back into the object recognition algorithms and used as a reinforcement signal, to focus the erroneous algorithms to other image regions *etc.*

Query using Semantic Similarity Typical CBIR retrieves only those images from a database which are *visually* similar. CBIR based on *semantic* similarity means that in the first step the query image is searched for objects which can be mapped to concepts in the ontology. In the second step the ontology is used to infer similar concepts and select images which deal with the same or similar concepts (for details see Sect. 4.5). Only after this step, the query is performed across all sorts of media which are stored semantically annotated in the knowledge base. Using semantic similarity allows a fully cross-modal as well as cross-lingual retrieval, since the document space is searched via the associated concepts and not the concrete contents.

Query Refinement During search the formal background knowledge can be used to analyze a query before the actual retrieval of results from the knowledge base is begun. Starting from a simple search phrase that is entered by the user, the reasoner can deduce related concepts in the ontology and provide the user with a list of concepts from which he can choose. The anatomical knowledge can also be used to deduce whether the entered concept is too generic (like, *e. g.*, searching for the concept "organ"), and request the user to refine the query. Thus, in an iterative process, the user can be guided in formulating his query without time consuming retrieval and relevance sorting of huge result sets.

4 System Architecture

Fig. 3 gives an overview of the proposed system architecture. As stated above we aim to combine various access methods to provide the user with a generic and scalable search and retrieval across medical information of all formats. Our experiences from integrating this framework are delineated in Sect. 5.

4.1 Graphical User Interface

The graphical user interface allows the user to formulate queries in different ways. These options can be divided into two main categories: Either the query is to be evaluated and executed in a *syntactic* or in a *semantic* way. Only for the purpose of explanation we will split them up into isolated components. In our implementation we allow combinations of all query types.

Query by Concept The most straight-forward type of Query by Concept interface is to allow the user to select concepts from ontologies. A more elaborate way of Query by Concept is to allow the user to type in keywords which he wants to search for in the knowledge base. Before starting retrieval of search results from the knowledge base, the keywords are mapped to concepts from the ontology. If disambiguation is needed the matching concepts are returned to the user and he can select those which he intents to search for. This method is closely related to what is described in Sect. 3.2 in the paragraph about Query Refinement.

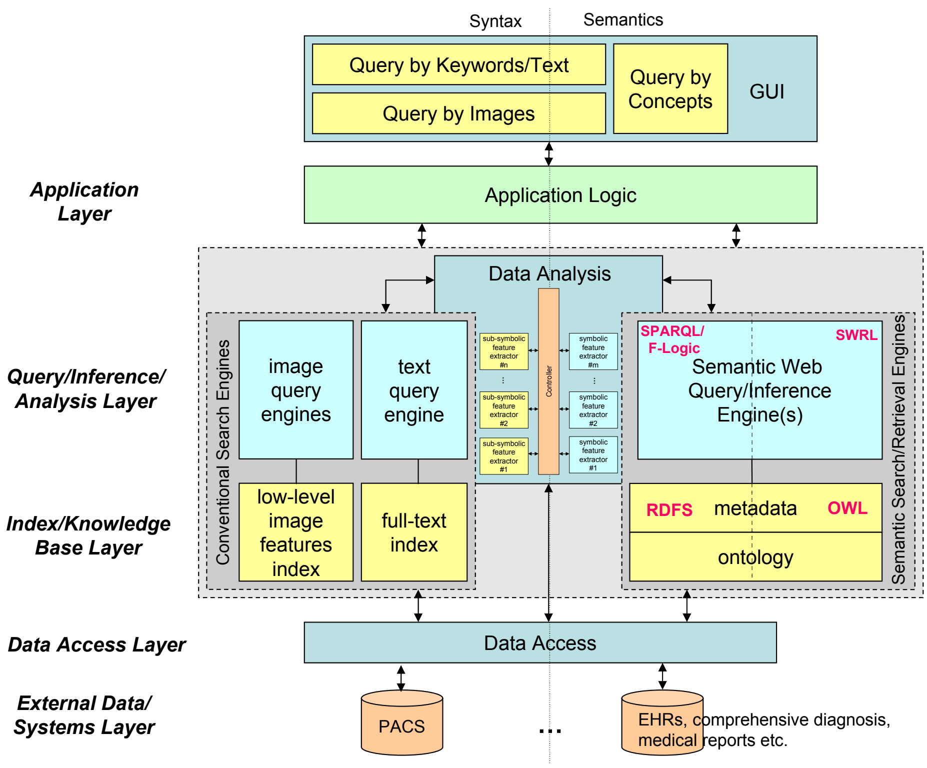

Fig. 3. Proposed System Architecture

Syntactic Query by Image CBIR can be performed both in a purely syntactic or a semantic way. In the case of syntactic CBIR the user is actually looking for *visually similar* images. Sect. 2 refers to existing work in this area. For this type of CBIR the search engine will only return images from the same imaging modality and only those which are visually similar. This technique requires an efficient index of low-level images features. The similarity measure is calculated exclusively based on comparison of low-level features extracted from the query image and features in the index.

Semantic Query by Image The other case is CBIR based on *semantic similarity*. That means, the image is analyzed with the Data Analysis component. This analysis results in annotations with concepts from the Medical Ontologies on the right-hand side of the ontology hierarchy (see Fig. 2. Only after this step, the query is performed across all sorts of media which is stored semantically annotated in the knowledge base. Only those documents (medical diagnosis reports, EHRs, images, *etc.*) are returned which were annotated with concepts that were extracted from the query image.

Query by Keywords/Text This case can be divided into (1) syntactic and (2) semantic cases. Like in most current search engines users formulate their queries using keywords. The most simple case is using one keyword or a logical combination of them to search for documents which contain these keywords. This type of search belongs to (1). It is well understood, robust, fast, and performed millions of times everyday via all major search engines on the Internet. Another advantage is that users are very familiar with this this way of searching. Therefore we will include this technique as one component in our system.

The main drawback of plain syntactical (1) search by keyword is that it is dependent on language and modality. Adding semantics to query by keywords (2) has already been described in the paragraph about *Query by Concept*.

Image retrieval based on search by keyword usually depends on the association of keywords which appear in the text around an image with the image itself—which is prone to errors—or on manual annotations which are strongly influenced by the subjective view of the person who is annotating the image.

Another approach is to allow users to formulate queries in longer text blocks. Again, this scenario can be divided into (1) a syntactic and (2) a semantic case. To answer such queries, (1) the Application Logic can use purely statistical methods to compare the text to all other documents in the index. The semantic approach (2) is to parse the query sentence(s) using techniques from natural language processing (NLP), represent their meaning using concepts from the domain ontology and perform a query based on semantic similarity. Thus, this type of search can be cross-modal and cross-lingual.

4.2 Application Logic

The Application Logic connects the graphical user interface with the Query and Inference Layer. As described above some of the GUI components require access to metadata and ontology. The Application Logic abstracts from the Semantic Search and Retrieval Engines (see Sect. 4.4) and provides an interface for such requests.

In most cases different types of queries will be linked together, *i. e.*, to perform a semantic CBIR which is limited to images of patients whose names match a certain (key)word. In such cases, the application logic has to demultiplex the complex query which is coming from the GUI and split it up into multiple requests for the Conventional Search Engines (see Sect. 4.3) and the Semantic Search and Retrieval Engines (see Sect. 4.4). After collecting the results of all sub-queries, the results have to be combined and filtered for the display through the GUI. For the Query Refinement (see Sect. 3.2) the Application Logic is controlling the iterative communication between user (via GUI) and search engines.

4.3 Conventional Search Engines

This component consists of various *image query engines* with associated indexes of low-level features. These query engines answer queries for images based on visual similarity. To allow better retrieval results we combine indexes of different low-level feature extraction algorithms (see Sect. 5).

The *text query engine* performs keyword-based search on a full-text index. After an initial indexing phase, searching this index is fast, robust, and accurate at the same time. As stated above, we rely on existing technology for this component.

Queries from the Application Logic are executed by the different engines, result lists are generated and returned via a well-defined interface. At the lower end the databases and storage systems (see Sect. 4.7) are interfaced which contain the data that is to be indexed and searched. This access is abstracted through the Data Access Layer (see Sect. 4.6).

4.4 Semantic Search and Retrieval Engines

This component provides all other sub-components with access to stored metadata, the ontology framework (see Sect. 3), and reasoning services. Depending on the expressivity of the different ontology components and metadata, different query interfaces are available. For ontology components that have a more database-like schema we use RDFS-based storage, a SPARQL [Prud'hommeaux and Seaborne, 2007] query interface and F-Logic [Kifer et al., 1995] as the rule language for retrieval of semantic data. For those ontology components which cover concepts with a higher expressivity and which are used with open-world reasoning (*i. e.*, the Medical Ontologies) we use OWL DL and DL reasoners like Pellet [Sirin and Parsia, 2004] which have support for rule languages, *e. g.*, SWRL [Horrocks et al., 2004].

4.5 Data Analysis

For the Data Analysis we can differentiate two basic situations: The (1) *query analysis* versus the (2) *data analysis from legacy systems in batch mode*. Situation

(1) occurs whenever a query has to be mapped to concepts from the ontology. When new documents have to be added to the index, they have to be analyzed as well. We assign the latter task to situation (2). In (1) the data originated from the GUI and flows top-down through the system architecture. In (2) the application logic controls the indexing process of new documents. In this case the data originates from the External Data and Systems Layer (see Sect. 4.7) and the data flow is bottom-up.

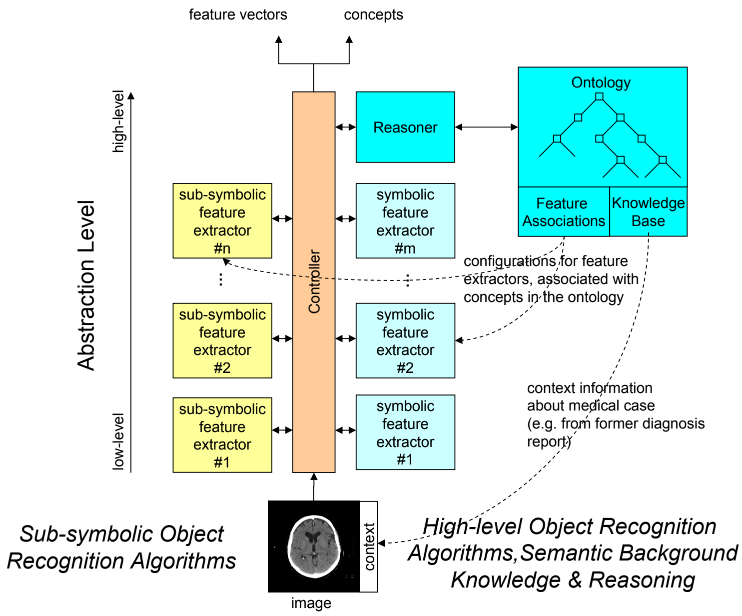

Fig. 4. Integration of Feature Extraction and Background Knowledge

Fig. 4 is an expansion of the miniature in the Data Analysis subcomponent in the System Architecture diagram (Fig. 3). For simplification we included the *Reasoner* and *Ontology/KB*. In fact—and as shown in the System Architecture diagram—, the *Semantic Search and Retrieval Engines* subcomponent is interfaced to connect the *Data Analysis* to the high-level knowledge.

As described in Sect. 2 we can build our system upon numerous existing specific object recognition algorithms. The crucial point is to add a tight fusion with background knowledge represented in ontologies. The final goal is a close interaction of low-level object recognition and reasoning on the high-level domain knowledge. Therefore we chose to integrate the functionality within multiple components as shown in Fig. 4. We use the term *symbolic feature extractor* to emphasize that the *sub-symbolic* algorithms on the left-hand side output only *sub-symbolic* feature vectors. In contrast the right-hand side algorithms' output can be mapped directly to symbolic *concepts* in the ontology.

For each image (as well as volume data set, video *etc.*) the raw input data is first passed to a central controller which starts analyzing the input with low-level algorithms from the left side, *i. e.*, to perform shape detection, calculate color histograms, and so on. The image can be accompanied by contextual data that is already stored in the knowledge base, *e. g.*, information from previous (semantically annotated) diagnosis reports for this patient (see dashed line). From the data flow perspective, this data will be available through the interface of the reasoner. The results from these low-level processing steps are passed back to the controller which can decide which other feature extractors to apply next. Since many of the more sophisticated object recognition algorithms depend on low-level preprocessing, the output of the first step can be reused in further processing steps. These low-level features are saved in the low-level image features index (Sect. 4.3) for search and retrieval based on visual similarity. Combined with the metadata from the DICOM header[8] the extracted low-level information is used to select a number of object recognition algorithms from the right-hand side of Fig. 4. In the *Feature Associations* part of the ontology we store associations between abstract medical concepts and feature extractor configurations to detect these concepts.

A similar architecture could also be applied for the extraction of high-level concepts from textual documents. Similar to the sub-symbolic feature extractors for images on the left-hand side of Fig. 4 statistics based methods can be used to extract simple features. Techniques from NLP can be used to extract more abstract information from the texts and map them to concepts in the ontology.

4.6 Data Access

This component provides an abstraction from existing clinical databases for our proposed retrieval system. The various data archiving systems which are already in use in modern hospitals are usually specially tailored to archive only documents from a specific modality or type of diagnosis. The Data Access Layer provides the rest of the system with a unified access to all data sources.

4.7 External Data and Systems

The External Data and Systems Layer consists of the various available data sources. Medical images are usually stored in PACS (Picture Archiving and Communications Systems). Electronic Health Records (EHRs) are stored in other database systems. Those records include comprehensive diagnoses, medical reports *etc.*Today, these databases are heterogeneous in the way they store information and in the interfaces they have to make the data available. But since all these sources contain valuable information for medical diagnoses we aim at integrating as many as possible. The semantic abstraction from the concrete document modality allows us to retrieve textual reports along with images from different medical imaging modalities as the result for a single query.

[8] *Digital Imaging and Communications in Medicine* (DICOM) is a standard for metadata storage for medical images (see http://medical.nema.org/).

5 Implementation

For our implementation we chose to extend the existing metadata extraction framework Aperture[9] which already supports indexing a large variety of document formats. It has a plug-in architecture which allows a simple extension by feature extractors for currently unsupported document formats or additional objects. Extracted features are returned in RDF [Hayes, 2004] format which is stored in a Sesame triple store.[10] The framework always extracts the full text from all supported text document formats. Depending on the type of document, additional attributes like author and date of creation are extracted. To allow matching of keywords and phrases within the extracted full text, it is stored in a Lucene[11] index. The key of the document in the Lucene index refers to the URI which is assigned during the crawling with the Aperture framework and links the RDF data to the elements within Lucene. For search by keyword, we rely on the functionality provided by Lucene. Retrieval based on other metadata attributes is performed by querying the triple store via a SPARQL endpoint.

To also cover images we extended the Aperture framework by three low-level MPEG7 feature extractors: *ScalableColor* (basic color distribution description), *ColorLayout* (spatial distribution of colors) and *EdgeHistogram* (edge distribution with a histogram based on local edge distribution).

Based on the implementation of Lucene Image Retrieval Engine (LIRE)[12] we can already support content based image retrieval using visual similarity by applying vector space based similarity measures on the extracted feature vectors. Higher sophisticated object recognition algorithms like those mentioned in Sect. 2, *e. g.*, to detect ventricles, have to be integrated to also allow a truly semantic annotation of the images.

6 Conclusion and Future Work

In this paper we proposed a close integration of sub-symbolic pattern recognition algorithms and semantic domain knowledge represented in formal ontologies. The vision is to combine the techniques from both fields to bridge the gap between a symbolic and sub-symbolic world for a generic understanding of medical images and text. The use of formal ontologies, together with the reasoning capabilities on top of them, forms the essence behind better information retrieval. By abstracting from the syntactic content representation, it is possible to perform semantic matching between queries and the content. Additionally, the user is provided with an extremely flexible interface which allows cross-modal as well as cross-lingual queries. In a number of scenarios we discussed benefits of the integration of syntactic and semantic techniques for a faster and more scalable information retrieval application.

[9] http://aperture.sourceforge.net/
[10] http://www.openrdf.org/
[11] http://lucene.apache.org/
[12] http://www.semanticmetadata.net/lire/

We proposed an ontology framework to model the background knowledge of our application domain which aims to reuse and map together existing ontologies from the medical domain. We have shown a system architecture detailing the integration of sub-symbolic and symbolic components. Basic components have already been implemented through extensions of existing software projects. At the current state our system is able to perform Query by Keyword and Query by Image based on visual similarity. Among our next steps will be the integration of existing object recognition algorithms as described in Sect. 2 and mapping their output to concepts in the ontology. Another area of work will be the selection of relevant fragments from the existing medical ontologies in cooperation with medical experts.

References

Brickley and Guha, 2004. Brickley, D. and Guha, R. (2004). RDF vocabulary description language 1.0: RDF Schema.

Buitelaar et al., 2006. Buitelaar, P., Sintek, M., and Kiesel, M. (2006). A lexicon model for multilingual/multimedia ontologies. *Proceedings of the 3rd European Semantic Web Conference (ESWC06).*

Castano et al., 2006. Castano, S., Ferrara, A., and Hess, G. (2006). Discovery-driven ontology evolution. *SWAP 2006: Semantic Web Applications and Perspectives - 3rd Italian Semantic Web Workshop.*

Chan et al., 2006. Chan, A. B., Moreno, P. J., and Vasconcelos, N. (2006). Using statistics to search and annotate pictures: an evaluation of semantic image annotation and retrieval on large databases. *Proceedings of Joint Statistical Meetings (JSM).*

Comaniciu et al., 2004. Comaniciu, D., Zhou, X., and Krishnan, S. (2004). Robust real-time myocardial border tracking for echocardiography: an information fusion approach. *IEEE Transactions in Medical Imaging*, 23 (7):849–860.

Deselaers et al., 2004. Deselaers, T., Keysers, D., and Ney, H. (2004). FIRE – flexible image retrieval engine: ImageCLEF 2004 evaluation. In *CLEF 2004, LNCS 3491*, pages 688–698.

Gruber, 1995. Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing. In *International Journal of Human-Computer Studies*, volume 43, pages 907–928.

Hayes, 2004. Hayes, P. (2004). RDF semantics. W3C Recommendation.

Hong et al., 2006. Hong, W., Georgescu, B., Zhou, X. S., Krishnan, S., Ma, Y., and Comaniciu, D. (2006). Database-guided simultaneous multi-slice 3D segmentation for volumetric data. In Leonardis, A., Bischof, H., and Prinz, A., editors, *Journal of the European Conference on Computer Vision (ECCV 2006)*, volume 3954, pages 397–409. Springer-Verlag.

Horrocks et al., 2004. Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., and Dean, M. (2004). SWRL: A semantic web rule language combining OWL and RuleML. Technical report, W3C Member submission 21 may 2004.

Kifer et al., 1995. Kifer, M., Lausen, G., and Wu, J. (1995). Logical foundations of object-oriented and frame-based languages. *J. ACM*, 42(4):741–843.

Kiryakov et al., 2001. Kiryakov, A., Simov, K., and Dimitrov, M. (2001). OntoMap: Portal for upper-level ontologies. In *Proceedings of the International Conference on Formal Ontology in Information Systems*, pages 47–58.

Lehmann et al., 2003. Lehmann, T., Güld, M., Thies, C., Fischer, B., Spitzer, K., Keysers, D., Ney, H., Kohnen, M., Schubert, H., and Wein, B. (2003). The IRMA project. A state of the art report on content-based image retrieval in medical applications. *Proceedings 7th Korea-Germany Joint Workshop on Advanced Medical Image Processing*, pages 161–171.

McGuinness and van Harmelen, 2004. McGuinness, D. L. and van Harmelen, F. (2004). OWL web ontology language overview.

Mechouche et al., 2007. Mechouche, A., Golbreich, C., and Gibaud, B. (2007). Towards an hybrid system using an ontology enriched by rules for the semantic annotation of brain MRI images. In Marchiori, M., Pan, J., and de Sainte Marie, C., editors, *Lecture Notes in Computer Science*, volume 4524, pages 219–228.

Müller et al., 2006. Müller, H., Deselaers, T., Lehmann, T., Clough, P., Kim, E., and Hersh, W. (2006). Overview of the ImageCLEFmed 2006 medical retrieval and annotation tasks. In *Accessing Multilingual Information Repositories*, volume 4022 of *Lecture Notes in Computer Science*. Springer-Verlag.

Noy and Rubin., 2007. Noy, N. F. and Rubin., D. L. (2007). Translating the Foundational Model of Anatomy into OWL. In *Stanford Medical Informatics Technical Report*.

Papadopoulosa et al., 2006. Papadopoulosa, G. T., Mezaris, V., Dasiopoulou, S., and Kompatsiaris, I. (2006). Semantic image analysis using a learning approach and spatial context. In *Proceedings of the 1st international conference on Semantics And digital Media Technologies (SAMT)*.

Prud'hommeaux and Seaborne, 2007. Prud'hommeaux, E. and Seaborne, A. (2007). SPARQL query language for RDF. Technical report, W3C.

Romanelli et al., 2007. Romanelli, M., Buitelaar, P., and Sintek, M. (2007). Modeling linguistic facets of multimedia content for semantic annotation. *Proceedings of International Conference on Semantics And digital Media Technologies (SAMT07)*.

Rosse and Mejino, 2003. Rosse, C. and Mejino, J. L. V. (2003). A reference ontology for bioinformatics: The Foundational Model of Anatomy. In *Journal of Biomedical Informatics*, volume 36, pages 478–500.

Semy et al., 2004. Semy, S. K., Pulvermacher, M. K., and Obrst, L. J. (2004). Toward the use of an upper ontology for U.S. government and U.S. military domains: An evaluation. Technical report, MITRE Corporation.

Sirin and Parsia, 2004. Sirin, E. and Parsia, B. (2004). Pellet: An OWL DL reasoner. In Haarslev, V. and Möller, R., editors, *Description Logics*, volume 104 of *CEUR Workshop Proceedings*. CEUR-WS.org.

Su et al., 2002. Su, L., Sharp, B., and Chibelushi, C. (2002). Knowledge-based image understanding: A rule-based production system for X-ray segmentation. In *Proceedings of Fourth International Conference on Enterprise Information System*, volume 1, pages 530–533, Ciudad Real, Spain.

Tu et al., 2006. Tu, Z., Zhou, X. S., Bogoni, L., Barbu, A., and Comaniciu, D. (2006). Probabilistic 3D polyp detection in CT images: The role of sample alignment. *IEEE CVPR*, 2:1544–1551.

Vompras, 2005. Vompras, J. (2005). Towards adaptive ontology-based image retrieval. In Stefan Brass, C. G., editor, *17th GI-Workshop on the Foundations of Databases, Wörlitz, Germany*, pages 148–152. Institute of Computer Science, Martin-Luther-University Halle-Wittenberg.