Semantic Navigation in Medical Heterogeneous Resources: Text and Video

Olfa Dridi
RIADI Laboratory
National School of Computer Sciences
Campus universiatire, Manouba, Tunis, Tunisia
dridi_olfa {at} yahoo.fr

Mohamed Ben Ahmed
RIADI Laboratory
National School of Computer Sciences
Campus universiatire, Manouba, Tunis, Tunisia

***Abstract*— In this paper, we present our approach of intelligent retrieval over repository of contextualized and semantically annotated resources. This approach indexes a collection of textual and video resources in three ways: descriptive, contextual and conceptual. This triple indexation is made possible through the generation of descriptive and contextual meta-information and semantic annotation of the resources contents. It allows more efficient mapping between resources and queries. Descriptive and contextual meta-information generation is based on usage context ontology, while semantic annotations are based on a domain ontology dedicated to "breast cancer". Our approach is addressed to the medical community: students, surgeon, patient, General practitioner, etc. All these users have the same interest domain "breast cancer", but they have different profiles, because they have different backgrounds.**

Keywords—descriptive annotation, contextual annotation, semantic annotation, breast cancer ontology, semantic navigation.

I. INTRODUCTION

In the recent years, the scientific community was interested to multimedia corpus containing video and textual resources. So, the idea of integration of these kinds of resources became crucial in order to overcome the insufficiencies of existing information retrieval systems (IRS).

Video resources are expressive and realistic media for communication and entertainment, but the high cost and difficulty of acquiring, manipulating, and distributing these media have traditionally limited exploitation in this domain. Recently, however, as cameras, bandwidth, and storage become increasingly affordable, a wider spectrum of users produces digital video in ever-increasing volumes. Unfortunately, existing approaches to interacting with digital video have not kept up with the deluge of raw material. Professional methods for searching, browsing, annotating, and editing video material are too complex and often lack the immediacy of interactivity required by users [1].

We have interested to the domain of video and textual annotation using ontology. In this paper, we present our work concerning knowledge extraction of medical video and textual resources. We propose a platform of intelligent retrieval over repository of contextualized and semantically annotated resources. This platform indexes a collection of textual and video resources in three ways: descriptive, contextual and conceptual. This triple indexation is made possible through the generation of descriptive and contextual meta-information and semantic annotation of the resources contents. It allows more efficient mapping between resources and queries illustrated by a high level of recall and precision. Descriptive and contextual meta-information generation is based on usage context ontology, while semantic annotations are based on a domain ontology dedicated to "breast cancer"; ontology which is enriched by pertinent terms, nouns phrases and concepts extracted from resources contents. Annotated resources will be targeted by a search engine in order to retrieve the resources and their fragments satisfying the needs and profiles of users. The developed prototype proposes two modalities of search: the first one is by queries and the second one is by navigation in semantic map capturing three semantics: referential, differential and inferential. The second search mode offers navigation in the semantic proximity field which nodes are annotated by indexing documents responding to precise, exploratory and connotative modes of search. So, the navigation mode integrates the cartography paradigms in the actual information retrieval systems which offer to the users more efficient search. The user has as response to his query a map that contains searched concepts as nodes.

For this, we present an overview of state-of-art of approaches that use ontology to perform semantic video annotation and semantic textual annotation. Then, we detail our approach that exploits the domain knowledge embedded into ontology in order to generate descriptive, contextual and semantic annotations. We have chosen a video and textual resource in the medical domain. Experiments have been performed in two different video: video for Breast Cancer Surgery and video for Breast cancer course.

II. CONTEXT OF OUR RESEARCH

Image and video assets constitute extremely rich information sources, ubiquitous in a wide variety of diverse applications and tasks related to information management, both for personal and professional purposes. Inevitably, the value of the endowed information amounts to the effectiveness and efficiency at which it can be accessed and managed. This is where semantic annotation comes in, as it designates the schemes for capturing the information related to the content.

As already indicated, two crucial requirements featuring content annotation are the interoperability of the created meta-information and the ability to automatically process them. The former encompasses the capacity to share and reuse annotations, and by consequence determines the level of seamless content utilisation and the benefits issued from the annotations made available; the latter is vital to the realisation of intelligent content management services [2].

Visual content semantics, as multimedia semantics in general, comes into a multilayered, intertwined fashion. It encompasses, amongst others, thematic descriptions addressing the subject matter depicted (scene categorisation, objects, events, etc.), media descriptions referring to low-level features and related information such as the algorithms used for their extraction, respective parameters, etc., as well as structural descriptions addressing the decomposition of content into constituent segments and the spatiotemporal configuration of these segments.

In this article, we present a framework for video annotation, visualization, and interaction that harnesses computer vision to aid users in visualising and exploiting digital video. We will discuss a series of related works that address some of the specific tasks common to many video applications.

These projects focus on the problems of video object annotation, static visualization, and temporal navigation within a video. Many video manipulation tasks – such as navigation, annotation, and editing – are still time-consuming processes, in some part due to the sheer volume of available video material that must be viewed and recalled. Commercial software for working with video generally offers only token mnemonic assistance, representing each video segment with only a single frame, typically defaulting to the segment's first frame. But this frame is often unrepresentative of the video content, and in any case does not illustrate other important aspects of the video, such as camera and subject motion [3].

III. VIDEO ANNOTATION TOOLS

The increase in the amount of video data deployed and used in today's applications not only caused video to draw increased attention as a content type, but also introduced new challenges in terms of effective content management. Image annotation approaches can be employed for the description of static scenes found in a video stream; however, in order to capture and describe the information issuing from the temporal dimension featuring a video object, additional requirements emerge [4].

In the following, we survey typical video annotation tools, highlighting their features. In addition to tools that constitute active research activities, we also examine representative video annotation systems that despite no longer maintained, are still accessible and functional;

A survey of existing video annotation systems are presented in the following :

- IBM – MPEG-7 Annotation Tool
- Ricoh – Movie Tool
- ZGDV – VIDETO
- COALA – LogCreator
- CSIRO's CMWeb tools
- Microsoft's MRAS
- VIA
- VideoAnnEx

A. IBM's MPEH-7

IBM's MPEG-7 Annotation Tool provides support for both MPEG-1 and MPEG-2 files as well as regional annotations. It also comes with a shot detection algorithm, an easy-to-use interface and a customisable lexicon [5]. If a video has a different format than it cannot be displayed correctly. The lexicon is also restricted to three default categories (event, static scene and key objects), although free text keywords can also be added. IBM's system doesn't support hierarchical video segmentation.

B. Ricoh's MovieTool

Ricoh's MovieTool does support hierarchical segmentation within a timeline-based representation of the video [6]. The automatic shot boundary detection algorithm permits changes to threshold settings. The MovieTool is the most mature and complete of the systems, but has a complicated user interface which is closely tied to the MPEG-7 specification. The user has to have a good knowledge of the large and complex XML Schema definition of MPEG-7 in order to browse using the MPEG-7 Editor.

C. ZGDV's VIDETO

In contrast, ZGDV's VIDETO hides the complexity of MPEG-7 basing the description properties on a simple description template, which can then be mapped to MPEG-7 using XSLT. Domain-specific description templates together with their corresponding XSLT mappings are generated. The resulting flexibility, customisability and user-friendliness of this approach are VIDETO's biggest advantages. VIDETO was developed as a research tool to generate video (XML) metadata for testing a video server and retrieval module [7].

D. COALA – LogCreator

The LogCreator of the COALA project is a web-based tool which supports video descriptions. It offers automatic shot detection and a good interface for hierarchical segmentation of videos that can be uploaded to the server, where it is saved as MPEG-7 in a native XML database [8]. However, it is a domain-specific tool, developed specifically for TV news documents with a predefined structure. The descriptors that are used to annotate the different video segments are predefined as well.

E. CSIRO's CMWeb tools

Two other web-based video annotation systems are: CSIRO's Continuous Media (CM) Web Browser which generates a proprietary HTML-format Annodex file [9]; and Microsoft's Research Annotation System (MRAS) [8] – a Web-based application designed to enable students to asynchronously annotate web-based lecture videos and to share their annotations.

F. Microsoft MRAS

None of the systems described above are designed to be used within a collaborative video-conferencing environment. Microsoft's Distributed Tutored Video Instruction (DTVI) [10] system does allow students to replay and discuss videos of lectures collaboratively. However it does not support real-time synchronous annotations. It is also based on a combination of Windows Media Player and Microsoft's NetMeeting [11].

Net Meeting is based on the T.120 protocol [12] for application sharing. Because T.120 has been designed for low bandwidth and only supports low frame rates (e.g., 10fps), the capture and transfer of mouse events, keyboard events and screen update to the display devices of the participants is too slow to adequately handle high quality MPEG-2 video (24-30fps).

G. VIA

The Video and Image Annotation (VIA) tool has been developed by the MK-Lab within the BOEMIE project [13]. A snapshot of the interface of the tool, during a shot annotation of a video file is shown in Figure 1. The shot records a pole vaulter holding a pole and sprinting at the jump point.

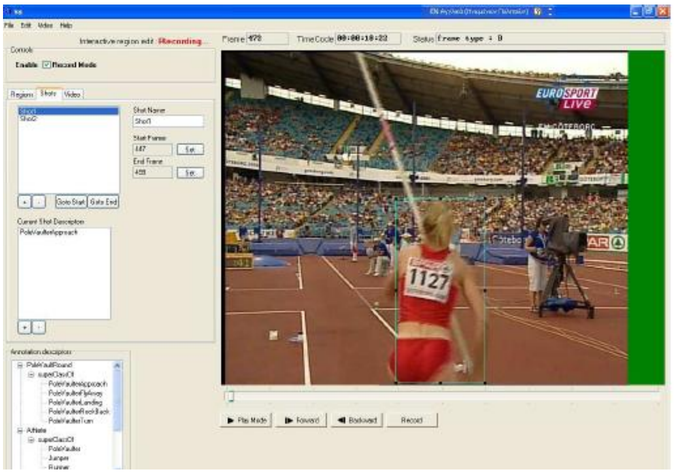

Fig. 1. Example of video annotation using VIA

VIA supports descriptive, structural and media metadata of image and video assets. Descriptive annotation is performed with respect to user loaded OWL ontology, while free text descriptions can also be added. Administrative metadata follow a customized schema internal to the tool, including information about the creator of the annotations, the date of the annotation creation, etc. A customized XML schema is also used for the representation of structural information, allowing for example to nest a video segment as part of a video and to define its start and end frame / time interval. The produced metadata can be exported either in XML or as in a more human readable format in textual format. Regarding image (and by consequence frame) annotation, the granularity levels supported include the entire image and specific still regions. The localisation of regions is performed either semi-automatically, providing the user a segmented image and allowing her to correct it by region merging, or manually, using one of the drawing functionalities provided, i.e. free hand, polygon, circle, rectangle. In the case of image annotation, the tool supports additionally the extraction of MPEG-7 visual descriptors per each annotated region, based on MPEG-7 XM [14], so the annotation outcome can be used as a training set for semantics extraction algorithms.

Regarding video annotation, the supported annotation granularity may refer respectively either to the entire video, video segments, moving regions, frames or even still regions within a frame. The annotation can be performed in real time, on MPEG-1 and MPEG-2 videos, using an interface consisting of three panels. The first one is concerned with region annotation, in which the user selects rectangular areas of the video content and subsequently adds corresponding annotations. The other two panels are used for annotation at shot and video level respectively. Shot boundaries are defined manually, by selecting its start and end frames. An important feature about region annotation is that the user can drag the selected region whereas at the same time the video is playing, so as to follow the movement of the desired region.

The annotations performed with VIA can be saved as annotation projects, so that the original video, the imported ontologies, and the annotations can be retrieved and updated at a later time. VIA is publicly available.

H. Vannotea – A Collaborative Video Indexing, Annotation and Discussion System For Broadband Networks

Vannotea is a tool for collaborative indexing, browsing, annotation and discussion of video content [15], developed by the University of Queensland. Vannotea's primary focus consists in providing support for collaborative, real-time, synchronous video conferencing services. Interoperability concerns, in conjunction with the requirements for simple and flexible annotations, led to the adoption of an XML-based description schemes. Building on a simplified translation of the respective MPEG-7 and Dublin Core descriptions, Vannotea metadata can be easily transformed into the corresponding standardised representations through the use of XSLT. It is worth noticing that Vannotea builds on the Annotea initiative, a W3C activity aiming to advance the sharing of metadata on the Web. Advocating W3C standards, Annotea adopts RDF based annotation schemes and XPointer for locating the annotations within the annotated resource.

I. VideoAnnEx

The IBM VideoAnnEx annotation tool [16] addresses video annotation with MPEG-7 metadata. Although the project within which VideoAnnEx was developed has finished and the tool is no longer maintained, VideoAnnEx is accessible and provides an illustrative case of content annotation in accordance to the MPEG-7 initiative.

Indexing and annotation systems for digital video files have been developed in the past - but only for use within stand-alone environments in which the annotations can be saved and shared asynchronously. Our first task was to carry out a detailed survey of these existing systems, determine their best and worst features and integrate the best features in a prototype which could be shared within a collaborative real-time high-quality video-conferencing environment.

J. Ontolog

Ontolog is a tool for annotating video and audio sources using structured sets of terms/concepts. It is a java application, designed and developed as part of a Ph.D. thesis in the Norwegian University of Science and Technology [17]. Though not maintained the past four years, the source code is available upon request. A screenshot of a video annotation process is shown in Figure 2.

Ontolog addresses various types of metadata, including descriptive, structural and administrative. Descriptive annotations are inserted according to one or more RDFS ontologies, imported or created by the user. The user can further enrich the subject matter descriptions by introducing additional properties. For the representation of administrative metadata, Ontolog provides by default two ontologies, namely the Dublin Core Element Set and the Dublin Core Qualified Element Set. Structural descriptions referring to video segments are created in correspondence with user-defined intervals, following the simplified structure representation defined in the Ontolog Schema[50] ontology. The produced annotations are in RDF.

Ontolog's interface consists of four components: a Media Panel, an Ontology Editor, a Logging Panel and a Property Editor. The media panel handles the video assets that are contained in an annotation project. For media loading either Quicktime (for Java) or the JMF framework can be used (and the corresponding media formats). The Ontology Editor provides mechanisms for the definition of concept hierarchies; properties defining relations between concepts can be specified in the Property Editor. Each property may optionally specify what kind of concept it may be applied to (domain) and what kind of values it may take (range).

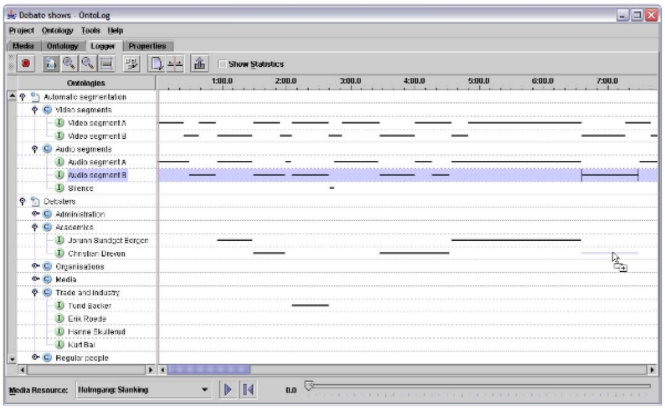

Fig. 2. Automatic segmentation in OntoLog

K. Anvil

Anvil is a tool that supports audiovisual content annotation, but which was primarily designed for linguistic purposes, in the same vein as the previously described tool. It was developed as part of a Ph.D. thesis at the Graduate College for Cognitive Science and the German Research Center for Artificial Intelligence (DFKI).

Anvil supports descriptive, structural and administrative annotations of video or audio objects that refer to the entire assets or to temporal segments of them. User-defined XML schema specification files provide the definition of the vocabulary used in the annotation procedure. The output is an XML file containing administrative information in its head segment, while its body includes the descriptive metadata along with structural information regarding the temporal localisation of the possible video segments. Recently, Anvil has been extended to support spatiotemporal annotation as well by allowing annotations to be attached to specific points; interpolation functionalities and arbitrary shapes constitute future extensions.

L. Semantic Video Annotation Suite

The Semantic Video Annotation Suite (SVAS), developed by Joanneum research Institute of Information Systems & Information Management, targets the creation of MPEG-7 video annotations.

VAS [18] encompasses two tools: the Media Analyzer, which extracts automatically structural information regarding shots and key-frames, and the Semantic Video Annotation Tool (SVAT), which allows to edit the structural metadata obtained through the Media Analyzer and to add administrative and descriptive metadata, in accordance with MPEG-7. The administrative metadata include information about the creator, the production date, the video title, shooting and camera details, and so forth.

IV. OUR SEMANTIC NAVIGATION APPROACH

For accessing to the semantic of resources contents, we have to realize semantic mediation between semantic of contents as described by the authors and semantic of queries as described by the users.

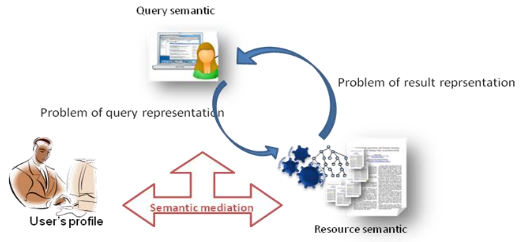

Fig. 3. General context of our approach

Our objective is, as mentioned inn figure 3, to define a model of knowledge representation from resources, users and services, and, at the second level, to map the content semantic and query semantic for presenting relevant resources and fragment resources via two kinds of services: semantic search and semantic navigation.

Our model is based on the result of the three important concepts, presented in figure 4:

- Semantic annotation and indexation of textual documents,
- Representation of a new model of semantic navigation,
- Offering intelligent services.

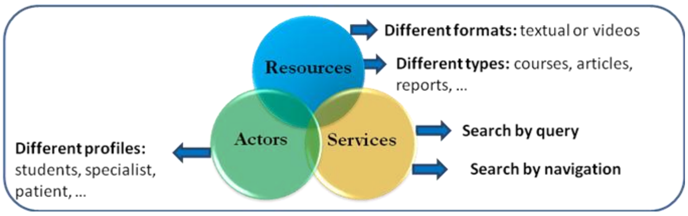

Fig. 4. Different conponents of our approach

A. Text annotation

The annotation process generates three types of annotations: descriptive annotations, contextual annotations and conceptual annotations.

- The descriptive annotations: metadata relative to title, authors, publication date, key-words, format, size, publisher, organization, language, format, country.
- The contextual annotations: metadata relative to the context of resources such as type of resources (practice case, course, recommendation guide, report), type of training (initial training, specialist training, preparation of competition or exam) and target public (student, surgeon, patient, specialist, general practitioner, researcher).

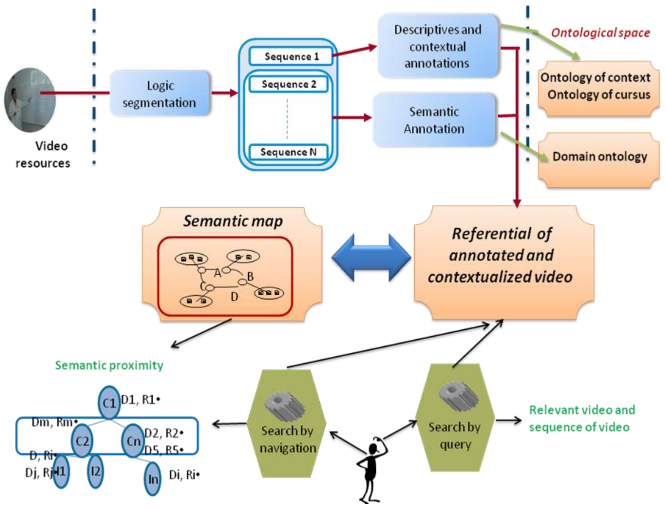

Fig. 5. General architecture of our approach

- The conceptual annotations: are relative to the major concepts and terms extracted from resources and their fragments. The concepts are presented with their degrees of representativeness CFIDF and the terms are presented with their degrees of pertinence TFIDF.

B. Video segmentation

If one of the defining properties of video is its temporal extent, then the treatment of time may be the most important part of a video content model. Videos may deal with many different topics at different times, and it is important for a content model to be able to reflect this. This is done by marking descriptions as valid only in certain temporal intervals – or, from the other point of view, divide the video into temporal fragments (logically, not necessarily physically) and attach descriptions to the fragments. A content model may put restrictions on these fragments – where their start and endpoints may be, their size, whether overlapping is allowed, whether adjacency is required, and so on. This dimension, we call the temporal expressiveness of the model.

With the term "segmentation", we mean logically dividing the video into adjacent, non-overlapping parts – a partition, in other words. Non-temporal (or content-independent) information is connected to the video as a whole, while temporal descriptions are connected to the relevant segment(s).

There are several paradigms for determining segment boundaries. Fixed segmentation entails that each segment has a fixed, content-independent size. The advantage of this approach is that it is simple – only the small, simple segment number needs to be stored to be able to compute the corresponding time interval, and the creation of segments requires no human interaction. The downside is that what happens in the video will probably not correspond very well to the arbitrary segmentation. Descriptions will be only partially valid if a segment contains changes in subject, and may need to be duplicated across segments if an interesting sequence straddles a segment boundary. As a simple example, consider annotating a video by marking the presence of different people in the picture: If a certain person is present in several consecutive segments, this has to be recorded separately for each segment; and if a person leaves the picture in the middle of a segment, the annotation is only partially correct.

Structure-based segmentation is a more common segmentation scheme: Using physical, structural properties of the video to determine segmentation boundaries.

Shot detection is an excellent candidate for this; audio analysis is another possibility.

This scheme is only marginally more complex than fixed segmentation, and has the advantage that the shots are often correlated to the semantic structure of the video. However, if the video (or significant parts of it) has no shots, or very many shots, this scheme has the same problems with redundant or partially invalid descriptions.

With user-defined segmentation, the user is free to choose the segmentation boundaries that are most useful to her, based on how we wish to describe the video – we can tune the segmentation to correspond exactly to the semantic content we want to model. The structure-based segmentation may be provided as a starting point, but segment boundaries can be set arbitrarily (as long as the subdivision still is a partition).

C. Video annotation

The more features that can be automatically understood in a video:

- Faces detected,
- Scenes recognized,
- Objects segmented.
- Descriptive annotation: Refers to the annotation dimension. For the purposes of this overview, we identify the following types:
    - content descriptive metadata addressing subject matter information,
    - structural metadata describing spatial, temporal and spatiotemporal decomposition aspects
    - media metadata referring to low-level features, and
    - administrative, covering descriptions regarding the creation date of the annotation, the annotation creator, etc.
- Contextual annotation: may refer to the entire video, temporal segments (shots), frames (temporal segments with zero duration), regions within frames, or even to moving regions, i.e. a region followed for a sequence of frames. It worths noting that due to the more complex structural patterns applicable for video, many tools besides the annotation functionality provide corresponding visualisation functionalities through the use of timelines. Thereby, the associations of subject matter annotations with respect to the video structure can be easily inspected.
- Semantic annotation : represent the content of the video as subject, object, time, etc.

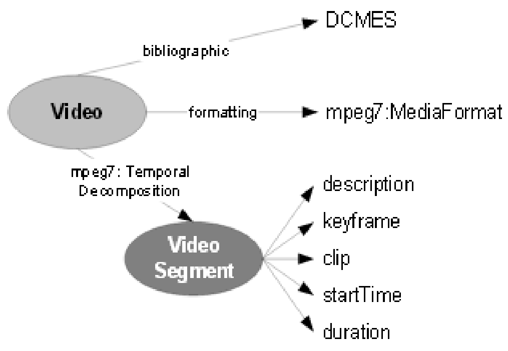

Fig. 6. A Generic Descriptive Metadata Model for Video

The annotation database stores the annotations (which may be associated with segments, keyframes or still regions within frames), as well as the source of the annotations (who, when, where). Annotations can be notes, explanations, or other types of external subjective remarks [].

Many video resources in the web have developed, and especially in the medical domain. We can find video representing courses, surgeon, experiences or recommendation guides. So, integration of these multimedia resources has been identified as important steps towards more efficient manipulation and retrieval of visual media.

Automatic annotation of video content at the semantic level has received a significant attention from the research community in the recent years, as a fundamental mean to face the explosive growth of video production and the associated growing request for search and retrieval by content of interesting elements.

In the last years many researchers have exploited ontologies to perform semantic annotation and retrieval from video. Ontologies are useful for semantic annotation of videos.

In this context, ontologies include a set of terms with their associated definitions that formally describe the application domain, through concepts, concept properties and relations, according to some particular view. Other ontologies provide structural and content-based description of multimedia data, similarly to the MPEG-7 standard.

Garcia and Celma [] have produced an OWL-Full ontology obtained through an automatic translation of MPEG-7; this approach has the limitation that computational complexity and decidability of reasoning are not guaranteed. Tsinaraki et al. [29] have manually developed an OWL-DL ontology that captures the full MPEG-7 Multimedia Description Schema (MDS) and the parts of the MPEG-7 video and audio schemas that are required for the complete representation of MDS.

OWL-DL ontology, designed to provide a high degree of axiomatization, ensuring interoperability through machine accessible semantics, and extensibility has been proposed. This ontology comprises parts of MPEG-7 descriptors such as visual low-level, spatiotemporal decomposition and media information descriptors.

D. Breast cancer Ontology

Recently ontologies have been regarded as an appropriate tool to bridge the semantic gap between the information that can be extracted from the visual data and the interpretation of the same visual data by a user in a given context. An ontology consists of concepts, concept proprieties, and their relationships to provide a formal description of a domain and provides a common vocabulary that overcome semantic heterogeneity of information.

Medical ontologies are developed to solve problems such as reusing and sharing of patient data, required of semantic-based queries/inference or the transmission of these data. The communication of complex and detailed medical concepts is a very important task in current medical information systems. In this way, more complex tools such as case-based retrieval or evidence-based medicine can be possible in medicine. A major weakness of usual ontological technologies is their inability to represent and to reason with uncertainty and imprecision [3]. However, medicine, being a science whose subject is people, is inherently a science of certainty, and mostly deals with uncertain knowledge and imprecise and vague information. In order to achieve maximum advantage from ontologies, we need an extension of ontologies, which has the capability of capturing uncertainty knowledge about concepts, properties and relations in domains to support reasoning with inaccurate information. Along this direction, researchers have attempted in the past to use different approaches on modeling uncertainty in ontologies. In this paper, we propose a new ontology-based mammography annotation system with a capability of uncertainty modeling in ontologies. To achieve this, we use Bayesian probability-based approach, without extending description logic and ontology languages. In addition, we also propose a rule-based BI-RADS score reasoning approach using Semantic Web technology of SQWRL (Semantic Query-enhanced Web Rule Language).

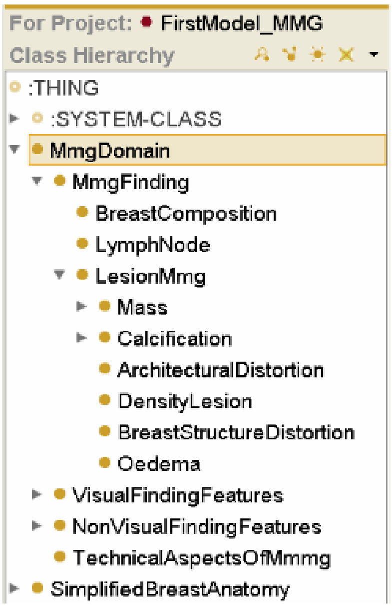

Fig. 7. Breast cancer ontology

E. Video ontology

To represent video objects, we have based on OVID (Object-oriented Video Information Database []) is a prototype system developed by Eitetsu Oomoto and Katsumi Tanaka. Part of their motivation was that meaningful scenes (temporal intervals) in a video document are identified incrementally and dynamically, according to various and changing user needs and domain requirements. It is not possible or desirable to a priori define an attribute schema that is suited for all kinds of scenes; each scene

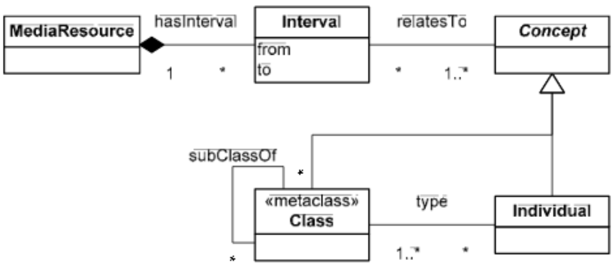

Fig. 8. An UML view of Video ontology

The MediaResource element represents the digital media objects, e.g. MPEG files. Each MediaResource contains an unbounded number of Intervals, with start time and end time. By default, no restrictions are put upon the temporal ordering of the intervals, so they may freely overlap.

The main mechanism is based on each Interval being connected (related) to one or more Concepts. There are two kinds of Concepts, classes and individuals, related to each other with specific semantics. Concepts may represent terms, topics, persons, places, events – anything that it is desirable to mark the presence of in the media object.

A class is a set (possibly unbounded) of values; as such, it is not much different from what is normally called a data type. However, unlike data types, the classes do not place any restrictions on the nature or representation of its values; a value is a member of a class simply by virtue of explicitly stating that it is. It follows that a value in general may be a member of multiple classes.

"Value" is a very general term. We therefore use the term individual for the user-defined values belonging to the user-defined classes used to annotate video. The "type" association be tween class and individual is the mechanism for individuals to state what class(es) they belong to.

Classes being sets, the "subClassOf" association is a subset relationship. It follows that members of a class are also members of its superclasses, and that a class may be a subclass of more than one class.

F. Semantic navigation

Our information representation and visualization model is based on the cartography paradigms studied in the state of the art. The aim is to reduce the cognitive effort of readers as regards to the classical result list representation mode. Indeed, graphical visualizations allow putting in evidence the pertinent information for users.

For representing and visualizing the information, we used a graph shaped representation based on the fisheye visualization techniques. This type of representation is adequate for representing semantic relations in the annotated domain ontology and the association networks (hierarchical relations between the concepts, the association relation between the concepts, the similarity relations between the documents, etc.). The fisheye technique allows putting in evidence the interest center of the user when he navigates in the graph.

The nodes represent concepts at the first level, terms at the second level, resources at the third level and segments or fragments in the last level. The arcs represent similarities between:

- Concept and concept
- Terms and resources
- Resources and segments or fragments

To measure similarity between concepts, we used CFIDF and between terms, we choose TFIDF. For the similarity between resources and segments, we calculated frequency of apparition.

In order to experiment our interactive visualization scenarios, we used the treebolic applet since it is based on graph representation [19].

Our new interaction mode offers to users a multi-approach of semantic navigation:

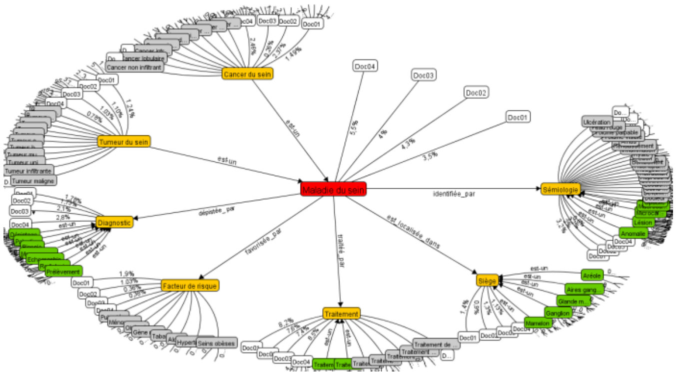

Fig. 9. Semantic map based on concepts

Domain ontology based navigation approach allowing users to make thematic searches to explore document informational space according to their themes of interest.

Concept association based navigation approach allowing the users to make connotative searches by navigating in the conceptual association graphs.

Similarity relation based navigation approach allowing users to make another type of connotative searches by navigating in the document similarity relation graphs.

The idea is to visualize the semantic content of the textual document corpus through a graphic representation of the annotated domain ontology. Initially the domain ontology is visualized as a hierarchy of themes and concepts, in which a user can navigate from one theme to another and from one concept to another in order to localize his interest center, mentioned in figure 10. For a given concept, the user can ask to display the titles of all documents indexed by this concept and to order them by their pertinence degree. The user can afterward consult the description of a document of his choice. This description represents a semantic summary of the selected document and contains descriptive, conceptual and thematic annotations already extracted during the annotation step.

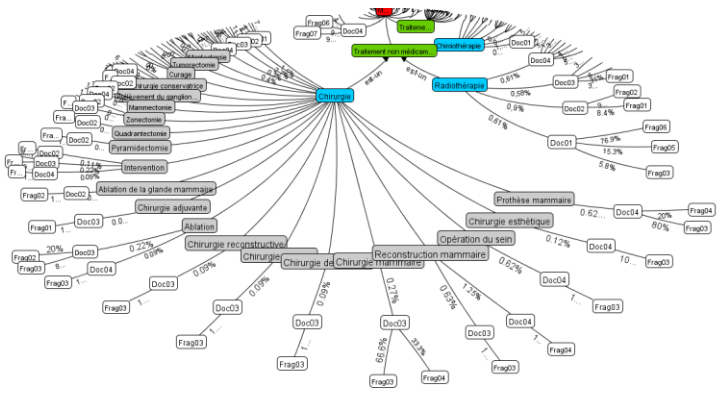

Fig. 10. Navigation in terms and resources

We decided to use OWL as the scripting language for constructing the ontology, and Protégé was used for visualising the semantic relationships. We also decided to use the Jena OWL API which is a Java framework for building Semantic Web applications and provides a programmatic environment for OWL.

The Kazuki extension to Jena was also used for handling concept instances within the generated ontology. Finally, we extended the well-established GATE text analyser system [20], by adding additional support for identifying, highlighting, and exporting the conceptual relationships identified within texts.

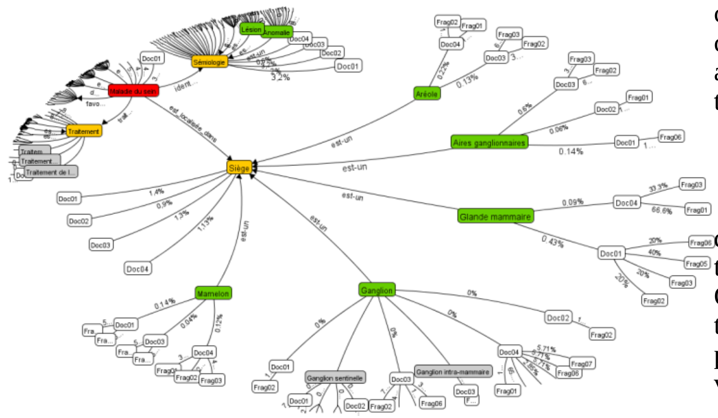

Fig. 11. Similarity measures between concepts

Several advantages ensue from this navigation approach. Effectively, this navigation approach offers a thematic search mode by reflecting for a given domain the semantic common to the majority of users. It offers to users a representation of knowledge close to the cognitive model which they have on the domain, what avoids them getting lost in the semantic map and allows them to localize quickly their interest center.

This navigation approach helps users answer these questions and minimize the problems of lost in information space and disorientation syndrome.

G. Semantic search

When navigating in the domain ontology, the user can focus his attention on a concept and wishes to know what are the concepts associated to it. The analysis of the conceptual association relations in the corpus allows answering this kind of needs. Our idea is to build for every concept an association graph allowing users to discover the association relations of their interest concept and to visualize documents relative to an association of their choice. For the identification of the conceptual association graphs we are based on the construction and the analysis of the cooccurrence networks.

So, for every concept of the ontology we determine the set of the concepts with which it is associated by a cooccurrence relation. We measure the degree of association of every relation according to the number of documents in which both concepts collocate. The analysis of the cooccurrence relations between concepts allows to index documents by conceptual associations.

The second level of nodes represents documents indexed by these conceptual associations. The label of an edge which connects a document node and an associated concept node represents the relevance degree of the document with regard to both concepts (associated concept and central concept).

The main interest of integrating conceptual association relations in the visualization process is to allow users to discover information related to their initial interest center what contributes to enlarge their domain knowledge. Besides, the visualization of association relations allows reflecting the real context in which concepts are evoked in documents. So users could refine their search according to the conceptual associations which are relevant to them (filter documents) and to discover new knowledge.

V. EXPERIMENTATION AND RESULTS

We choose medical domain for experimentation because doctors and radiologists are often exposed to high pressure of time and pressure to perform in the present clinical processes. One of these processes in the radiology is the examination of three-dimensional medical images. However, there are two problems that occur continuously on interactions with these volume images.

Diagnosis of breast cancer normally involves multi-disciplinary meetings with experts from different medical backgrounds, e.g. radiologists, surgeons, oncologists, histologists and other clinical staff. A typical procedure for cancer assessment starts with a report from routine X-ray check or a self-report of abnormal symptoms followed by a X-ray scan. X-ray mammography is thus an obvious starting point for the knowledge modelling effort [21]. In this paper, we present features of the mammographic ontology, the conceptual issues faced and the lessons learnt in the process.

We have obtained medical imaging data and records to evaluate our system from the University of South Florida digital mammography dataset [22]. The cases have been annotated and inserted to the repository according to the available explanations. The evaluation of the system has been based on submitting test quires to the system to observe relevancy of results in each of the systems.

For example, the keyword search for "Microcalcification" returned 39 images. All of which were examined, 24 images amongst them described some sort of relevancy to "Microcalcification", one image had "Necrosis" which is relevant to "Microcalcification", and 14 images had annotations that contained "No Microcalcifications" or "No Suspicious clusters of Microcalcification".

The term "Microcalcification" was selected from the ontology view and the system was able to retrieve two direct relevant images, both of which were from the same medical case. A search based on the retrieved resources (semantic browsing) was performed. The test results evaluated for the "Disorder", "Treatment" and "Chemical" attributes. The system was able to retrieve several other related resources based on similar treatment type and also the cases which shared the same disorder but used different treatment and also other attributes.

There were related cases which had same treatment type, but with a different disorder. This type of results helps the medical expert to have broader view of the case and enables the users to see and compare the relevancies based on different criteria.

However, semantic browsing functionality of the system, with referring to meaningful relationships specified in the ontology, enables the users to retrieve relevant resources based on semantic relations. Another important advantage of the developed system is providing a unified vocabulary to annotate and also to search for the resources. Different medical experts could have different terminology for annotating and searching the resources. It would be difficult to formulate a query that matches all the keywords used by several medical experts. Using an ontology as a standard controlled vocabulary provides a unified meta-data representation that is used in information search and retrieval process.

VI. CONCLUSION

The principal idea of this work is to propose a model allowing to put in evidence semantic inherent to the heterogeneous corpus. Our model is based on the result of the semantic annotation and indexation of textual and video documents, and represents a new model of graphic visualization and semantic navigation.

The annotation process generates three types of annotations: descriptive annotations, conceptual annotations and thematic annotations.

One of the biggest challenges of the visualization conception is that there is no strategy of "ideal" visualization; the conception is always specific to the application. Different systems are efficient for users having different backgrounds and needs (expert or novice, scientist or general information). A universal model is difficult to be generalized.

So, we presented an approach to semantic navigation based on multiple ontologies. Our aim is to support users with semantic descriptions during some their tasks, as modeling, presentation and selection of knowledge classified in documents and according to their needs. The integration and distinction of different topic ontologies allow for effective richness and variety of views on a same domain, felt as a crucial aspect within an organization.

We concluded that semantic text and video annotation constitute particularly active research fields, faced with intricate challenges. Such challenges issue not only from implications related to the sheer volume of content available, but also from the dynamically evolving context of intelligent content management services as delineated by the growth of Semantic Web technologies, as well as by new powerful and exciting concepts introduced by initiatives such as Web 2:0,

As the first perspective, The evaluation of information visualization is a very problematic task. Several challenges could rise when researchers conduct an information visualization evaluation. These challenges can be related to many factors: the context of use, participant gathering, data collection, existence of evaluation environment (standard, reference tool for comparison, etc.).

As another perspective, we plan to construct a framework allowing users to exploit different kind of resources: test, video and image, select interactively the visualization paradigm to be used in their maps and to make conversion between visualization paradigms if they are not satisfied.

References

[1] M. Koskela and A.F. Smeaton. Clustering-Based Analysis of Semantic Concept Models for Video Shots In Proc. IEEE International Conference on Multimedia & Expo (ICME 2006). Toronto, Canada. July 2006.

[2] A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[3] IBM MPEG-7 Annotation Tool. http://www.alphaworks.ibm.com/tech/videoannex

[4] Ricoh MovieTool. http://www.ricoh.co.jp/src/multimedia/MovieTool/

[5] Zentrum fuer Graphische Datenverarbeitung e.V. (ZGDV). VIDETO - Video Description Tool. http://www.rostock.zgdv.de/ZGDV/Abteilungen/zr2/Produkte/videto/index_html_en

[6] Swiss Federal Institute of Technology (EPFL). COALA (Content-Oriented Audiovisual Library Access) – LogCreator. http://coala.epfl.ch/demos/demosFrameset.html

[7] CSIRO. The Continuous Media Web (CMWeb). http://www.cmis.csiro.au/cmweb/

[8] David Bargeron, Anoop Gupta, Jonathan Grudin, and Elizabeth Sanocki. "Annotations for Streaming Video on the Web: System Design and Usage Studies". Microsoft Research, Redmond. http://www.research.microsoft.com/research/coet/MRAS/WWW8/paper.htm

[9] C. Tricot, Cartographie sémantique, des connaissances à la carte, thèse de doctorat, Université Savoie, 2006.

[10] http://mklab.iti.gr/project/via

[11] MPEG-7 XM: MPEG-7 Visual eXperimentation Model (XM), Version 10.0, Doc. N4062. ISO/IEC/JTC1/SC29/WG11 (2001)

[12] http://www.itee.uq.edu.au/reresearch/projects/vannotea/index.html

[13] http://www.research.ibm.com/VideoAnnEx/index.html

[14] http://www.idi.ntnu.no/ heggland/ontolog/

[15] http://www.joanneum.at/en/fb2/iis/products-solutions-services/semantic-video-annotation.html

[16] R. Rao, S.K. Card, The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus+Context Visualization for Tabular Information, in Beth, A., Susan, D. et Judith, S.O. (eds.), ACM Press, 1994, p.318-322.

[17] Jon Heggland. OntoLog :Flexible Management of semantic Video centent annotations; thesis, Faculty of Information Technology, Mathematics and Electrical Engineering, 2005.

[18] Relationships, In IEEE Transactions on Visualization and Computer Graphics, volume 8, pages 9–20.

[19] http://treebolic.sourceforge.net/treebolic1/fr/home.htm

[20] Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V., Aswani, N., Roberts, I., Gorrel, G., Funk, A., Roberts, A., Damljanovic, D., Heitz, T., Greenwood, M. A., Saggion, H., Petrak, J., LI, Y. and Peters, W. (2011). Text Processing with GATE (Version 6).

[21] American College of Radiology. Breast Imaging Reporting and Data System: BI-RADS, third edition.

[22] Digital Database for Screening Mammography (DDSM). Available at: http://marathon.csee.usf.edu/Mammography/Database.html

[23] MedPix, Free Online Medical Image Database, Available at: http://rad.usuhs.mil/medpix/medpix.html