A BOOTSTRAPPING APPROACH TO KNOWLEDGE ACQUISITION FROM MULTIMEDIA CONTENT WITH ONTOLOGY EVOLUTION

Vangelis Karkaletsis, Georgios Paliouras, Constantine D. Spyropoulos

Software and Knowledge Engineering Laboratory,
Institute of Informatics and Telecommunications,
National Centre for Scientific Research "Demokritos",
15310 Aghia Paraskevi Attikis, Athens, Greece
{vangelis, paliourg, costass}@iit.demokritos.gr

ABSTRACT

We propose a bootstrapping approach to knowledge acquisition, which uses multimedia ontologies for fused extraction of semantics from multiple modalities, and feeds back the extracted information, aiming to automate the ontology evolution process. This paper presents the basic components of the proposed approach and discusses the open research issues focusing on the synergy of extraction and evolution that will enable the development of scalable and precise knowledge acquisition technology.

1. INTRODUCTION

The potential in the use of ontologies for extracting semantic information from multimedia content has been highlighted in many recent research articles. However, efficient and effective methodologies for information extraction from multimedia content using ontologies are still missing. Meanwhile, the potential of an iterative approach to ontology evolution, combining information retrieval and extraction techniques for textual content has also been proposed in the literature. On this basis, we propose a specific bootstrapping approach to knowledge acquisition, which uses multimedia ontologies for fused extraction of semantics from multiple modalities, and feeds back the extracted information, aiming to automate the ontology evolution process.

The proposed approach is unique in that it links multimedia extraction with ontology evolution, creating a synergy of enormous yet unrealized potential. In recent years, significant advances have been made in the area of automatic extraction of low-level features from visual content. However, little progress has been achieved in the identification of high-level semantic features or the effective combination of semantic features derived from different modalities. Driven by domain-specific multimedia ontologies, the information extraction systems implementing the proposed approach will be able to identify high-level semantic features in image, video, audio and text and fuse these features for optimal extraction. The ontologies will be continuously populated and enriched using the extracted semantic content. This is a bootstrapping process, since the enriched ontologies will in turn be used to drive the multimedia information extraction system.

Section 2 highlights the state of the art in each of the major technologies involved. Section 3 presents the main aspects of the proposed approach, the architecture designed for its implementation and the basic components of the architecture. Section 4 outlines an application scenario we are currently examining for the evaluation of the proposed approach. Section 5 discusses some of the issues that arise under this bootstrapping framework and need to be searched. The paper concludes presenting our next steps for the implementation of the proposed approach.

2. STATE OF THE ART

The proposed approach towards the automation of knowledge acquisition from multimedia content, through ontology evolution, is based on the synergy of various technologies. This section highlights the state of the art in each of the technologies involved.

2.1. Semantics extraction from multimedia content

Semantics extraction from multimedia content is the process of assigning conceptual labels to either complete multimedia documents or entities identified therein. In general, extraction can be performed at three different levels:

- Layout: the syntactic structure an author uses for multimedia documents (camera shots, audio segments, text syntax).
- Content: relates layout segments to elements that an author uses to create in a multimedia document.
- Semantics: expresses the intended meaning of the author.

In the case where content is available in multiple related modalities, these can be combined for the extraction of semantics. The combination of modalities may serve as a verification method, a method compensating for inaccuracies, or as an additional information source [1]. The processing cycle of combination methods may be iterated allowing for incremental use of context. The major open issues in the combination approaches concern the efficient utilization of prior knowledge, the specification of open architecture for the integration of information from multiple sources and the use of inference tools for efficient retrieval.

Most of the multimedia extraction approaches encountered in the literature are based on learning methods, e.g. naive Bayes classifiers, decision tree induction, k-Nearest neighbour, Hidden Markov model [2, 3]. However, with the advent of promising methodologies in multimedia ontology engineering, knowledge-based approaches are expected to gain in popularity and be combined with the machine learning methods. This is also the case we will study in the proposed approach.

2.2. Multimedia Ontologies

Ontologies can play a major role in multimedia content interpretation because they can provide high-level semantic information that helps disambiguating the labels assigned to multimedia objects. Indicative approaches for constructing multimedia ontologies are the ones presented in [4, 5, 6]. The major open issues here concern the automatic mapping between low level audio-visual features and high level domain concepts, the automated population from unconstrained content and when there are no metadata attached to the content.

Constructing ontologies is a non-trivial process that involves the following main steps: (a) selection of concepts to be included in the ontology, (b) specification of concept properties and relationships between concepts, (c) addition of concept instances. Steps (a) and (b) together are usually termed ontology enrichment, while step (c) is termed ontology population. Given the difficulty of constructing ontologies manually, various approaches to assist the experts in this task have been proposed. The most recent of these approaches involve the use of machine learning techniques to identify regularities, which could lead to interesting concepts and relations. These approaches are almost exclusively designed for unstructured natural language text or semi-structured HTML pages [7, 8]. The extent to which the text is structured determines the depth of the required linguistic analysis, in order to identify concept instances. Very little work has been done on other modalities [9, 10].

In cases of complex domains, multiple ontologies may be present and ontology coordination techniques [11, 12] have to be employed. To enable ontology coordination and enrichment with a multitude of homogeneous and heterogeneous ontologies, appropriate matching techniques are required to determine semantic mappings between concepts of different ontologies that are semantically related. Matching can be syntactic, structural or semantic, according to the type of knowledge used. The problem of ontology matching plays a crucial role in the ontology engineering process at two levels: a) ontology merging in ontology design in order to define a new ontology using a set of previously defined ontologies, in particular in the case of homogeneous ontologies; b) ontology alignment: in a multi-ontology environment a main problem is to deal with heterogeneous ontologies that have to maintain their heterogeneity. The solution is to establish mappings between ontologies, simultaneously maintaining their original specification. A survey of the most relevant methodologies and methods used for ontology matching can be found in [13].

2.3. Synergy between information extraction and ontologies

The interaction between information extraction and ontology learning has also been modelled at a methodological level as a bootstrapping process that aims to improve both the conceptual model and the extraction system through iterative refinement.

In [14] the bootstrapping process starts with an information extraction system that uses a domain ontology. The system is used to extract information from text. This information is examined by an expert, who may decide to modify the ontology accordingly. The new ontology is used for further information extraction and ontology enrichment. Brewster et al. [15] propose a slightly different approach to the bootstrapping process. Starting with a seed ontology, usually small, a number of concept instances are identified in the text. An expert separates these as examples and counter-examples which are then used to learn extraction patterns. These patterns are used to extract new concept instances and the expert is asked to re-assess these. When no new instances can be identified, the expert examines the extracted information and may decide to update the ontology and restart the process.

3. METHODOLOGY AND ARCHITECTURE

We advocate an ontology-driven multimedia content analysis (semantics extraction from images, video, text, audio/speech) through a novel synergistic method that combines multimedia extraction and ontology evolution in a bootstrapping fashion. This method involving on the one hand, the continuous extraction of knowledge from multimedia content sources in order to populate and enrich the ontologies and, on the other hand, the deployment of these ontologies to enhance the robustness of the multimedia information extraction system.

Our main goal is to achieve, through the proposed synergistic approach, large-scale and precise knowledge acquisition from multimedia content. In order to achieve this, we propose the development of innovative methodologies and toolkits in ontology evolution and information extraction. More specifically, on the side of ontology evolution, we propose:

- A unifying representation for multimedia ontologies and related knowledge. This "*multimedia semantic model*" will link domain-specific ontologies, in which concepts are represented by domain-specific terms, with multimedia content and descriptor ontologies that represent content structure in multimedia documents and describe characteristics of multimedia objects in terms of low-level features and structural descriptions.

- A methodology for ontology evolution to coordinate the various tools that will use the extracted data to populate and enrich the ontologies. The resulting architecture will use the representation for the multimedia ontologies and will specify the interfaces for the various types of tools for ontology evolution.
- A toolkit for ontology evolution. Based on the specification of the ontology evolution architecture, tools will be developed to support ontology learning, ontology merging and alignment, semantic inference for consistency maintenance, and ontology management.

On the side of information extraction, we propose:

- A methodology and an open architecture for information extraction from multimedia content using data fusion techniques. The proposed methodology will specify how information from the multimedia semantic model can be used to achieve extraction from various media (text, image, video and audio). Additionally, it will combine extracted information from multiple media, using a probabilistic evidence-based framework, in order to improve the extraction performance.
- A toolkit for semantic extraction from multimedia content. Within the extraction architecture, tools will be developed to support extraction from image, audio, video and text, as well as information fusion.

We also propose an open architecture of a system that integrates the components for ontology evolution and semantics extraction in order to realise the synergistic bootstrapping approach. As it is depicted in Figure 1, the major components of the proposed architecture include:

- The multimedia ontology which links domain-specific ontologies with multimedia content and descriptor ontologies. This will be evolving through the ontology evolution component. An ontology initialization tool will be developed to provide a friendly user interface for the creation of the initial ontology.
- The semantics extraction component which will provide tools for the analysis of single modalities (visual, text and audio extraction tools) as well as tools for fusing information from multiple media sources (information fusion tools). The whole extraction process will be ontology driven in the sense that the ontology will provide the initial knowledge to the extraction process and will also be used to disambiguate the extraction results.

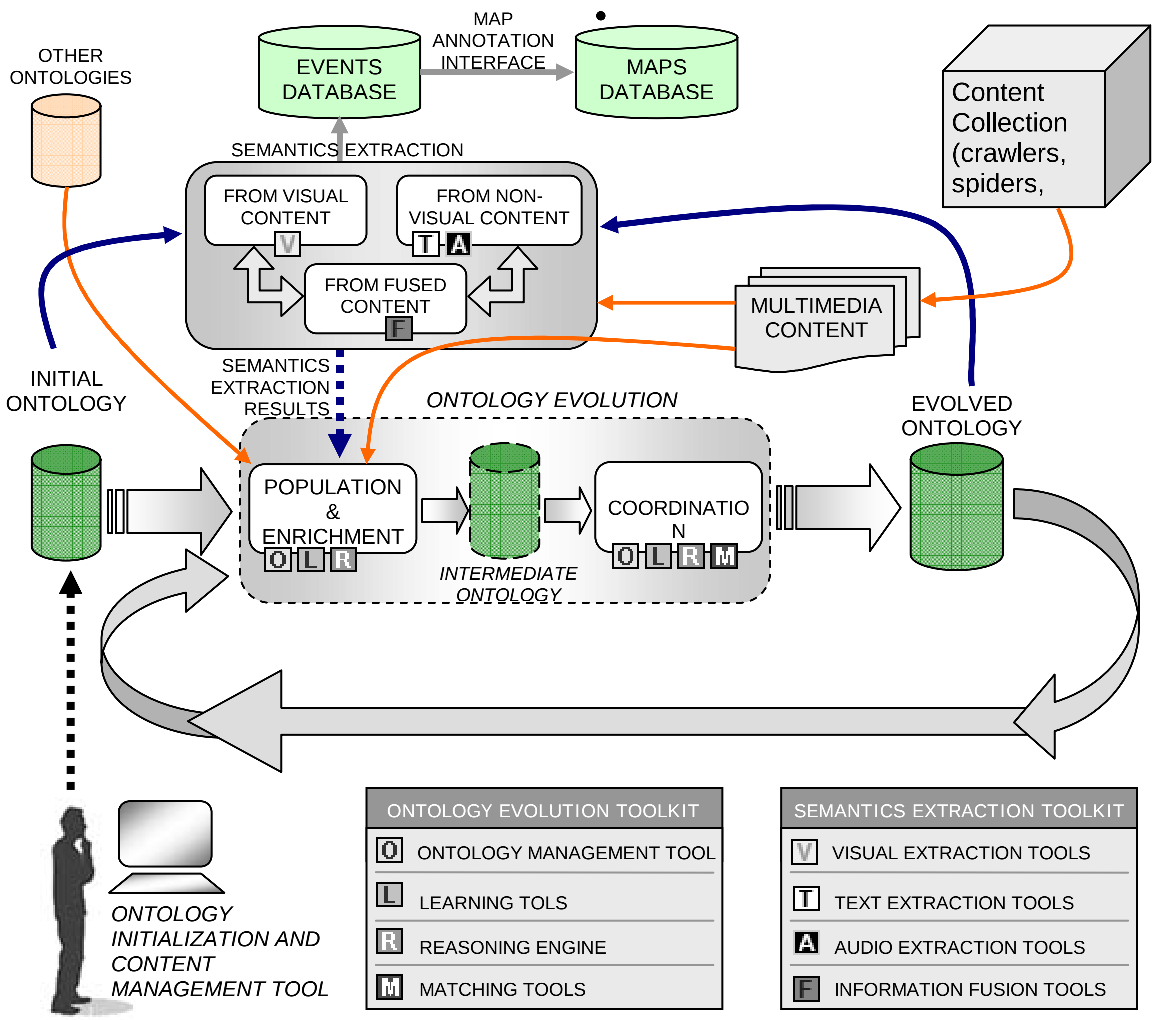

Figure 1. Architecture of the integrated system

- The ontology evolution component which will use the results of the extraction process to populate the multimedia ontology with instances of the various concepts, to enrich the ontology with new concepts and relations, as well as to coordinate the ontologies composing the multimedia ontology. Special emphasis will be given to the maintenance of semantic consistency, since any changes may generate inconsistencies in other parts of the same ontology, in the linked ontologies or in the annotated content.

In the following sub-sections more details are provided on the features of the above components.

3.1. Semantics Extraction from Multimedia Content

The high complexity that characterises the multimedia content along with the lack of precise modelling for multimedia concepts makes automatic semantics extraction a very difficult and challenging task. Although latest advances in multimedia content analysis have improved capabilities for effective searching and filtering, a gap still remains between the low-level feature descriptions that can be automatically extracted such as colours, textures, shapes, motions, and high-level semantic descriptions of concepts like objects, scenes and events that set the basis for meaningful multimedia content description.

A suitable approach to fill this gap is to use a semantic model in the extraction process. Moreover, the analysis of single modalities, in particular of visual content alone, is inadequate in all but a small number of restricted cases. The effort required to provide problem-specific extraction tools makes single-media solutions non-scalable, while their precision is also rarely adequate.

In the proposed approach, on the level of individual modalities, particular emphasis will be given to visual content, from images and video, due to the richness of this source and corresponding difficulty of extracting useful information. Non-visual content, audio/speech and text, will provide supportive evidence, in order to improve extraction precision.

Since no single modality is powerful enough to encompass all aspects of the content and identify concepts precisely, fusing information from multiple media sources is needed. Investigation into fusion techniques for data from heterogeneous sources using an ontology-based approach will be carried out and a toolkit for semantics extraction from fused multimedia content will be developed, based on the following components:

- Ontology-based disambiguation: The results of the extraction process are used to build hypotheses. These will be then matched to information included in the multimedia ontology to build more precise higher-level hypotheses. During this process, conflicts or inconsistencies may be found, prompting the revision of intermediate results, and, possibly, the adjustment of parameters for low-level processing modules to achieve more precise results at higher levels. This will be implemented as a closed-loop extraction process.
- Probabilistic evidence-based interpretation framework: As single-modality extraction is prone to imprecision and uncertainty, a probabilistic evidence-based interpretation framework will be developed. The evidence in this case will be provided by low-level features and the extracted semantics may belong to intermediate levels of abstraction.

3.2. Multimedia Ontologies

Although significant progress has been made in recent years on automatic segmentation or structuring of multimedia content and recognition of low-level features within such content, comparatively little progress has been made on the construction of ontologies that contain multimedia information.

In our approach, we propose the development of a unifying representation for multimedia ontologies and related knowledge. This “multimedia semantic model” will serve as an integrated model for the different ontologies that are necessary to support the semantics extraction process:

- Multimedia content ontology: It represents the structure of the content of the multimedia documents. The top level hierarchy of a multimedia document is classified into: Image, Video, Audio, Audiovisual and Multimedia. Each of these types has its own segment subclasses. These subclasses describe the specific types of multimedia segments, such as video segments, moving regions, still regions and mosaics.
- Multimedia descriptor ontology: This ontology models concepts and properties that describe visual characteristics of objects, especially the visualizations in still images and videos in terms of low-level features and media structure descriptions. Sub-concepts will include MPEG-7 standard features like colour, shape, texture, motion, localization and basic descriptors.
- Domain-specific ontologies: These ontologies contain concepts and properties related to the knowledge of the domain of interest (see Section 4 “Application scenario”). In these concepts we assign instances, which are used to recognize semantic objects using the results of the content analysis process. These ontologies also contain detailed descriptions of objects using spatiotemporal and partonomic relations defined in the multimedia semantic model.

Support is needed for an initial extraction of characteristics of multimedia objects as instances of the multimedia descriptor ontology and linking of these descriptors with actual concepts of the domain-specific ontology in a user friendly-way. This can be considered as an initialization phase, where descriptors are linked to the concepts of the domain ontologies and these multimedia ontologies will be used later for the analysis of new content and the evolution procedure. Therefore, a specific tool must be developed to provide a friendly user interface allowing the management of multimedia content, the extraction of visual descriptor instances and their linking with domain ontology instances.

3.3. Evolution of Multimedia Ontologies

According to [16] ontology evolution is "*the timely adaptation of an ontology to the arisen changes and the consistent propagation of these changes to dependent artefacts*". Thus, ontology evolution is a complex process, involving the following sub-processes:

- ontology population and enrichment, i.e., addition and deletion of concepts, relations, properties and instances,
- coordination of homogeneous ontologies, e.g. when more than one ontologies for the same domain are available, and heterogeneous ontologies, e.g. updating the links between a modified domain ontology and a multimedia descriptor ontology,
- maintenance of semantic consistency, since any of the above changes may generate inconsistencies in other parts of the same ontology, in the linked ontologies or in the annotated content base.

Our approach for ontology population and enrichment will be based on machine learning techniques using the information from the semantics extraction process. More specifically, the extraction process will populate the ontologies with instances of the various concepts, together with their properties and will also provide unclassified entities extracted from the multimedia content which may lead to suggestions for the enrichment of the ontologies with new concepts and relations, through novelty detection. This novelty detection is based on information from all different types of media being processed.

Ontology coordination approaches will be devised to interlink ontologies with different levels of heterogeneity. Ontology coordination involves the use of matching techniques and tools for mapping, alignment and merging. In particular, merging techniques will be used for coordinating homogeneous ontologies, e.g. two ontologies for the same domain, while alignment/mapping techniques will be used for coordination of heterogeneous ontologies, e.g. link a domain-specific ontology with the multimedia content ontology. We will also use machine learning techniques in order to detect similarities between concepts exploiting features in syntactic, structural and semantic level.

During ontology evolution, any of the changes may generate inconsistencies in other parts of the same ontology, in the linked ontologies or in the annotated resources. At the current state of the art, description logic reasoning systems (e.g. RACER [17]) are not tailored to these "incremental changes". We will investigate how such changes can be much more efficiently supported. The aim is the development of models, techniques, and tools for semantic consistency checking of ontology content throughout the evolution process.

An ontology management tool will be developed to provide a user-friendly interface for managing the multimedia ontologies. This will be based on existing ontology editors. It will also control ontology versioning, maintaining the different versions of the ontology throughout the evolution process.

4. APPLICATION SCENARIO

The application we are currently examining for the evaluation of the proposed bootstrapping approach concerns the enrichment of digital maps with semantic information. In other words, the results of the semantics extraction process will be displayed to the end-user, through an interactive digital map. The specific application scenario we have in mind involves an automatic content collection and annotation service for public events in a number of major cities. The domain of public events includes commercial exhibitions, sport events, concerts etc.

This service will continuously collect and annotate large amounts of dynamic content, from the Web and proprietary sources. The results of the annotation process, i.e., the identified entities and their properties, will be linked to geographical locations and stored in a content server. The application can be considered as a monitoring service for public events in a number of big cities. The user will be provided with immediate access to the annotated content base, through the user-friendly interface of digital maps, which will also provide immediate navigation guidance to the place of interest. The domain-dependent semantic model will be used by the extraction architecture to identify multimedia information related to the concepts in the ontologies. Further, from the extracted information, new concepts will be generated to extend the ontologies, using the evolution architecture.

As a concrete example of the application scenario, consider the domain of vehicle exhibitions. Given such a domain, the following stages will be followed to customize and use the envisaged system (see Fig. 1):

- Intialization: We will start by collecting, extending and merging existing ontologies for sub-domains referring, for example, to car, motorcycle and yacht exhibitions. These ontologies will also be linked to the appropriate multimedia descriptor ontologies. This process will be accomplished using the ontology initialization and content annotation tool and will result to the initial multimedia semantic model for the domain.
- Training: The various semantics extraction and ontology evolution tools are trainable to the domain. Therefore, a training dataset needs to be collected and used to customise the system. This training set should contain representative and annotated multimedia content, as expected to be encountered by the system at run time. The ontology initialization and content annotation tool will be used for the construction of the training set.

- Information gathering: Having customised the system, the first step of its run-time use is to collect content from various Web and proprietary sources. In the case of car exhibitions, such sources will include TV and news programmes, on-line magazines, the sites or proprietary databases of car suppliers and dealers, specialized discussion fora and Weblogs, as well as generic content sources. Different sources will provide different types of content, which when combined under the fused semantics extraction approach can lead to a rich description of concepts and their instances.
- Semantics extraction: The trained semantics extraction tools will be applied at regular intervals to the incoming stream of multimedia content, performing two parallel tasks: (a) extracting the relevant information from each piece of content, such as event venue, event dates, event organisers, exhibits and multimedia related to it, exhibitors, etc., (b) modelling the various concepts in the ontologies, by identifying their characteristic elements in associated multimedia content, e.g. terms in the text and audio, objects in visual content, crowd density, lighting, indoor/outdoor characteristics, etc.
- Ontology evolution: The former, extraction task, performed in the previous step, will populate the ontologies with instances of the various concepts, together with their properties. This process will also be accompanied by the appropriate annotation of content in the server, in order to provide semantic access to the content by the end-user. The latter, concept modelling task, performed by the extraction methods, will lead to suggestions for the enrichment of the ontologies, through novelty detection. For example, if new content about a new type of vehicle exhibition, e.g. van and truck exhibitions, becomes available, new concepts and relations will have to be added to the semantic model. This novelty detection is based on information from all different types of media being processed. For example, van and truck exhibitions will share several, but not all, visual and audio characteristics with car exhibitions. This notion of similarity will drive the evolution process.
- Information positioning and retrieval: The concept instances annotated on the multimedia content will be linked to the map data in the digital maps server. A user-friendly interface will allow the user to navigate to the various exhibitions in a city and retrieve appropriate content. Furthermore, the user will be able to issue queries about exhibits, events, etc. The results will always be associated to places on the digital map. Time plays an important role, as different exhibitions take place in the same venue at different time periods.

Although the scenario focuses on exhibitions it is easily extensible to other types of public event, like sport events, cultural events or political/news events. Furthermore, the basic bootstrapping approach is applicable to a variety of domains, other than public events, such as natural sciences, e.g. tracking of animal species and individuals, traffic control, etc. Exhibitions are proposed as an application that has significant commercial and social interest, while at the same time it is associated with a wealth of complementary multimedia content that is evolving over time.

5. DISCUSSION

The synergy of ontology evolution and multimedia information extraction under a bootstrapping framework seems to be a promising path towards the automation of knowledge acquisition from multimedia content. However, there are various issues that arise under this framework and need to be researched, in order to enable the development of scalable and precise knowledge acquisition technology. Some of these issues are raised in this section.

In terms of semantics extraction from multimedia content, we propose the integration of an ontology-based approach with a probabilistic inference scheme. We need to examine carefully the role of the ontology in fusing information extracted from multiple media. We also have to examine new ways to learn the optimal combination of features derived from multimedia content. Synchronization and alignment of the different modalities is another issue, since all modalities must refer to a common timeline.

Ontologies must be sufficiently expressive in order to describe the construction space for possible interpretations in general and for specific interpretation results in terms of a particular piece of media. Current ontology description languages (such as OWL) do not provide specific representation constructs, for instance, for temporal information which is very important for the specific application we are currently examining. We need to study how existing ontology languages such as OWL can be extended to more adequately deal with temporal information. In addition, multimedia applications have highlighted the need to extend representation languages with capabilities which allow for the treatment of the inherent imprecision in multimedia object representation, matching, detection and retrieval. Existing standard web languages do not provide such capabilities. Therefore, considerable research effort needs to be directed towards representation and management of uncertainty, imprecision and vague knowledge that exists in real life applications.

In terms of ontology population and enrichment, we will exploit the multimedia semantic model as well as current research on learning and inference techniques aiming to develop a generic framework for ontology learning and inference from multimedia content. We need to examine more carefully this task due to the complexities introduced by the multimedia context. Addition of instances in the multimedia descriptor ontology may also require updating the corresponding link with the domain-specific ontology. The semantics extraction process will provide unclassified entities extracted from the multimedia content which may lead to the enrichment of the ontologies with new concepts and relations based on information from all different types of media being processed. Concerning inference techniques for ontology population and enrichment, we need to optimize and enhance description logic inference technology such that learning and retrieval requirements are optimally supported. Furthermore, extensions with respect to spatial data and conceptual data models involving space and time need to be developed.

The task of ontology coordination in a multimedia context must also be examined carefully since it involves the coordination of both homogeneous and heterogeneous ontologies. We also propose the use of machine learning techniques to assist coordination in this context and we need to investigate which methods are most appropriate for such a task. This depends very much on the type of training data that is available. Supervised learning of complex representations requires data that may not be possible to acquire manually. Unsupervised or partially supervised methods may prove more useful in these cases.

Concerning semantic consistency checking in ontology evolution, two main problems are involved when considering multimedia ontologies and the bootstrapping process adopted, which we need to examine carefully. The first problem occurs at the instance level and requires techniques for efficiently handling incremental additions of instances, while checking integrity constraints. A second problem occurs at the concept level and requires techniques for checking the consistency of new concepts against the current ontology, in order to choose a valid and consistent enrichment solution among a set of possible alternatives.

6. CONCLUDING REMARKS

We proposed a new approach towards automation of knowledge acquisition from multimedia content, by introducing the notion of evolving multimedia ontologies which will be used for the extraction of information from multimedia content. This is a synergistic approach since it combines multimedia extraction and ontology evolution in a bootstrapping process involving, on the one hand, the continuous extraction of semantic information from multimedia content in order to populate and enrich the ontologies and, on the other hand, the deployment of these ontologies to enhance the robustness of the extraction system.

The main measurable objective of this initiative is to improve significantly the performance of existing single-modality approaches in terms of scalability and precision. Towards that goal, our aim is to develop a new methodology for extraction and evolution, using a rich multimedia semantic model, and realize it as an open architecture. The architecture will be coupled with the appropriate set of tools, implementing the advanced methods that will be developed.

The resulting technology has a wide range of applications in commerce, tourism, e-science, etc. We are currently examining the evaluation of the technology through the development of an automatic content collection and annotation service for public events in a number of major European cities. The extracted semantic information will enrich a digital map, which will provide a friendly interface to the end user.

7. REFERENCES

[1] Cees G.M. Snoek, M. Worring Multimodal Video Indexing: A Review of the State-of-the-art Multimedia Tools and Applications, 25, 5–35, 2005

[2] C.D. Manning and H. Schutze, Foundations of Statistical Natural Language Processing. The MIT Press, Cambridge, USA, 1999.

[3] L.R. Rabiner, A tutorial on hidden markov models and selected applications in speech recognition, Proceedings of the IEEE, Vol. 77, No. 2, pp. 257–286, 1989.

[4] J. Hunter, "Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology", International Semantic Web Working Symposium (SWWS), Stanford, July 30 - August 1, 2001

[5] V.Mezaris, I.Kompatsiaris, N.V.Boulgouris and M.G.Strintzis: "Real-time compressed domain spatiotemporal segmentation and ontologies for video indexing and retrieval", IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Audio and Video Analysis for Multimedia Interactive Services, vol. 14, no. 5, pp. 606-621, May 2004.

[6] R. Troncy, “Integrating Structure and Semantics into Audio-Visual Documents”, In the second International Semantic Web Conference, ISWC 2003, LNCS 2870, pp. 566-581, 2003.

[7] D.W. Embley. Towards semantic understanding – an approach based on information extraction ontologies. In Klaus-Dieter Schewe and Hugh E. Williams, editors, ADC, volume 27 of CRPIT, page 3. Australian Computer Society, 2004.

[8] A. Valarakos, G. Paliouras, V. Karkaletsis, G. Vouros, Enhancing Ontological Knowledge through Ontology Population and Enrichment, In Proceedings of the 14th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2004), Lecture Notes in Artificial Intelligence, volume 3257, pp. 144-156, Springer Verlag, 2004.

[9] A. Maedche, and S. Staab, “Ontology Learning for the Semantic Web”, IEEE Intelligent Systems, 16(2), 2001.

[10] G. Paliouras, “On the Need for Bootstrapping Ontology Learning with Extraction Grammar Learning”, Invited paper, Proceedings of the International Conference on Conceptual Structures (ICCS), Kassel, Germany, July 2005, Springer-Verlag. (to appear)

[11] S. Castano, A. Ferrara, S. Montanelli, and G. Racca, “Matching Techniques for Resource Discovery in Distributed Systems Using Heterogeneous Ontology Descriptions”, IEEE Proc. of the International Conference on Coding and Computing (ITCC04), Las Vegas, Nevada, USA, 2004

[12] K. Kotis, G. Vouros. HCONE approach to Ontology Merging. ESWS'04. The Semantic Web: Research and Applications, LNCS, Vol. 3053, Springer-Verlag, (2004)

[13] OntoWeb. Deliverable D1.3. A survey on ontology tools, May 2002 (ed. A. Gómez Pérez)

[14] A Maedche and S. Staab. Mining ontologies from text. In R.Dieng and O.Corby, editors, EKAW, volume 1937 of Lecture Notes in Computer Science, pages 189–202. Springer, 2000.
[15] C. Brewster, F. Ciravegna, and Y. Wilks. User-centred ontology learning for knowledge management. In B. Andersson, M. Bergholtz, and P. Johannesson, editors, NLDB, volume 2553 of Lecture Notes in Computer Science, pages 203–207. Springer, 2002.
[16] L. Stojanovic. Methods and Tools for Ontology Evolution. PhD thesis, University of Karlsruhe, 2004.
[17] http://www.sts.tu-harburg.de/~r.f.moeller/racer/index.html