Introduction/summary

For what application are you using/plan to use a multimedia ontology (e.g. annotation, retrieval, media production, summarisation etc)?

* Multimedia Information Extraction and Semantic Annotation: a synergistic approach that 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.

Reference to your web site and/or publications describing the application

* C. D. Spyropoulos, G. Paliouras, V. Karkaletsis, D. Kosmopoulos, I. Pratikakis, S. Perantonis, B. Gatos, "**BOEMIE: Bootstrapping Ontology Evolution with Multimedia Information Extraction**", In *Proceedings of the 2nd European Workshop on Integration of Knowledge, Semantic and Digital Media Technologies,* London, UK, November 2005.

* V. Karkaletsis, G. Paliouras, C. D. Spyropoulos, "**A Bootstrapping Approach to Knowledge Acquisition from Multimedia Content with Ontology Evolution**," In *Proceedings of the International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning (AKRR),* pp. 98-105, Helsinki University of Technology, Finland, June 2005.

* D. Kosmopoulos, S. Petridis, I. Pratikakis, B. Gatos, S. Perantonis, V. Karkaletsis, G. Paliouras Knowledge Acquisition from Multimedia Content using an Evolution Framework, submitted 3rd. IFIP Conference on Artificial Intelligence Applications & Innovations, Athens 2006

Requirements of the “minimum set” involving the analysis of:

Scope and Usage:

* Knowledge representation for multimedia content, combining low-level with high-level semantic features, as well as providing reference to geographic concepts and objects.

* Fusion techniques for data from heterogeneous sources (e.g., 2D geometric objects, phonemes, etc.) using an ontology-based approach.

* **Multimedia content ontology:** It represents the structure of the content of the multimedia documents.

*** Multimedia descriptor ontology:** This ontology will model concepts and properties that describe visual and non-visual characteristics of objects.

*** Domain-specific ontologies:** These ontologies will contain concepts and properties related to the knowledge of the domain of interest (geographical information) and more specific application domains that will be selected according to user requirements analysis.

*** Geographic information ontology:** It will contain the domain specific geographic concepts and relations.

* Representation and management of uncertainty, imprecision and vague knowledge that exists in real life applications.

* Automated ontology evolution with heterogeneous multimedia information: ontology population, enrichment and coordination, ontology management and versioning as well as the preservation of semantic consistency under different requirements and constraints.

* Novelty detection that is based on information from all different types of media being processed.

Resources/Concepts including information on what concepts you need to describe :

* **Multimedia content ontology:** The top level hierarchy of a multimedia document is classified into five types: Image, Video, 3D graphics, Audio, Audiovisual and Multimedia. Each of these types has its own segment subclasses. MPEG-7 provides a number of tools for describing the structure of multimedia content in time and space. A number of specialized subclasses are derived from the generic Segment Description Scheme. These subclasses describe the specific types of multimedia segments, such as video segments, moving regions, still regions and mosaics, which result from spatial, temporal and spatiotemporal segmentation of the different multimedia content types.
* **Multimedia descriptor ontology:** This ontology may model concepts and properties that describe audiovisual characteristics of objects, especially audio segments and visualizations in still images and videos in terms of low-level features and media structure descriptions. Structure and semantics should be largely consistent with the existing MPEG-7 multimedia description standard and several modifications should be investigated and carried out in order to adapt the XML Schema provided by MPEG-7 to an ontology. Subconcepts should 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 specific application domains of varying granularity. These ontologies also contain detailed descriptions of objects using spatiotemporal and partonomic relations defined in the multimedia semantic model.
*** Geographic information ontology:** It should contain the domain specific geographic concepts and relations modelling in an adequate and smart way the position dependent information needed to geo-reference multimedia content to the map. The construction of this ontology requires investigation in the following areas: foundational ontology of spatial properties and relations, vague and ambiguous spatial and geographic concepts, the relationship between geographic and physical concepts, modelling spatial granularity, events and spatio-temporal processes, identity of spatial/geographic objects through time, aggregate objects.

Requirements for cross-domain linking, if applicable

* Ontology coordination 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 can be used for coordinating homogeneous ontologies, e.g. two ontologies for the same domain, while alignment/mapping techniques can be used for coordination of heterogeneous ontologies, e.g. link a domain-specific ontology with the multimedia content ontology. Machine learning techniques can be used in order to detect similarities between concepts exploiting features in syntactic, structural and semantic level.

Tools (e.g. annotation, reasoning)

* Tools to handle uncertainty and support the probabilistic evidence-based interpretation framework (e.g. inference using a Bayesian Network).

* Ontology-based fusion tool : The results of the extraction process are used to build hypotheses. These should be then matched/linked with information included in the multimedia ontologies 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 should be implemented as a *closed-loop extraction process*.

* Tool to support the initial extraction of characteristics of multimedia objects as instances of the Multimedia Descriptor Ontology and linkage of these descriptors with actual concepts of the domain ontology in a user friendly-way.

* Ontology population and enrichment tools

* Ontology coordination tools

* Description logic reasoning systems, to support “incremental changes”.

* Despite many advances in description logic inference technology, for representation languages with high expressivity (e.g. OWL-DL), practical application scenarios with large ontologies require the development of new optimization techniques for reasoning at the concept level.

* Ontology management and versioning tool: support the designer in maintaining an ontology during the entire bootstrapping process in the context of an interactive environment. In addition, the ontology management tool will also control ontology versioning, by maintaining the different versions of the ontology throughout the evolution process.

Languages – standards

* Explicit fuzzy extensions of OWL-DL can be derived, while this approach should be extended for SWRL and OWL-QL to support uncertain rules and queries as well.

* Modifications are needed in order to adapt the XML Schema provided by MPEG-7 to an ontology and the data type representations available in OWL.

Identification of the characteristics to be represented:

* *Support for Multiple Audiovisual Descriptions*: Audiovisual characteristics of domain concepts cannot be described using one single instance of the multimedia descriptors in question. The approach followed will provide means for multiple prototypical descriptions of a domain concept using the MPEG-7 standard.

* *Spatiotemporal Relation Representation*: Modelling of spatiotemporal and partonomic relations is required apart from simple audiovisual properties. Based on concepts taken from the 'region connecting calculus', Allen's interval calculus and directional models, directional and topological relationships in the spatial and temporal domains can be represented in the knowledge structure. Existing spatio-temporal relation of ontologies (e.g. as in the DOLCE ontology) can be extended to accommodate topological and directional relations between regions of different types.

* *Multimedia Structure Representation*: The result of the annotation or content analysis should be able to express the structure of a multimedia document itself, depending on the type of document, e.g. image, video, audio, or multimedia presentation. A hierarchical structure of multimedia segments in the Multimedia Content Ontology should thus be developed in order to capture all possible types of spatiotemporal or media decompositions and relations aligned with the MPEG-7 Standard.

Harmonisation approaches

* The role of a unifying multimedia semantic model is to serve as an integrated model for a variety of ontologies. This may include relations, constraints and rules representing spatiotemporal qualities of the multimedia content which further assist the analysis of the multimedia content.