

# Semantic, Constraint & Preference Based Multimedia Presentation Authoring

Chrisa Tsinaraki<sup>1</sup>, Andrea Perego<sup>2</sup>, Panagiotis Polydoros<sup>1</sup>, Athina Syntzanaki<sup>1</sup>,  
Alessandro Martin<sup>3</sup>, Stavros Christodoulakis<sup>1</sup>

<sup>1</sup> Technical University of Crete,

Laboratory of Distributed Multimedia Information Systems and Applications (TUC/MUSIC)

University Campus, 73100 Kounoupidiana, Crete. Greece.

{chrisa, panpolyd, athina, stavros}@ced.tuc.gr

<sup>2</sup> University of Insubria, Department of Computer Science and Communication

(UNINSUBRIA/DICOM), Via Manzoni 5, 21100 Varese. Italy.

andrea.perego@uninsubria.it

<sup>3</sup> University of Milan

Department of Computer Science and Communication (UNIMI/DICo)

Via Comelico 39/41, 20135 Milano. Italy.

martin@dico.unimi.it



Journal of Digital  
Information Management

Uncorrected Proof

**ABSTRACT:** We present in this paper an integrated system that allows the management and annotation of multimedia objects stored in MPEG-7/21 repositories, and the specification and semi-automatic generation of multimedia presentations based on the content relationships that exist between the multimedia objects. This system is the outcome of the collaboration between the Technical University of Crete (TUC-MUSIC) and the University of Milan (UNIMI) in the CoCoMA (Content and Context Aware Multimedia Content Retrieval, Delivery and Presentation) project of the DELOS II European Network of Excellence on Digital Libraries. The resulting system is one of the main components of the CoCoMA infrastructure that aims to provide content- and context-aware rich interactive multimedia presentations by controlling data fusion and metadata reuse. The integrated system utilizes the SyMPA management and presentation authoring system developed by UNIMI and the DS-MIRF framework developed by TUC-MUSIC.

## Categories and Subject Descriptors

H.5.1 [Multimedia Information Systems] Wireless Communication;

H.3.7 [Digital Libraries]: Standards

## General Terms

Multimedia Content Retrieval, Metadata, MPEG 7, MPEG 21

**Keywords:** Digital libraries, Multimedia standards, Multimedia presentations, Authoring tools

Received 8 April 2006; Reviewed and accepted 24 April 2006

## 1. Introduction

The high penetration of the digital multimedia in several domains of the everyday life (e.g. education, entertainment, etc.) has resulted in the increasing production of multimedia objects and their synthesis in multimedia presentations. A **multimedia presentation** essentially is a graph, where each node corresponds to a set of heterogeneous multimedia objects (e.g., text, images, audio, and video files) that are grouped according to their content relationships and organized according to a given spatial and temporal disposition. The edges connecting the graph nodes denote the execution flow of the presentation — i.e., the sequence according to which the objects in each node are displayed to the user.

Multimedia presentation authoring is quite a complex task and involves several different issues. The authors must collect the objects to be used, specify their spatial and temporal disposition inside each presentation node, and, finally, build the presentation graph (the 'structure' of the presentation). The existing authoring models and systems provide different strategies for carrying out such tasks, but they require authors to specify explicitly the objects to be used, their disposition, and the presentation structure. This may be acceptable when we have a limited number of objects, but this approach is not applicable when multimedia presentations are a means to access the contents of multimedia repositories such as Digital Libraries (DLs). In this case, mechanisms should be provided in order to automate as much as possible the composition of a multimedia presentation. In order to achieve this, the authoring procedure should focus on the 'content' of the presentation itself. In fact, a multimedia presentation can be seen as a discourse on a

given topic, where each node corresponds to a sub-topic, and the presentation structure represents the order according to which the sub-topics are addressed. Thus, the content relationships that exist between the objects, which may be either extracted from the objects themselves or derived from the metadata associated with them, can be used in order to generate both the nodes and the structure of the multimedia presentations. Author intervention cannot be avoided, but it mainly concerns the revision of the presentations automatically generated by the system.

In this paper, we present an integrated system that allows the management and annotation of multimedia objects stored in MPEG-7/21 repositories, and the specification and semi-automatic generation of multimedia presentations based on the content relationships which exist among multimedia objects. This system is the outcome of the collaboration between the Technical University of Crete (TUC-MUSIC) and the University of Milan (UNIMI) in the context of the CoCoMA (Content and Context Aware Multimedia Content Retrieval, Delivery and Presentation) project of the DELOS II European Network of Excellence on Digital Libraries. The resulting system is one of the main components of the CoCoMA infrastructure (Christodoulakis & al., 2005; Christodoulakis & al., 2006), which aims to provide content- and context-aware rich interactive multimedia presentations, by controlling data fusion and metadata reuse. The integrated system utilizes the **SyMPA (System for Multimedia Presentation Authoring)** management and presentation authoring system developed by UNIMI and the **DS-MIRF (Domain-Specific Multimedia Indexing, Retrieval and Filtering) framework** developed by TUC-MUSIC.

The DS-MIRF framework (Tsinaraki & al., 2007) utilizes and extends the MPEG-7 (Chang & al., 2001) and MPEG-21 (Pereira, 2001) standards in order to facilitate the development of knowledge-based multimedia applications. In the DS-MIRF framework, ontology-based semantic annotation of the multimedia objects is supported using annotation interfaces integrated with the GraphOnto semantic multimedia annotation component (Polydoros & al., 2006), that utilizes the ontological infrastructure of the DS-MIRF framework which is expressed in OWL. The OWL (McGuinness & van Harmelen, 2004) annotations are then transformed, using the **DS-MIRF transformation rules**, to **MPEG-7/21 metadata descriptions**. The MPEG-7/21 metadata descriptions are stored in the **DS-MIRF Metadata Repository**, which is accessed by the end-users through appropriate **application interfaces** built on top of the MP7QL query language that is being developed in the DS-MIRF framework for querying MPEG-7 descriptions (Tsinaraki & Christodoulakis, 2006). Semantic user preference descriptions, based on the MP7QL syntax, are also stored in the DS-MIRF Metadata Repository.

**SyMPA** is a Web-based management and presentation authoring system, which consists of three main components: a database, a set of modules for object management, annotation, and presentation specification, and two modules in charge, respectively, of presentation generation and of the retrieval of objects or presentations. SyMPA allows users to acquire and annotate objects, using multiple metadata vocabularies (which may be plain sets of descriptors, conceptual hierarchies, and ontologies), concerning both high- and low-level features. These annotations are then used to assist authors in building multimedia presentations. SyMPA allows the specification of multimedia presentations that have a dynamic

structure, unlike the existing approaches, which are designed for building fixed-structure presentations comprised of a fixed set of objects. Because of the dynamic structure of the presentations, when alternative versions of the same presentation are required, for example varying in duration or using different sets of multimedia objects, the author does not have to specify them explicitly. This not only reduces the complexity of the presentation specification task, but it also allows personalizing a presentation based on the interests and the skill levels of the end users. In order to address this issue, SyMPA utilizes the multimedia presentation authoring model described in (Bertino & al., 2005), where content relationships among objects are used to identify the objects associated with each node of the presentation, and to build automatically different execution flows of the same presentation. Content-based and semantic metadata that are associated with multimedia objects can be used to automatically carry out the presentation specification task in SyMPA. The metadata may be used from the authoring model to infer the content relationships that exist between the objects, which will then determine the presentation structure and the multimedia objects that should be associated with each node of the presentation. Thus, although the utilization of metadata cannot make completely automatic the presentation specification procedure, it may improve its efficiency, especially when dealing with large collections of multimedia objects, where locating objects may be a difficult and time-consuming task.

As the metadata stored in the DS-MIRF MPEG-7/21 Metadata Repository can be utilized by the authoring model of SyMPA, we are working on the integration of the DS-MIRF Metadata Repository with SyMPA, in order to store in the repository the multimedia presentations and the multimedia object annotations defined using the SyMPA, and for locating multimedia objects that will be utilized in the presentations. In addition, the ontological infrastructure of the DS-MIRF framework, which already includes domain ontologies for soccer and formula 1 and is currently being extended with an art ontology, is utilized in SyMPA as a set of metadata vocabularies for multimedia object annotation. Finally, the semantic user preference descriptions stored in the DS-MIRF MPEG-7/21 Metadata Repository will be systematically utilized in order to allow for semantic presentation personalization.

The remainder of this paper is organized as follows: Section 2 discusses related work issues; Section 3 provides an overview of the integrated architecture of the proposed system, whereas section 4 illustrates the proposed semantic, constraint and preference based authoring approach; Finally, section 5 concludes the paper and outlines future research directions.

## 2. Related Work

We present in this section research efforts relevant to our integrated multimedia authoring system. First, we present in subsection 2.1 research efforts in multimedia presentation authoring, and then we describe in subsection 2.2 research efforts in knowledge-based multimedia application support.

### 2.1. Multimedia Presentation Authoring

The existing systems for multimedia presentation authoring and generation can be grouped into two main approaches: the **operational approach** and the **constraint-based approach**. The former asks authors to specify explicitly the **absolute** spatial and temporal disposition of the objects in each node of the presentation by using, for example, (x, y) coordinates and/or timelines, whereas the latter provides a set of spatio-temporal constraints expressing the **relative** position of each object with respect to another one (for instance, the temporal T\_Before(a, b) constraint states that object a must be played before object b). The operational approach, due to the fact that it is easy to implement and that it allows authors to completely control the final presentation, is the most often adopted by the existing systems, such as Macromedia Director, ZyX (Boll & Klas, 2001), MET++ (Ackermann, 1994), and CMIF (Hardman & al., 1993). On the other hand, the constraint-based approach has the advantage of requiring just a high level specification of the spatio-temporal disposition of the objects, which is then used by the system to generate the final presentation. Among the systems that adopt this approach are CHIMP (Candan & al., 1996), CUYERS (Geurts & al., 2001; van Ossensbruggen & al., 2003), and MADEUS (Jourdan & al., 1998; Tardif & al., 2000). Despite their differences, both the operational and the constraint-based approaches force authors to

focus on how objects are presented to the end users, and not on the content of the presentation. For this purpose, in Bertino & al. (2005) a multimedia presentation authoring model has been proposed, which makes use of content-based constraints for the authoring and the semi-automatic generation of a multimedia presentation. In particular, three content-based constraints have been defined, which are binary relations expressing, respectively, that two objects belong to the same topic (C\_Same), that two objects belong to different topics (C\_Different), and that one object belongs to a topic which 'conceptually' follows the topic of the other one (C\_Link). Such constraints are then used in order to generate the final presentation, where the C\_Same and the C\_Different relations identify the objects which belong to each presentation node, whereas the C\_Link relations determine the structure of the presentation graph. Moreover, thanks to the C\_Link constraint, it is possible to automatically determine different execution flows of the same presentation, which can be chosen by end-users depending on their preferences and/or skill levels, whereas in the 'traditional' approaches each possible execution flow of the same presentation must be specified explicitly.

### 2.2 Knowledge-based Multimedia Application Support

As MPEG-7 and MPEG-21 are the dominant standards for multimedia application support, the research in this area is mainly based on them. Although the well-accepted MPEG-7 standard allows, in the MPEG-7 MDS (ISO/IEC; 2003a), the semantic description of the audiovisual content using both keywords and structured semantic metadata, several systems follow the keyword-based approach (Rogers & al., 2003; Tseng & al., 2004; Wang & al., 2004; Graves & Lalmas, 2002). The keyword-based approach is limiting, as it results in reduced precision of the audiovisual content retrieval. As an example, consider a fan of the Formula-1 driver Alonso, who wishes to retrieve the audiovisual segments containing the overtakes that Alonso has performed against Michael Schumacher. If the user relies on the keyword "overtake" and the names "Alonso" and "Schumacher", (s)he will retrieve, in addition to the segments containing the overtakes that Alonso has performed against Schumacher, also the segments containing the overtakes that Schumacher has performed against Alonso.

The above problem may be solved (at least at some extent) using the structured semantic description capabilities provided by MPEG-7. The major shortcoming of most of the systems adopting this approach is that the general-purpose constructs provided by MPEG-7 are used without a systematic effort for domain knowledge integration in MPEG-7 (Agius & Angelides, 2004; Hammiche & al., 2004; Lux & Granitzer, 2005), so that standard MPEG-7 software may utilize it. An approach that allows integrating, in semantic MPEG-7 descriptions, domain knowledge expressed in domain ontologies formed using MPEG-7 constructs, is discussed in (Tsinaraki & al., 2003; Tsinaraki & al., 2005a). As the utilization of existing OWL domain ontologies makes interoperability support within user communities easier, a methodology for the integration of OWL domain ontologies in MPEG-7 has also been developed (Tsinaraki & al., 2007).

A structured semantic content description model cannot be fully exploited by keyword-based user preferences; As this is the case in MPEG-7/21 (the user preferences allow only keyword-based descriptions of the desired content), the MPEG-7/21 based systems either utilize keyword-only metadata, thus ignoring the structured MPEG-7 semantic metadata (Rogers & al., 2003; Tseng & al., 2004; Wang & al., 2004), or they ignore the MPEG-7/21 user preference model and follow proprietary filtering approaches on top of the structured MPEG-7 semantic metadata (Agius & Angelides, 2004). In order to allow the full exploitation of structured semantic audiovisual content descriptions, a semantic user preference model for MPEG-7/21 has been proposed in (Tsinaraki & Christodoulakis, 2006).

## 3. System Architecture

We provide in this section an overview of the major components of our integrated system, namely the DS-MIRF framework (presented in subsection 3.1) and the SyMPA multimedia presentation authoring system (presented in subsection 3.2), as well as an overview of the architecture of our integrated system (presented in subsection 3.3).

### 3.1 The DS-MIRF Framework

We present in this subsection the DS-MIRF framework (Tsinarakis & al., 2003; Tsinarakis & al., 2005; Tsinarakis & al., 2007), a software engineering framework that aims to facilitate the development of knowledge-based multimedia applications utilizing and extending the MPEG-7/21 standards. The **multimedia content annotator** is a special type of user in DS-MIRF, who performs the semantic annotation of multimedia documents using an **annotation interface** integrated with the **GraphOnto** semantic multimedia annotation component (Polydoros & al., 2006). GraphOnto is a Java application that allows OWL ontology management, ontology-based semantic annotation, and utilizes the ontological infrastructure of the DS-MIRF framework. The DS-MIRF ontological infrastructure includes:

- (a) An OWL **Upper Ontology** that fully captures the MPEG-7 MDS (Multimedia Description Schemes) and the MPEG-21 DIA (Digital Item Adaptation) Architecture (ISO/IEC 2003b);
- (b) A set of OWL **Application Ontologies** that provide additional to the upper ontology functionality in OWL. The functionality provided by the application ontologies either makes easier for the user the use of the MPEG-7/21 (like, for example, a typed relationship ontology that captures the typed relationship semantics implied in the MPEG-7 MDS text) or supports advanced multimedia content services (like, for example, a semantic user preference ontology); and
- (c) OWL **Domain Ontologies**, which extend the Upper Ontology and the Application Ontologies with domain knowledge (like, for example, sports ontologies for soccer, formula 1 etc.).

Since all the ontologies in the DS-MIRF framework are expressed in OWL, the result of the annotation is an OWL description of the multimedia content. The OWL annotations are then transformed, using the **DS-MIRF transformation rules** that are implemented in the GraphOnto component, to **MPEG-7/21 metadata descriptions**. The MPEG-7/21 metadata are stored in the **DS-MIRF Metadata Repository**.

The DS-MIRF Metadata Repository has been developed on top of the Berkley DB XML, contains MPEG-7/21 metadata descriptions associated with multimedia objects and provides semantic retrieval capabilities based on the MP7QL language for querying MPEG-7 descriptions. In addition to the multimedia object annotations, MPEG-7/21 user preference descriptions and semantic user preference descriptions, structured according to the model specified in (Tsinarakis & Christodoulakis, 2006), which is based on the MP7QL query model, are stored in the DS-MIRF Metadata Repository, in order to allow for

the personalization of the services offered to the users.

The MP7QL query language has been expressed using both XML Schema and OWL syntax and the XML Schema syntax is being supported on top of the DS-MIRF Metadata Repository. The MP7QL allows one to query every aspect of an MPEG-7 multimedia object description and also allows to utilize the user preferences as context, in order to support personalized multimedia content retrieval. The MP7QL allows queries about: (a) multimedia content that satisfies specific criteria (for example, "give me the multimedia objects where a goal is scored"); (b) semantic entities that satisfy specific criteria and may be used for the semantic descriptions of multimedia content (for example, "give me the players affiliated to the soccer team Barcelona"); and (c) domain ontology constructs expressed using MPEG-7 syntax (for example, "give me the subclasses of the Player class").

### 3.2 The SyMPA Authoring System

We present in this subsection SyMPA, a Web-based management and presentation authoring system for multimedia objects stored in distributed repositories. SyMPA consists of three main components: a database, a set of modules for multimedia object management, annotation, and presentation specification, and two modules in charge, respectively, of presentation generation and objects' / presentations' retrieval. In the SyMPA architecture, depicted in Figure 1, multimedia objects are stored in distributed repositories, whereas their metadata are stored and managed by a centralized database. Multimedia objects and the associated metadata are managed through a Web-based administration interface, which is supported by a set of software modules that collaborate in order to perform the supported tasks.

The SyMPA administration interface allows the users to acquire and annotate objects using multiple metadata vocabularies (which may be plain sets of descriptors, conceptual hierarchies, and ontologies) which describe both high- and low-level features. The annotations are then used to assist authors in building multimedia presentations, according to a content-based approach, formally defined in (Bertino & al., 2005), where content relationships among objects are used to identify the objects associated with each node of the presentation and to build automatically different execution flows of the same presentation. This way, presentation specification becomes a task similar to object annotation, and the approach also allows for specifying presentations using the contents of large multimedia object repositories, such as DLs. The authors first select the topic to be addressed by the presentation from the available metadata vocabularies. The existing multimedia object annotations

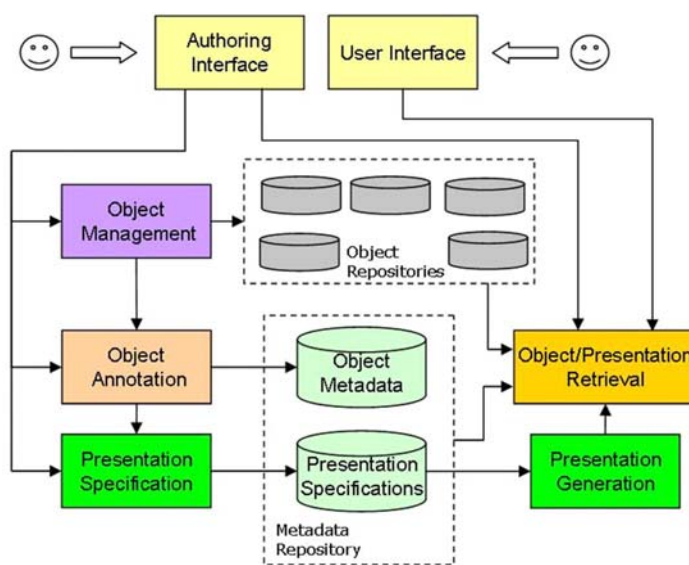


Figure 1. The SyMPA architecture

are then evaluated by the system and the multimedia objects relevant to the topic as well as a proposed structure of the presentation are returned. The presentation structure is obtained by grouping objects sharing similar characteristics (i.e., metadata) into subsets, whose nested structure is used to build the proposed presentation graph, where the nodes correspond to possible 'sub-topics' of the presentation.

The example depicted in Figure 2a describes the generation of the presentation graph starting from a user query. The user specifies the topic of the presentation (e.g., a presentation about German gothic oil paintings) by selecting a set of descriptors (e.g.,  $d_1$ =artist nationality,  $d_2$ =technique,  $d_3$ =movement). The system then retrieves the objects associated with the specified descriptors and groups them in a set of clusters. The result is a set of nested clusters, corresponding to a subset of the powerset of the specified descriptors. Given the 3 descriptors  $d_1$ ,  $d_2$ , and  $d_3$ , we have the following  $2^3$  combinations:  $D_1 = \{d_1, d_2, d_3\}$ ,  $D_2 = \{d_1, d_2\}$ ,  $D_3 = \{d_1, d_3\}$ ,  $D_4 = \{d_2, d_3\}$ ,  $D_5 = \{d_1\}$ ,  $D_6 = \{d_2\}$ ,  $D_7 = \{d_3\}$ ,  $D_8 = \{\}$ . Clusters correspond to the subsets of objects associated with one or more elements of the powerset, excluded the empty set  $\{\}$ . Thus, given  $n$  descriptors, we will have  $k$  clusters of objects, with  $0 \leq k \leq 2^n - 1$ . The structure of the presentation is determined by building an oriented edge from the including cluster to the included one. For example, the cluster corresponding to  $D_1$  (i.e.,  $C_1$ ) will be the presentation root, since it includes all the other ones. Then, the second level of the presentation structure will consist of the clusters corresponding to  $D_2$ ,  $D_3$ , and  $D_4$  (i.e.,  $C_2$ ,  $C_3$ ,  $C_4$ ), whereas the third level will consist of the clusters corresponding to  $D_5$ ,  $D_6$ , and  $D_7$  (i.e.,  $C_5$ ,  $C_6$ ,  $C_7$ ). The resulting presentation structure is depicted in Figure 2 b.

Besides inclusions, which are used to build the main structure of the presentation, the intersections between clusters are also considered in order to build alternative paths between nodes of the presentation. In particular, a not-oriented edge connects intersected, but not included, clusters. In our example, the intersected, but not included, clusters are: (a)  $C_2$  and  $C_3$ ; (b)  $C_2$  and  $C_4$ ; and (c)  $C_3$  and  $C_4$ . Thanks to these edges, the end users can directly jump between content-related nodes at the same or different level of a presentation, without the need of walking through the upper level(s) of the presentation.

It is important to note that the resulting presentation graph consists of all the possible execution flows of the presentation, which can be used to provide different versions of the same presentation, addressing the different preferences (in terms of content, duration, etc.) of the end-users. Finally, the author is asked to revise the proposed presentation by possibly modifying its structure and/or adding/removing objects in the presentation nodes.

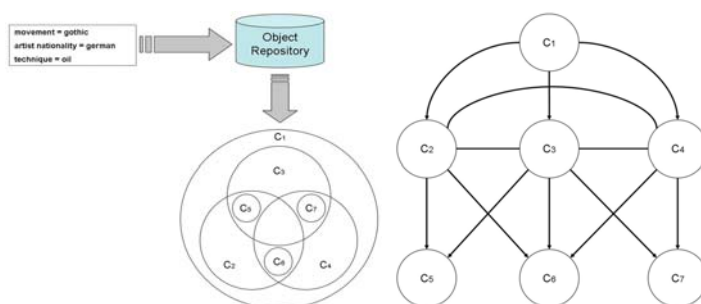


Figure 2. (a) Grouping the retrieved objects; (b) Presentation graph

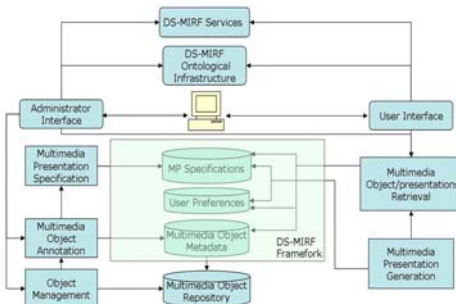


Figure 3. Overview of the System Architecture

This approach has the advantage of supporting a **declarative** specification of a multimedia presentation, based on its 'content', whereas the existing systems enforce **procedural** strategies, where a multimedia presentation is specified in terms of the spatial-temporal disposition of the objects in each node, and on how presentation nodes are organized. Moreover, content relationships can be used to automatically obtain multiple execution flows of the same presentation, which can be chosen by the end-users depending on their interests or skill levels. By contrast, in the existing systems the presentation graph is fixed, and possible alternative paths must be specified explicitly by building different presentations.

### 3.3 Architectural Overview

The integrated architecture of our system is depicted in Figure 3. Users and administrators interact with the system using the SyMPA Graphical User Interface (GUI), which utilizes the DS-MIRF ontological infrastructure and interacts, through the services offered, with the DS-MIRF metadata repository. The services offered on top of the DS-MIRF metadata repository include: (a) The insertion of ontologies, semantic entities for multimedia object annotations, multimedia object annotations and multimedia presentations in the DS-MIRF metadata repository; (b) MP7QL-based retrieval services, which allow locating multimedia objects based on every available aspect of the MPEG-7 multimedia object descriptions; and (c) Ontology and semantic entity access services for the ontologies and the semantic entities (to be used in the multimedia object annotations available in the DS-MIRF metadata repository). These services are offered in the same way with the retrieval services, as the ontology and semantic entity access are supported by MP7QL as special query types.

This way, the users may perform multimedia object management operations on the multimedia object repository, annotate multimedia objects — thus producing multimedia object metadata — and specify multimedia presentations, which are stored in the DS-MIRF metadata repository. During multimedia presentation specification, multimedia objects and presentations may be retrieved in order to be used in the newly-created presentations.

The integrated architecture of Figure 3 shows the interplay of DS-MIRF with SyMPA, in order to support semantic, constraint and preference based multimedia authoring. In order to achieve this, the DS-MIRF framework and the SyMPA multimedia presentation authoring system have been integrated in the following points:

1. The DS-MIRF ontological infrastructure is utilized by SyMPA as a set of metadata vocabularies and is being extended with an art ontology. The users utilize the metadata vocabularies during the selection of knowledge domain and topic of interest.

2. The DS-MIRF MPEG-7/21 Metadata Repository is used: (a) For storing multimedia presentations and multimedia object annotations defined using SyMPA; (b) For locating, using the MP7QL-based retrieval capabilities of the DS-MIRF framework, multimedia objects that will be utilized in the presentations; and (c) For retrieving the metadata associated with the multimedia objects, which are used by the authoring model to infer the content relationships that exist among objects. These relationships determine which multimedia objects should be associated with each node of the presentation as well as the presentation structure.
3. The semantic user preference descriptions stored in the DS-MIRF MPEG-7/21 Metadata Repository are system-atically utilized by the authoring model in order to allow presentation personalization. In particular, the content-related likes and dislikes of the users, as they are expressed in their preference descriptions, allow tailoring the presentation to them, especially if some of the multimedia objects should be discarded for one of the following reasons: (a) Too many multimedia objects related to the topic(s) the user selected exist in the repository; and (b) Duration and/or space (regarding the maximum allowed number of multimedia objects to be used in the presentation) constraints exist.

The multimedia presentation definition scenario of the integrated system has as primary agent the end user, who interacts with the SyMPA GUI. When a user chooses the knowledge domain and the topic(s) of interest from the metadata vocabularies, SyMPA queries the DS-MIRF metadata repository on the topic(s) selected and their sub-topics and builds, according to the query results, the graphical user interface. The queries are expressed in MP7QL, are automatically generated by the SyMPA authoring system and are then posed, using the semantic query services offered by DS-MIRF, on the DS-MIRF repository in order to locate multimedia objects as well as semantic entities and ontology terms related to the topic of interest. Then, the authoring model of (Bertino & al., 2005) is used and a proposed presentation specification structure is displayed to the end user, who may then modify it.

In order to improve the performance of the interface generation process, the SyMPA user interface generation has been redesigned as shown in Figure 4 in order to use domain-specific user interface templates. The SyMPA GUI builds on the following components: (a) The **Interface Selector**, which allows the user to select a specific query interface through the selection, first, of a knowledge domain of interest, and the selection of the topic(s) of interest afterwards; (b) The **Interface Builder**, which is in charge of querying the DS-MIRF metadata repository in order to get all the needed information regarding ontologies and semantic entities used in the multimedia object annotations, and of deciding which is the appropriate interface template that will be used; and (c) The **Template Repository**, which contains a set of interface templates in order to speed up the query interface building process. Thanks to the templates, the interface builder does not need to rebuild the whole interface every time a user selects a different knowledge domain or different topics.

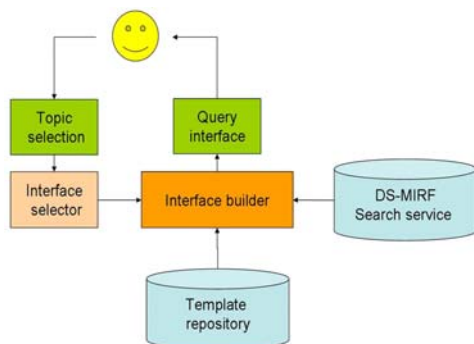


Figure 4. The SyMPA Graphical User Interface (GUI)

#### 4. Semantic, Constraint and Preference based Authoring Support

We present in this section the multimedia presentation authoring approach supported by our integrated system. As already mentioned, the MPEG-7/21 content-based and semantic metadata stored in the DS-MIRF Metadata Repository, which are associated with multimedia objects, are used to automatically carry out the presentation specification task in SyMPA. Although the utilization of the metadata stored in the DS-MIRF metadata repository cannot make completely automatic the presentation specification procedure, they can be used for improving its efficiency, especially when dealing with large collections of multimedia objects, where locating objects may be a difficult and time-consuming task. The presentation authors may specify a presentation utilizing the semantic retrieval capabilities of the DS-MIRF framework by defining a set of topics that should relate with the semantic metadata that describe the multimedia objects to be used in the presentation. The metadata are used in the authoring model to infer the content relationships that exist among objects, which will in turn determine the multimedia objects that should be associated with each node of the presentation and the presentation structure itself. This way, the multimedia presentation authoring approach of (Bertino & al., 2005) is enhanced by inferring content constraints from the metadata associated with multimedia objects. Based on this, the integrated system returns the objects belonging to each presentation node; then the author decides which objects should be used and their spatial and temporal disposition. Finally, the possible execution flows of the presentation are obtained by evaluating the semantic relationships existing among the selected objects.

Although the metadata utilization assists the presentation generation, it does not allow avoiding author intervention for two main reasons:

1. Content constraints can be used only for grouping objects and building the presentation structure, but the spatial and temporal disposition of the objects in each node cannot be determined automatically.
2. There is no control over the number of objects which will be automatically assigned to a presentation node. For instance, assume that in the DS-MIRF repository 20 objects exist, associated with metadata describing them as images reproducing Impressionist paintings: in this case, the node of the presentation corresponding to the Impressionism should contain 20 images, independently from the size of the display area. This issue may be addressed by associating with each object a relevance level, which can be used to discard the less relevant objects, thus reducing the number of objects in each node of the presentation. Nonetheless, in most cases, the relevance of an object cannot be determined a priori, but it depends on the context—i.e., the topic and/or the presentation. A possible solution may be to decide the relevance level of an object by taking into account the existing presentations, according to the principle that the more the object is used in a given context, the more it is relevant to it. For instance, if we have several presentations concerning Impressionism where a given object is always used, this object may be considered relevant for this topic. Note, however, that, although this strategy allows us to possibly reduce the number of objects, we still may have a too large number of equally relevant objects. For instance, if among the set of objects concerning Impressionism, 10 of them are equally relevant, they may be still too many for a single node of a presentation. Moreover, this procedure can be applied only when we already have a sufficiently large and heterogeneous set of presentations in the system, so that it will be possible to evaluate statistically the relevance of the objects for any available topic.

In order to overcome the second limitation, the user preference descriptions are utilized in the integrated system, so as to discard objects that, in addition to the topics of interest selected by the user, they are also related with one or more topics for which the user has expressed dislike in his/her user preference description. In addition, if the objects related with the topic(s) of interest are too many and there exist (implicit or explicit) duration and/or space constraints, only the most relevant to the user preferences multimedia objects will be used in the presentation. This way, the author intervention is further limited.

## 5. Conclusions – Future Work

We have presented in this paper an integrated system that allows the management and annotation of multimedia objects stored in MPEG-7/21 repositories, and the specification and semi-automatic generation of multimedia presentations based on the content relationships which exist among multimedia objects. This system is the outcome of the collaboration between the TUC-MUSIC and the UNIMI in the CoCoMA project of the DELOS II European Network of Excellence on Digital Libraries. The resulting system is one of the main components of the CoCoMA infrastructure, which aims to provide content- and context-aware rich interactive multimedia presentations by controlling data fusion and metadata reuse. The integrated system utilizes the SyMPA management and presentation authoring system developed by UNIMI and the DS-MIRF framework developed by TUC-MUSIC. In particular, the DS-MIRF framework and the SyMPA multimedia presentation authoring system have been integrated in the following points:

- (a) The DS-MIRF ontological infrastructure is utilized by SyMPA as a set of metadata vocabularies that are used during the selection of knowledge domain and topic of interest. In addition, the SyMPA GUI utilizes the ontology and semantic entity access capabilities provided by the MP7QL on top of the DS-MIRF MPEG-7/21 Metadata Repository.
- (b) The DS-MIRF MPEG-7/21 Metadata Repository is used: (a) For storing multimedia presentations and multimedia object annotations defined using SyMPA; (b) For locating, using the MP7QL-based retrieval capabilities of the DS-MIRF framework, multimedia objects that will be utilized in the presentations; and (c) For retrieving the metadata associated with the multimedia objects, which are used by the authoring model to infer the content relationships that exist among objects. These relationships determine which multimedia objects should be associated with each node of the presentation as well as the presentation structure.
- (c) The semantic user preference descriptions stored in the DS-MIRF MPEG-7/21 Metadata Repository are systematically utilized in order to allow presentation personalization so as to take into account the user likes and dislikes and to meet duration and/or space (implicit or explicit) constraints.

A first demonstrator of the integrated system is available, and the finer integration of the components, including the integration of the GraphOnto component with the SyMPA user interface for the management of the OWL ontologies, is being finalized. Our future plans include the integration of our multimedia authoring system with the other components of the CoCoMA architecture (Christodoulakis & al., 2005; Christodoulakis & al., 2006).

## Acknowledgements

The work presented in this paper was partially funded in the scope of the DELOS II Network of Excellence in Digital Libraries (IST – Project Record Number 507618).

## References

1. Ackermann, P (1994). Direct manipulation of temporal structures in a multimedia application framework. *In: Proc. of the 2nd ACM Int. Conf. on Multimedia*, p. 15–58.
2. Agius, H., Angelides, M (2004). Modelling and Filtering of MPEG-7-Compliant Meta-Data for Digital Video. *ACM SAC*, 2004, p. 1248-1252.
3. Bertino, E., Ferrari, E., Perego, A., Santi, D (2005). A constraint-based approach for the authoring of multi-topic multimedia presentations. *In: Proc. of the IEEE International Conference on Multimedia & Expo (ICME 2005)*, Amsterdam (The Netherlands), July 6-8, 2005, IEEE CS Press.
4. Boll, S., Klas, W (2001). ZyX – A multimedia document model for reuse and adaptation. *IEEE Transactions on Knowledge and Data Engineering* 13 (3) 361–382.
5. Candan, K.S., Prabhakaran, B., Subrahmanian, V.S (1996). CHIMP: A framework for supporting distributed multimedia document authoring and presentation. *In: Proc. of the 4th ACM Int. Conf. on Multimedia*, p. 329–340.
6. Chang, S.F., Sikora, T., Puri, A (2001). Overview of the MPEG-7 standard. *IEEE Transactions on Circuits and Systems for Video Technology* 11 (2001) 688–695.
7. S. Christodoulakis, S., Tsinaraki, C., Breiteneder, C., Eidenberger, H., Divotkey, D., Boll, S., Scherp, A., Bertino, E., Perego, A (2005). CoCoMA: Content and Context Aware Multimedia Content Retrieval, Delivery and Presentation. *European Conference on Research and Advanced Technology for Digital Libraries (ECDL) based Digital Libraries. European Conference on Research and Advanced Technology for Digital Libraries (ECDL) 2005*, Vienna, Austria.
8. Christodoulakis, S., Tsinaraki, C., Gioldasis, N., Breiteneder, C., Eidenberger, H., Divotkey, D., Boll, S., Scherp, A., Bertino, E., Perego, A (2006). CoCoMA: Content and Context Aware Multimedia Content Retrieval, Delivery and Presentation. *European Conference on Research and Advanced Technology for Digital Libraries (ECDL) 2006*, Alicante, Spain.
9. Graves, A., Lalmas, M (2002). Video Retrieval using an MPEG-7 based Inference Network. *SIGIR 2002*, Tampere, Finland.
10. Geurts, J., van Ossenbruggen, J., Hardman, L (2001). Application-specific constraints for multimedia presentation generation. *In: Proc. of the Int. Conf. on Multimedia Modeling (MMM01)*, p. 247–266, 2001.
11. McGuinness, D. L., van Harmelen, F (eds.) (2004). *OWL Web Ontology Language: Overview. W3C Recommendation*, Accessed 10 Feb. 2004. (<http://www.w3.org/TR/owl-features>).
12. S. Hammiche, S. Benbernou, M.-S. Hacid, A. Vakali (2004). *Semantic Retrieval of Multimedia Data (2004). 2<sup>nd</sup> ACM International Workshop on Multimedia Databases (MMDB)*, Arlington, VA, USA, November 2004.
13. Hardman, L., van Rossum, G., Bulterman, D. C. A (1993). Structured multimedia authoring. *In: Proc. of the 1st ACM Int. Conf. on Multimedia*, p. 283–289.
14. ISO/IEC: 15938-5:2003 (2003a). *Information Technology – Multimedia content description interface – Part 5: Multimedia description schemes. 2003, First Edition*.
15. ISO/IEC: JTC 1/SC 29/WG 11/N5845 (2003b). *MPEG-21 Multimedia Framework, Part 7: Digital Item Adaptation (Final Committee Draft). ISO/MPEG N5845*, July 2003.
16. Jourdan, M., Layaïda, N., Roisin, C., Sabry-Ismaïl, L., Tardif L (1998). MADEUS, An authoring environment for interactive multimedia documents. *In: Proc. of the 6th ACM Int. Conf. on Multimedia*, pages 267–272.
17. Lux, M., Granitzer, M (2005). Retrieval of MPEG-7 based Semantic Descriptions. *BTW-Workshop “WebDB Meets IR” in context of the “11. GI-Fachtagung für Datenbanksysteme in Business, Technologie und Web”*, University of Karlsruhe, Germany, March 2005.
18. van Ossenbruggen, J., Geurts, J., Hardman, L., Rutledge, L (2003). Towards a formatting vocabulary for time-based hypermedia. *In: Proc. of the 12th Int. World Wide Web Conf.*, p. 384–393, 2003.
19. Pereira, F (2001). The MPEG-21 standard: Why an open multimedia framework? *In: 8<sup>th</sup> International Workshop on Interactive Distributed Multimedia Systems (IDMS 2001)*, LNCS 2158, Lancaster, September 2001, Springer-Verlag, Heidelberg, pp. 219–220.
20. Polydoros, P., Tsinaraki, C., Christodoulakis, S (2006). GraphOnto: OWL-Based Ontology Management and Multimedia Annotation in the DS-MIRF Framework. *Workshop on Multimedia Semantics (WMS)*, Chania, Crete, June 2006.
21. Rogers, D., Hunter, J., Kosovic, D (2003). The TV-Trawler Project. *In: Special Issue of the International Journal of Imaging Systems and Technology on Multimedia Content Description and Video Compression*, March 2003.
22. Tardif, L., Bes, F., Roisin, C (2000). Constraints for multimedia documents. *In: Proc. of the Int. Conf. on Practical Applications of Constraint Techniques and Logic Programming*.
23. Tseng, C.-Y. Lin, B., Smith, J (2004). Using MPEG-7 and MPEG-21 for personalizing video. *IEEE Multimedia*, 11 (1) 42-52, Jan.-Feb. 2004.
24. Tsinaraki, C., Christodoulakis, S (2006). A User Preference Model and a Query Language that allow Semantic Retrieval and Filtering of Multimedia Content. *Semantic Media Adaptation and Personalization Workshop (SMAP 2006)*, December 2006, Athens, Greece.

25. Tsinaraki, C., Fatourou, E., Christodoulakis, S (2003). An Ontology-Driven Framework for the Management of Semantic Metadata Describing Audiovisual Information. 15<sup>th</sup> International Conference on Advanced Information Systems Engineering (CAISE), p. 340-356, June 2003, Klagenfurt/Velden, Austria.

26. Tsinaraki, C., Polydoros, P., Christodoulakis, S (2007). Interoperability support between MPEG-7/21 and OWL in DS-MIRF. Transactions on Knowledge and Data Engineering (TKDE), Special Issue on the Semantic Web Era, 2007.

27. Tsinaraki, C., Polydoros, P., Kazasis, F., Christodoulakis, S (2005). Ontology-based Semantic Indexing for MPEG-7 and TV-Anytime Audiovisual Content. In: (MTAP), Special Issue of on Video Segmentation for Semantic Annotation and Transcoding, *Multimedia Tools and Application Journal* 26, 299-325.

28. Wang, Q., Balke, W.-TW., Kiessling, W., Huhn, A (2005). P-News: Deeply Personalized News Dissemination for MPEG-7 based Digital Libraries. European Conference on Research and Advanced Technology for Digital Libraries (ECDL) 2005, Bath, UK.



**Chrisa Tsinaraki** holds a diploma (1995) and a M.Eng degree (2000) in Computer Engineering. She is a PhD student, has been working as a researcher at the Technical University of Crete and has been involved in many European RTD projects. Her research focus is on interoperability support for multimedia content services, based on MPEG-7/21 and ontologies.



Andrea Perego holds a Post-Doc position at the University of Insubria (Varese, Italy). He received an MA degree in the Humanities, a Master's degree in Computer and Communication Sciences, and a PhD in Computer Science in the University of Milan. His main research interests concern the Semantic Web, Web content rating and filtering, enhanced access and personalization of digital library contents, trust and access control in social networks. Results of his research work have been published in proceedings of international conferences and workshops. He has been involved in the EU projects EUFORBIA and QUATRO, and in the DELOS EU Network of Excellence on Digital Libraries.



**Panagiotis Polydoros** holds a diploma (2004) in Computer Engineering. He is a M.Eng student and works as a researcher at the Technical University of Crete. His research focus is on Multimedia Information Management Systems & Applications.



**Stavros Christodoulakis** is Professor of the Department of Electronic and Computer Engineering of the Technical University of Crete, Greece, and Director of the Laboratory of Distributed Multimedia Information Systems and Applications (TUC/MUSIC). He holds a PhD in Computer Science from the Department of Computer Science of the University of Toronto.



**Athina Syntzanaki** is a student in the Computer and Electronics Engineering Department of the Technical University of Crete and is currently working in her diploma thesis.



Alessandro Martin holds a Post-Doc position at the University of Milan. "He received an undergraduate degree in Social Science and a Master's "degree in Computer and Communication Sciences. His main research "interests concern knowledge management systems, usability and the "Semantic Web. In 2006 he was Visiting Researcher at the Technical "University of Crete. Results of his research work have been published in "proceedings of international conferences and workshops. He has been "involved in the EU projects QUATRO, and in the DELOS EU Network of "Excellence on Digital Libraries."