Ontology-Based Multimedia Presentation Generation

Payam M. Barnaghi
School of Computer Science and Information Technology
The University of Nottingham (Malaysia Campus)
Kuala Lumpur, Malaysia
e-mail: payam.barnaghi@nottingham.edu.my

Sameem Abdul Kareem
Faculty of Computer Science and Information Technology
The University of Malaya
Kuala Lumpur, Malaysia
e-mail: sameem@um.edu.my

Abstract—Multimedia data are illusory entities for the machines. Their contents include interpretable data as well as binary representations. Understanding and accessing the content-driven information for multimedia objects allow us to design an efficient multimedia querying and retrieval system. In this paper, we propose a framework to represent the multimedia information and object roles in order to generate automatic multimedia presentations. The proposed architecture attempts to represent the semantic information and the relations amongst the multimedia objects in a disclosure domain. Thus, the system is domain dependent. The represented data associates with the presentation mechanisms to create an integrated presentation generation system. A multi-layer design defines the various levels of abstraction for the proposed framework.

1. INTRODUCTION

The widespread usage of multimedia data in recent years has led an enormous interest in the research of multimedia databases and support for content based multimedia information retrieval and presentation generation systems. Multimedia data are types of objects which specify real world circumstances such as an event, a place and so on. Sometimes they are provided with manually (or semi-automatic) explanatory information about their features and attributes which is referred as metadata in this paper. Although metadata helps retrieval systems to address the objects in the repository, it is still not sufficient. An efficient querying and retrieval system for multimedia objects should address the semantic information and relations among objects in the repository as well as the manually provided data.

Furthermore, an integrated multimedia generation system should be able to present the query results to the users. The presentation generation process requires special attention to multimedia data. Multimedia data has spatial and temporal characteristics. Temporal considerations are needed for audio, video and other continuous media types. The layout design of the presentation is related to the objects’ temporal characteristics. Hence, generating conceivable and effective hypermedia presentations depend on the system appropriate decisions based on the user’s profile, contextual knowledge and managing the various interdependencies between these choices [1].

In this paper we propose a knowledge conceptualisation structure in order to process the multimedia objects relations and arrange the media items to generate meaningful multimedia presentations. An ontology is used to process the relations between the multimedia objects based on their roles in the presentation scenario. The paper is structured as follows. The next section reviews the layered data abstraction methods and meta-data representation for the objects. Section 3 describes presentation structure for the objects and section 4 reviews the related work. In section 5 an instance scenario is explained. Section 6 provides a conclusion for the paper and describes the future work.

2. INFORMATION REPRESENTATION AND ANNOTATED DATA

A multimedia data model is a collection of concepts that describes the multimedia objects and operations to extract and represent the features of the multimedia data. It should be able to represent logical and physical structures and the relationships between multimedia objects. Vries [2] described the multimedia objects structure in the DBMS by addressing the following three aspects:

- The digitised representation of the object
- Database abstraction
- Content abstraction

A digitised representation of the object describes the multimedia object in the computer. Data abstraction refers to traditional data modeling, representing attributes such as the painter of a portrait. Data abstraction is also used to represent the media dependent structure of the multimedia object. Metadata represents the content of the multimedia object and forms the content abstraction of the multimedia object. It can be a textual description of the semantic content, and also a feature representation of perceptual content which is extracted automatically from the digitised representation object.

This study focuses on information retrieval and presentation. Thus, we assume all the media items are annotated and they include content based meta-data in textual format. The information representation model for the multimedia items employs a multi-paradigm structure in order to describe different aspects of multimedia objects and their related semantics. The structure employs an XML based format which cooperates with an ontology to interpret the contextual relations. The details of data model is discussed in [3].
3. PRESENTATION PROGRESSION PROCESS

The data representation approaches to different views of the annotated multimedia objects. The annotated data which describes the content-driven, media dependent attributes and manually provided information for the multimedia objects. This structure and its related query and retrieval processes provide facilities to address the appropriate objects in the repository. In addition, we need a presentation structure to express the results to the user. The presentation structure must support temporal and spatial characteristics of multimedia data. The role based view to the objects refers to their assigned classes and the proposed character in the desired presentation. The domain ontology represents the relations and concepts of these roles in the disclosure domain. The following describe the system’s work process and elaborates the relationships between different aspects of the proposed architecture.

The system addresses the objects based on the user’s query and adds them to a graph which we call it “results graph”. Then, the referred objects are used to address related objects (using relation schema which is provided in information representation model) [3]. To express the roles, the system refers to the domain ontology and applies the relation rules to the items. Therefore the ontology conceptualises the domain artefacts and their relationships. Hence, the presentation unit needs to deliver the presentation in a meaningful order and structure to the users. Thus, the system searches the XML representation for the requested query and finds the relevant objects. Then it addresses the objects based on the query (direct candidates) and also related objects (indirect candidates). We need to process this structure, delete the repetitions and give weights to each node of the results tree to specify priorities for the presentation sequence (presentation scenario).

In other terms, after addressing the related objects, this information will also be added to the results tree. In the next step the system traverses the results tree and finds the repetitions (addressing an object for several times) and removes the repetitions. The final step is preparing the results tree in a meaningful scenario. The weights are used to define the order. Fig. 1 depicts the results tree generation process. We consider two different approaches based on the user’s preference. They come as static and dynamic scenarios.

In the static scenario, we use predefined templates with unbound variables. For example we define an educational genre which begins with introduction, objectives, details and future works. The system will try to use the knowledge base and ontology for each of the defined stages and assign the weights based on the class they belong to. In fact, the ontology represents the rules regarding to the object roles (for example, we may define that “Mendel” is belonged to “who” and the rules specify that the artefacts in this class have a relation to “History” class in the genetic disorders ontology with a “research work” role).

In the dynamic scenario, the system asks the user about the main target of the presentation. It means that the users clarify an event-based type of the results (user is looking for “what” or, “how” or, “who” or, “when” or, “which”, etc.). The weights for the nodes in the results graph would be assigned based on their relevance to the presentation target event and their description class. Interpreting an object and referring it to an event depends on the system knowledge and ontology definition. In this way if the system could not elucidate the object, it will get the lowest weight for the presentation order. The current version of the system does not provide an implementation for this method. The weight assignments and setting the priorities are still in progress research.

Another issue is defining presentation delivery structure. We define an XML-based representation as our presentation specification structure. The system defines templates with unbound variables. The presentation layer assigns the objects to the templates based on the arranged and normalised presentation structure. After that this structure with an XSLT transformation would be delivered to the user in HTML or any other presentation formats (current version of the system only supports HTML presentations).

4. RELATED WORK

Celentano et al [4] provided a schema modeling for automatic generation of multimedia presentations. Their schema for multimedia reports is based on XML representations. This schema describes the spatial and temporal layers of the multimedia objects. A visual authoring system is incorporated to provide the presentation template. The model is used to support content independent structure processing. For this purpose, the XML-based schema keeps multimedia data and structural information separated from spatio-temporal information. Celentano et al focus on the presentation side and the suggested model uses schemas to tailor a template based on spatio-temporal characteristics of multimedia objects.

The Topia project [5] explores an automatic hypermedia generation system based on semantic representations. In this project an “explanation structure” which is derived from semantic information processing is transformed to hypermedia presentations. In relation with this project, Geurts et al [6]...
provided a document generator to create multimedia presentations. The presentation generation is based on an ontology driven transformation process. The system builds a semantic graph from the information and then, processes this semantic graph. In the next step, the semantic graph is transformed to a presentation structure and subsequently to the multimedia presentation. The proposed method is presented based on an artist’s biography ontology.

There are discourse, presentation and design ontology’s that are cooperating in the presentation generation process. In fact, they use design and representation time knowledge rather than information retrieval level description for the scenario presentation. The domain ontology describes the semantics between the artefacts in the discloser domain. The media items are processed and selected based on a user query and annotated meta-data. The semantics are used to cite the related objects.

Little et al [7] provided a semi-automatic and intelligent multimedia presentation generation through semantic inferring. They proposed a high-level architecture which generates multimedia presentations by using both semantic inferring and multimedia presentation generation tools. The meta-data schema represented as Dublin core [8] and OAI [9] archives and used to inference metadata descriptions and semantic relations. Although Little et al provide a dynamic approach to generate multimedia presentation using AI methods and the presentation model is compatible with public standard like SMIL, but there are some limitations in their approach. Indeed, the Dublin Core is not designed to represent multimedia items and using Dublin Core for metadata description of multimedia items is not an efficient and sufficient approach. There are temporal and spatial attributes for the items which are not addressed through Dublin Core specifications. The system employs AOI archives based knowledge and in this context focuses on “people”. What if the system wants to extend the functionality to “events”, “places”, “objects” and etc?

In terms of inferring rules, the system refers to a set of predefined rules and utilises a mapping table to map the extracted semantic relations to a MPG-7 semantic relationship description form. The functionality could be enhanced using a domain ontology to conceptualise the semantic and rhetoric relations. Thus, the ontology could extend the inferring rules by addressing the concepts and their semantic relations. In this way the functionality of system could be also extend to different sort of queries such as “places” and “events”.

We follow the same idea as Celentano et al [4] to separate different aspects of the multimedia data from each other. Namely, a multi-paradigm schema supports the representation of the multimedia information in abstracted layers. The ontology-based presentation generation approach is similar to the Topia [5] project, except we use a different approach to define the presentation scenario and preparing the result structure. In the next section, using an instance scenario, we discuss how the desired framework provides the presentation progression according to the object representation schema.

5. AN INSTANCE SCENARIO

The illustration given describes a scenario for a medical multimedia information database. In this scenario we assume all of the multimedia data are annotated with contextual meta-data. Fig.2 shows XML-based representation for two instance objects.

```xml
<mm_Objects_Space name="Medical Data Objects">
  <mm_Object id="0" type="image">
    <title>Fabery</title>
    <keyword>Fabery</keyword>
    <keyword>genetics</keyword>
    <keyword>disorder</keyword>
    <keyword>inheritance</keyword>
    <desc>Fabery is an X-linked recessive disease</desc>
    <event type="what">diagnosis</event>
    <class type="disorder"/>
  </mm_Object>
  <mm_Object id="1" type="text">
    <title>Fabery Clinical symptoms</title>
    <keyword>Fabery</keyword>
    <keyword>genetics</keyword>
    <keyword>diagnosis</keyword>
    <keyword>inheritance</keyword>
    <desc>Clinical heterogeneity</desc>
    <event type="what">diagnosis</event>
    <class type="treatment"/>
  </mm_Object>
</mm_Objects_Space>
```

Figure 2. Simplified annotation for instance multimedia objects

In this scenario we only consider some of the attributes for the objects and other information is ignored to simplify the illustration. The query module traverse the XML based structure to find the related data while the data abstraction layer provides facilities to interface between the logical (objects which are represented in XML framework) and the physical (stored objects in database) structures. A sample query for the represented information can be obtained through the search for a specific phrase (for example: “Fabery”). After addressing the related objects a presentation structure must be provided. To achieve this, we offer predefined templates for the presentations and placing the objects in the template. Eventually the query result would be organised in the presentation structure.

Hence, the retrieval and representation processes after query processing and transformation are similar to the methods used by Geurts et al [6]. Geurts et al define the graph structure for the retrieved results and use an ontology to arrange the presentation structure based on the object’s semantics. We also define a Document Object Model (DOM) structure for our results. The presentation preparation component extracts the semantic relations amongst the results nodes based on the system ontology. For each node, the system assigns a relevance weight based on the query and assigned class. Finally, the system rearranges the structure regarding to the weights and a
predefined scenario. As an instance for an educational presentation, the presentation sequence would be the description of "what", "where" or "who" or "when" (or all), "how", and "future" respectively. It means we first provide introductory information about the query. After that the next slides would be about the history and origin of the query object then the documentary information and finally information about the future events. We use a set of sample data for genetic disorders and the system is supposed to automatic presentation for educational and professional purposes. Therefore, the sample sequence for the presentation scenario is based on the system information domain. It is obvious for other domains we need different ontology and also different content and scenario descriptions.

The retrieval process is defined in through the following steps: in the first step, the system searches the XML structure and then selects the objects. In the next step it combines and organises the retrieved objects and eventually the system provides the template based representations.

The proposed solution uses object descriptions and predefined templates to generate the presentations. The multimedia objects with their temporal and spatial constraints define the presentation elements. For example the system describes object A must be played after object B and object C must be played simultaneously with B. In fact, the multimedia object descriptions would be filtered to construct the presentation components. The user’s preferences and feedbacks are other aspects that should be considered in constructing the presentation layout. Fig. 3 depicts the presentation preparation process with a static scenario. Furthermore, the switching elements could be defined as different versions of the selected object based on resolution, size and other like properties or they could be different objects which describe similar facts or closer meanings. Hence, using the switching elements during the presentation gives more adaptability to the system [10]. This relations and concepts could be defined in the relations schema of the model or could be represented through an ontology.

![Diagram of presentation preparation process](image)

**Figure 3.** The presentation preparation process with a static scenario

6. CONCLUSION

We have proposed a framework for information query and retrieval for multimedia objects in order to provide automatic multimedia presentation generation. We focused on a multi-paradigm information description for the multimedia data. The descriptors define a dynamic structure for multimedia objects in the system. Processing and querying the data is based on this dynamic structure. The desired architecture for the framework follows a layered data abstraction approach. This approach enables the system to be independent of the database software and also facilitates scalability and flexibility in the architecture. The layered and component based view to the system gives more flexibility and improves the multimedia generation process based on the represented contents. We have implemented a static mapping to bind the addressed objects to the predefined templates.

At the current stage, we are developing the system using Java 2 platform and PostgreSQL database server [11] to store the BLOB’s and dbXML database [12] to store the XML representations. We intend to improve our data model with an attribute-oriented content representation approach (logical structure) similar to Ren [13]. In this model we plan to use the attributes and relation schema to integrate the structure and specify the dynamic model and define the structure and domain for each of the attributes. The presentation scenario and the arrangement of the retrieved information in a more significant order is also another stage for the improvement of the framework.

REFERENCES


[9] OAI Community, Open Archives Initiative. [Available at](http://www.openarchives.org/).

