A system for multimedia information management and retrieval

Vincenzo Moscato
Dipartimento di Informatica
Via Claudio 21, 80125 Naples, Italy
vmoscato@unina.it

Antonio Picariello
Dipartimento di Informatica
Via Claudio 21, 80125 Naples, Italy
picus@unina.it

Antonio M. Rinaldi
Dipartimento di Informatica
Via Claudio 21, 80125 Naples, Italy
amrinald@unina.it

Abstract

The integrated management of multimedia information such as images, graphics, video, audio, and text, is of great interest in a lot of application fields. The major challenges in this non-trivial task are due to structural, syntactic and semantic heterogeneity of distributed multimedia data repositories. In this paper we present the design of a system, IDEM, based on an object-oriented data model and a wrapper/mediator architecture, which integrates data from heterogeneous multimedia repositories in a single, unified database view, in order to build a multimedia information system capable of integrating data that resides in different data base systems as well as in a variety of non-database data servers.

1 Introduction

The integrated management of multimedia information such as images, graphics, video, audio, and text, is at the moment of great interest in a lot of application fields like Information Retrieval, Office Automation, E-learning, Virtual Museums, Newspaper and Magazines production, Video and Cinema Editing, Medical Applications, Geographical Information Systems Management, Biometric Security Application, and so on. Unfortunately such produced media are not available in a unique content container, but in heterogeneous and distributed repositories as World Wide Web, professional and personal databases of different kinds (Relational Database, Object Database, XML Database, etc.), digital libraries and archives and so on. The major challenges in this non-trivial task are due to structural, syntactic and semantic heterogeneity of distributed multimedia data repositories that make complex the data management processes. To solve such problems it is necessary to define a data model that is capable of representing in unique logical view the multimedia data, together with an architecture able for support in an efficient manner the management of such data. For what concerns the data model, it’s natural to model multimedia data using the object oriented paradigm in order to capture both the different variety of real data and the different related functionalities. This kind of approach may be useful not only for modeling single multimedia structured or unstructured repositories, but especially for integrating different data sources in order to give a simple and powerful mechanism for managing and retrieving the data. For what concerns the architecture, a well known strategy for supporting distributed data integration is the adoption of the “wrappers/mediator” middleware. In this framework we present a system, called IDEM (Integration of Distributed and hEterogeneous Multimedia repositories), wich adopts both the object oriented and wrapper/mediator. The system also supports, by combining textual-based (data are represented as a collection of one or more keywords or with phrases, sometimes called “metadata”, which describe their content) and content-based techniques (retrieval is performed by using the visual content of multimedia data, usually considering some low/medium level features, such as colors, texture, spatial position, objects shape [12]), a semantic search on managed objects that makes more effective the retrieval process.

The paper is organized as follows. In Section 2 the main systems for multimedia data integration are described. In section 3, we outline the data model for multimedia objects management and a functional overview of system architecture, based on Wrapper/Mediator schema. In section 4, we describe the retrieval process for image databases based on a semantic search. In section 5 an example of the system operation, with the experimental protocol and related results, is provided. Eventually some concluding remarks are given in Section 6.

2 Related Works

In the literature, several multimedia integration systems have been proposed. In the following we report a short description of the main systems.

MediaLand [11] is a database system aiming to provide a “true” support for multimedia data man-
The objective of MediaLand is to provide an integrated framework for users with different levels of experiences to manage and search multimedia repositories easily, effectively, efficiently and intelligently. For satisfying these objectives, each multimedia data is represented as a particular “object” (described by apposite metadata) and the correlation among different objects is obtained by means of “links, in order to construct “multimedia object graphs”. In this way, the authors give a unique conceptual structure for describing the multimedia data, successively clustered in domains called “media class. The system presents a 4-tier architecture and supports a multi-paradigm query approach to retrieval aims.

Video Anywhere [9] is a web-agents system capable of providing a syntactic and structural interoperability between the systems that support information providers and information consumers (specifically video asset). To this aim, the core of the system is represented by a special agent called “Web Content Manager”, which manages interactions between video sources and a central database containing the metadata description of the different sources. The retrieval process uses such metadata organized in specific ontologies to obtain more accurate results.

InfoSleuth [2] is an agents system that proposes a semantic approach to provide heterogeneous data integration. In particular, the data integration is obtained extracting a common view of the semantic content from multimedia repositories. This approach gives an independence of requests from information structures, resolving the heterogeneity of data by means of an apposite ontology. InfoSleuth also uses a specific language, KQL (Knowledge Query and Manipulation Language), for the agent communication.

Garlic is a more complete and complex system [8] that, similarly to some of previous projects, uses an object-oriented approach to represent in an uniform way the data from different content servers. But, differently from the other approaches, Garlic provides an efficient query processing and data access layers, for efficiently managing user queries. The data model is based on Odamg – 93 and defines an apposite language for the data definition.

Impact [10] is an agents-based system capable of integrating heterogeneous information. Impact Architecture is based on two entities: “Agents” - software modules created from users or other agents and having high level functions - and “Impact Servers” - representing the services infrastructure created by agents. A Multi-Agents Paradigm allows to integrate heterogeneous information using different agents with particular functions and services. A “yellow pages” mechanism is used to manage the services discovery.

3 Overview of system functionalities

3.1 Data model

In the same vein of MediaLand project, we have adopted an object oriented model for the management of multimedia information: in the literature the advantages of o.o. technology both for multimedia data modeling and for supporting efficient query processing are well known. The proposed data model is based on two fundamental concepts: (i) MultiMedia Object: it is an entity of the multimedia world (a sunset image or a particular movie are example of multimedia objects); (ii) MultiMedia Class: it is a set of multimedia objects sharing the same proprieties (the Image or the Video class are examples of multimedia classes).

Each multimedia object (MMO) can be intuitively described by means of the tuple: MMO=<OID,OL,AS,AML>, O(object) ID(entifier) being the object identifier that distinguishes itself from the other objects, O(object) L(ocator) the physical pointer to locate the actual source of the multimedia data so providing a fast and direct access to data, A(attributes) S(set) the set of attributes (DataBase Attributes, Information Retrieval features or Content Based Retrieval features) instances that describes the object, A(attributes) M(ethods) L(ist) the list of available methods defined on the object. Each multimedia object requires the definition of these elements throughout its system lifetime.

Similarly, a generic multimedia class (MMC) can be described by means of the set: MMC={OID,OL,AS,AML,CR}, O(object) ID(entifier) being the domain of object identifiers that distinguish unambiguously the multimedia objects, O(object) L(ocator) the domain of the physical pointers to locate the actual sources of the multimedia data, A(attributes) S(set) the set of attributes DataBase Attributes, Information Retrieval features or Content Based Retrieval features) defined on the class, A(attributes) M(ethods) L(ist) the implementation of available methods defined on a particular object and Class(R)elationship the set of relationships between the class being specified and the other existing classes.

3.2 Mediator/Wrapper Architecture

The proposed architecture is based on the classic Mediator-Wrapper, also used in [4], and tries to satisfy the main requirements of a multimedia database management system. In this kind of approach, the wrapper explores and examines the several multimedia repositories and send the mediator an appropriate description of the related information in terms of multimedia
classes. From the other side, the mediator receives and organizes these information in order to create a single view on all repositories in order to satisfy the user queries processing. The system architecture has three functional layers: a client layer to submit queries, a mediator layer to manage data, a wrapper layer to extract data.

The Mediator middleware (whose logical architecture is shown in figure 1) has the following functionalities: (i) classify and manage the multimedia class sent by wrapper; (ii) manage the user query; (iii) manage the communication with wrapper systems.

![Figure 1: Mediator-Wrapper architecture](image)

In the classification task, a CLASSIFIER module takes the class-description of multimedia data from the wrappers and, by means of a STANDARD MULTIMEDIA THESAURUS, tries to link each multimedia class to one of the standard classes contained dedicated database called STANDARD CLASS DB; such database contains the set of fundamental multimedia classes managed by the system. If the classification process is successful, the class attributes and the methods are then stored in another system database called METADATA DB, while activating an EXCEPTION HANDLER module in the opposite. In particular the METADATA DB contains the list of effective class-attributes and of all the available methods, for each multimedia repository, and is usefully used to solve the syntactic heterogeneity during query processing. A SEMANTIC MANAGER, based on a MULTIMEDIA KNOWLEDGE BASE, is also used after the classification stage, in order to associate the repository data with a semantic concept, organized in multimedia semantic domains, to be stored in the METADATA DB. To these aims, the module uses the information contained in some standard IR features and CBR features (e.g., for images title, category, color, shape, texture) of multimedia objects. During the query processing task the user queries are submitted by means of a MULTIMEDIA USER INTERFACE; such software component is also used to show related query results. The user queries are then processed by means of a QUERY ANALYZER, whose results and related information are stored in a system database, called QUERY DB. The queries are then “taken and compiled” by means of a QUERY ENGINE, using information stored in the METADATA DB, which sends the appropriate wrappers the given query. The partial and global query results are then stored in another system database called RESULT DB. The results are managed by a RESULT HANDLER and by a TOP-K SELECTION module, which analyzes the results and chooses, by means of appropriate strategies, the best K results and reports them to an OBJECT RECOVERY module. In order to choose the best results also object semantic information (user keywords) are considered by means of a SEMANTIC REFINING module that uses a SEMANTIC NETWORK to discover hidden associations between objects having a similar semantic meaning. The SEMANTIC NETWORK is dynamically generated by means of a general knowledge base (in our case WordNet [7]) and used for measuring the semantic similarity as described in section 4.

The communication with wrappers is carried out by an apposite component called WRAPPER INTERFACE, that send mediator requests and picks up the “meta-information”, in terms of multimedia classes related to the different multimedia repositories, and the related query results according to a XML-based protocol. The mediator system is also provided with an advanced CONFIGURATION MANAGER module. It is used to define, during the initialization phase of the system, the standard classes structure (e.g. selection of standard multimedia classes and of the related attributes and methods) and the system policies (e.g. events managing, Top-K selection strategy, user authentication, etc...). In a such way, the system architecture is fully flexible in different application environments. For some multimedia classes, the use of certain standard methods is necessary. For example, the possibility of executing on the attributes of a relational database “sql queries” is fundamental. To this aim, in our implementation there exist special methods named “default methods”. The CLASSIFIER expects that such method is defined in multimedia class. If a multimedia class is not provided with this method, the Wrapper has the task of implementing on the related Multimedia Repository the desired functionality.

The Wrapper middleware (whose logical architecture is shown in figure 1), entirely developed in JAVA and XML, has the following functionalities: (i) it has to classify and manage the multimedia class defined by the repository administrators, (ii) it has to manage the mediator queries; (iii) it has to manage the
communication with the mediator system. The CLASS EXTRACTOR has the responsibility to create the multimedia class representing the objects (video, images, etc...) managed by the repository. All the information about the attributes and methods and used for the query resolution, are stored in a system database called WRAPPER DB. A RECLASSIFICATION MANAGER is a special module that is necessary when the mediator is not able to classify a given class. Such module is activated by the Analyzer after a mediator request, sending the new class-description to the mediator. A WRAPPER ADMINISTRATOR INTERFACE is used to describe the repositories (i.e. if they are Relational Database, Object Database, Web Page, etc...) a nd the location of data. For example, in the case of relational database, the administrator must indicate the tables and the stored or external procedure which must be exported for the class definition. This interface is also called every time there are some errors in the wrapper automatic tasks.

The query processing task is carried out by means of a QUERY EXECUTOR, activated when the wrapper receives a mediator query request. It analyzes the WRAPPER DB, taking information about the queried class and runs the query, in the optimal manner, on the repository. At the end of the query processing the results are sent to Analyzer which converts them in the communication standard format and activates the Communicator for the data transfer. The communication with mediator system is performed by means of a MEDIATOR INTERFACE module, that is composed of two fundamental sub-parts: the Communicator and the Analyzer. The Communicator has the task of managing the physical communication with the Mediator. From the other side, the Analyzer has the task of interpreting the mediator requests.

We want to notice that the Wrapper Administrator has the possibility of creating complex multimedia classes from one or more classified classes, by means of application of class-relationship constructs. In this case the new class must be classified by means of a matching with the possible system standard complex multimedia classes.

Eventually, for what concerns multimedia queries supported, referring to a given multimedia class, the following queries can be expressed: (i) queries with “where” clause; (ii) keywords-based queries; (iii) queries with conditions on multimedia contents (content-based query); (iv) queries with Multimedia Objects (query by example).

The first class of queries can be solved by the classical SQL approach, the second class with an Information-Retrieval approach, while for the last ones the use of techniques for multimedia data management is required.

4 Annotated image matching

Even if our system is capable of managing all kinds of multimedia data, we have primary directed our efforts towards integration of image repositories. In this section we describe an innovative metric which combines semantic information, based on textual annotations, with image content features, to perform a more accurate retrieval process on image repositories.

From one hand, the first component of the discussed metric allows to determine a grade of relatedness between user search-keywords and the multimedia object description by using a semantic net generated in an automatic way thanks to WordNet. The construction of this net is exploited dynamically considering the information and the structure of WordNet (i.e., in WordNet the terms are organized through their linguistic properties: each term can have different meaning or sense, “polisemy”, depending on the topic area, and, each sense is then organized in “synsets” constituted by synonyms). To this aim, the system supplies an interface that helps the user in choosing the “right sense” of a term through the description retrieved from the WordNet structure (gloss). Once chosen the sense and the appropriate synset, it is possible to build a first core of the semantic net being considered all the terms contained in the synset; successively, by exploiting the other WordNet linguistic properties related to the type of a given term (e.g., names, adjectives, verbs, adverbs), the semantic net can be extended obtaining a strongly connected net, in which the relation between the different terms are labeled using some normalized weights that take in account the “strength” of the relation. For measuring the correlation between terms in the image metadata and user data we extend a metric defined in [6]:

$$S_{t_i,t_j} = e^{-\alpha l} e^{\beta d} - e^{-\beta d}$$

where $\alpha \geq 0$ and $\beta > 0$ are two scaling parameters whose values have been defined by an experimental setup and $i$ and $j$ are the indexes of the considered terms. The above equation is expressed by combining the normalized distance between two terms $l = \min_j \sum_{i=1}^{h_j} \frac{1}{\sigma_j}$, where $\sigma_j$ spans over all the paths between the two considered terms, $h_j$ is the number of hops in the $j$-th path and $\sigma_j$ is the weight assigned to the type of relations in the $j$-th path, and, the depth of its subsynset $d$ from the root of WordNet hierarchy ($d$ is computed using WordNet and considering the IS-A hyponymy-hyperonymy hierarchy only).

Eventually the final relatedness measure is obtained by combining $S_{t_i,t_j}$ quantity with the weight of a given term $t_i$ calculated as $w_{t_i} = \frac{1}{\text{poly}(i)}$, being $\text{poly}(i)$ is the polisemy grade of $t_i$.
More details about the semantic net construction and the used metric are reported in [1]. Such approach is useful to solve a semantic heterogeneity in the description of the multimedia objects since the user can specify, in its query, keywords not directly present in the metadata but related to them by linguistic relations.

From the opposite hand, to evaluate the “content similarity” between two images, the second compo-

nent of the discussed metric, we have used a normalized metric $\mu_{i,j}$ based on Animate Vision theory, also adopted in the video segmentation process [3]. Such content metric takes into account the flow of image salient points, called “information path” and uses color histogram and wavelet based techniques to determine the similarity degree between different images.

5 Experimental Results

From one side, we suppose the presence of two multimedia repositories containing some scenic (sunsets, deserts, coasts, etc...) and animals (birds, fishes, mammals, etc...) digital images. In particular the first repository is a MySql 4.0.16 Relational Database running on a Linux RedHat 7.3 operating system. The second repository is an Access XP Relational Database running on a Windows XP operating system. The wrapper functionalities are provided with the USER INTERFACE. It is possible to: (i) register a new wrapper and configure the wrapper for repository communication; (ii) create a multimedia object; (iii) set communication parameters with mediator.

The definition of the “Image” multimedia class is performed by the OBJECT CREATOR that proposes to the wrapper administrator the possible class attributes and methods in terms of database tables, table fields and stored procedure. The class tree structure is after converted in a XML format which is sent to mediator with the related data, necessary to associate to the object a semantic meaning. Eventually the description of the object is stored in the WRAPPER DB. Possible classification errors are solved by the RECLASSIFICATION MANAGER and WRAPPER ADMINISTRATION modules.

The query processing to/from mediator is managed by the QUERY EXECUTOR: such module uses the information in WRAPPER DB in order to translate mediator “SQLWhere” requests in the local DBMS SQL format and the ANIMATE VISION module to retrieve the similarity measures between the target image and the DB images. Eventually the query results and the related scores are sent to the mediator. From the other side, we suppose the the mediator is able to manage a unique standard class called “Image”.

The mediator functionalities are provided with two different graphical interfaces. By means of the USER INTERFACE, it is possible to: (i) submit a multimedia query; (ii) view the query results; (ii) set some communication parameters in order to optimize the transmission flow towards the user. By means of the ADMINISTRATOR INTERFACE, it is possible to: (i) define the standard classes managed by the system; (ii) define the system policies (e.g., events managing, wrappers managing, Top-K strategies, data flow security, etc...); (iii) set general parameters (multimedia thesaurus, possible object resolutions for devices transmissions, etc...). The main task of mediator is to classify the classes originated by the two wrappers. If the classification has a success, the mediator inserts the new class in the METADATA DB. The SEMANTIC MANAGER tries to associate a semantic concept to the wrapper objects and the semantic associations with the other objects. The semantic concept is obtained for multimedia semantic domain and by object feature values (shape, color, texture), choosing the association rule with the maximum probability. The semantic object relationships are obtained by using the described SEMANTIC NETWORK.

The query processing is carried out in three phases: (i) by means of USER INTERFACE all data (standard class, attributes and methods constraints, Top-K dimension, query weights), necessary to the query execution, are picked up and stored in the QUERY DB; (ii) the query, stored in the QUERY DB, is compiled by the QUERY ENGINE and sent to wrappers; (iii) the query results are picked up form Wrapper and, by means of the TOP-K SELECTION [5], the best ones are shown to the user. During the download, as specified in the user settings, the object resolution is adapted to the user device type.

Eventually, some user query cases are reported. The query processing is performed by means of the FindSimilarity and SQLWhere methods and by expressing one ore more search keywords that lead the described semantic retrieval. For each case, the query set up parameters (see table 1) and the query results (see figure 2), in terms of multimedia objects, matching scores and wrapper sources, are shown (in the second query the image is not present in any wrapper).

<table>
<thead>
<tr>
<th>Query</th>
<th>Content Similarity</th>
<th>SQL</th>
<th>Semantic</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.55</td>
<td>0.4</td>
<td>0.15</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>0.4</td>
<td>0.1</td>
<td>0.5</td>
<td>5</td>
</tr>
</tbody>
</table>
6 Conclusions

A multimedia data integration system is presented. It allows a single unified database view from more multimedia repositories. The proposed data model describes a generic multimedia entity as an object belonging to a particular class with similar features and general methods. The system architecture is based on the Mediator/Wrapper schema in order to have a fine description and organization of multimedia data. The user query is processed by ad hoc module and sent to appropriate wrappers. The query results are calculated taking into account their semantic meaning and using a dynamic multimedia semantic network. The reported preliminary results show the efficiency and effectiveness of the proposed architecture performances.

7 Acknowledgment

This work has been carried out partially under the financial support of the Ministero dell’Istruzione, dell’Università e della Ricerca (MIUR) in the framework of the FIRB Project “Middleware for advanced services over large-scale, wired- wireless distributed systems (WEB-MINDS)”

References