MULTIMEDIA ARCHIVES AND MEDIATORS

Manolis Wallace, Yannis Avrithis and Stefanos Kollias
National Technical University of Athens, GREECE

During the last decade, the cost of storage and wide area communication services has decreased, while their capacity increased dramatically. This fact, along with the increasing penetration of e-commerce applications, has made digital storage, annotation and access of multimedia information a mature and viable choice for content holding organizations and individuals. Numerous multimedia archives have, either totally or incrementally, turned to the utilization of digitized archive technologies. The content of these archives can be made accessible, depending upon copyright, policy and security decisions, over the Internet in a cost efficient, time efficient, anyplace, anytime fashion.

However, one of the main problems of multimedia archiving has been inherited to their digital descendants. For traditional archives, where raw media were stored in the form of analog hard copies, search was not an easy task as a human had to either go through a separate annotation archive, or, ideally, search using keywords in a custom, proprietary, metadata database. Much similarly to the case of books in a library that have not been indexed, information stored in a multimedia archive that cannot be searched, identified and accessed easily is practically unavailable.

Trends in research and standardization

In order to provide for more efficient search services in the always augmenting space of available digital multimedia content, several systems have been proposed and several research projects and initiatives have been funded, making important contributions to theoretical fields ranging from multimedia signal processing, computer vision, multimedia database and knowledge management to artificial intelligence, human computer interaction and information retrieval. Still, considering the number and diversity of multimedia archives existing worldwide, being able to search in each one of them independently but not in all of them at the same time through a common interface is much like having independent indices for each corridor in a library. When the library becomes larger, data is once again as good as non existing.

Current and evolving international standardization activities, such as MPEG-4 [1] for video, JPEG-2000 [2] for still images, and MPEG-7 [3], MPEG-21 [4], SMIL [5] for generic multimedia, deal with aspects related to audiovisual (a/v) content and metadata coding and representation, aiming to provide a framework for uniformity and interoperability between developed systems. Still, mainly due to the fact that digital archives have pre-existed the developed standards, very few of them fully comply with them. In most cases, multimedia archives operate using proprietary data structures as well as administrator and end user software. Thus, the integration of multimedia archives through a common, unified access point for end users, always considering their particular copyright and access policies, emerges as a necessary step for the preservation of their content and their financial viability.

In order to achieve this goal, several research activities are currently active in the direction of knowledge acquisition and modeling, capturing knowledge from raw information and multimedia content in distributed repositories to turn poorly structured information into machine-processable knowledge [6][7]. A second future direction is knowledge sharing and use, combining semantically enriched information with context to provide infrencing for decision support and collaborative use of trusted knowledge between organizations [8]. Finally, in the intelligent content vision, multimedia objects integrate content with metadata and intelligence and learn to interact with devices and networks [9].

It is becoming apparent in all the above research fields that integration of diverse, heterogeneous and distributed – pre-existing – multimedia content will only be feasible through the design of mediator systems. In [10] for instance, a multimedia mediator is designed to provide a well-structured and controlled gateway to multimedia systems, focusing on schemas for semi-structured multimedia items and object-oriented concepts, while [11] focuses on security requirements of such mediated information systems. On the other hand, [12] deals with media abstraction and heterogeneous reasoning through the use of a unified query language for manually generated annotation, again without dealing with content or annotation semantics. A semantically rich retrieval model is suggested in [13], based on fuzzy set theory with domain-specific methods for document analysis and allowing natural language queries. Finally, [14] focuses on the design of a single intuitive interface supporting visual query languages to access distributed multimedia databases.

Mediator systems

A mediator is "one who resolves or settles differences by acting as an intermediary between two conflicting parties", "an impartial observer who listens to both sides separately and suggests how each side could adjust its position to resolve differences", "a negotiator who acts as a link between parties" and so on, depending on which dictionary one consults. In information systems, a mediator is a independent system residing between the end user and an information system. Its role is to provide a kind of translation between the two ends, by setting up a user friendly interface towards the user which serves information acquired from the typically more rigid information system [15][16].

In the case of multimedia archives, the main hurdle to overcome is not the rigidness of the end systems, but rather the fact that numerous representation standards have been developed up to date. Even if the MPEG series seem to be gaining momentum and setting themselves up as the definitive choice for the future, numerous existing archives follow other standards, or even totally proprietary representation and storage models. As a result, unified access is not feasible. The role of a mediator in this framework is to provide a single interface to the end user, while at the same time making sure that the user query is properly translated in order to be communicated to all sorts of different multimedia archives, thus offering unified multimedia access services. The typical mediator follows a 3-tier architecture such as the one depicted in figure 1. The central idea in its design is the strict and formal definition of the interface points between the three components.

From the *developer*'s point of view (*presentation-tier*) this allows for the design of multiple, customized user interfaces that are all able to connect and interact with the central system. This way, for example, the same mediator can be used by uses in different countries, by simply providing user interfaces with different languages. This can also allow for the generation of different user interfaces for first time users and experts, as well as for customized user interfaces for users with different access rights.

From the *researcher*'s point of view (*application-tier*) this allows for the constant update and extension of the core system with the inclusion of knowledge and intelligent processing modules. The application-tier is the processing heart of the mediator where all intelligence is included. The definition of formal and strict interfaces to the other tiers allows for the standardization of the data accepted as input and provided at the output. By choosing a flexible and descriptive data model for the application-tier we allow for maximum extensibility though the integration of emerging and future technologies in the processing stages, as, for example, contextual and ontological reasoning.

Finally, from the *archivist*'s point of view (*data-tier*) this allows for the easy inclusion of more archives. Having a strictly defined interface between the application-tier and the data-tier allows for

- i) the participation in the same system of archives that are diverse in storage structure,
- ii) the inclusion of more archives at a time later than the design and development of the overall system, with minimal cost and effort, and
- iii) the development of new archives that are able to be connected directly to the mediator system

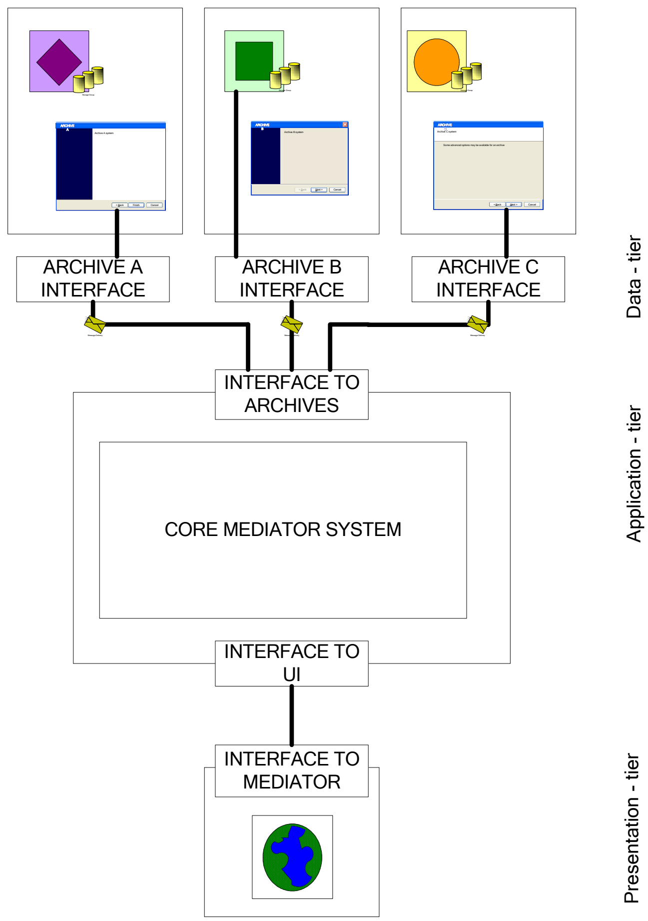

Figure 1. A typical 3-tier architecture for a multimedia mediator.

Unified access: mediators vs. integration

A good way for one to fully grasp the importance of mediator technology is to consider its predecessor. In the case of a business take over, or in strategic alliances, archives had the need to have and/or to offer unified access to their data. Prior to mediator systems this need for unified access was satisfied through data integration. In this approach, existing data from two, or more, existing systems had to be ported to a single system.

The storage structure of this system would ideally be flexible enough to support all annotation features existing in either of the participating archives. Then, in the merging phase, all existing data had to be processed, manually to the greatest part, in order for the existing annotation to be transformed into the supported structure. In practice, due to the fact that this manual processing is strenuous, time consuming and expensive, the structure chosen was typically that of the largest – in amount of contained data – archive, thus minimizing the size of the manually processed data.

The data integration approach to data unification has important drawbacks, most of which are rather obvious:

i) the cost of data integration is often prohibitive, so unified access is only provided in the limited cases that the foreseen financial benefits are important,
ii) the storage structure chosen for the integrated system is not the optimal one, so there is limited descriptive power and limited support for future extension,
iii) the integration of more archives at a later time is not straight forward and
iv) the integration of large scale archival systems, or equivalently of large numbers of archival systems, is not practically and financially feasible.

What is most important is that archives are required to either abandon their existing systems, thus loosing the ability to serve a part of their existing clientele, or maintain duplicated data in both systems, thus having to suffer the burden of manually assuring that the two versions remain identical trough all future updates and insertions of new data. Mediator technology provides an answer to all these issues, by separating the notions of data and access integration; mediators allow for unified access while at the same time avoiding the costly and problematic procedure of integrating archived data into a single system.

Syntactic vs. semantic integration

The role of a mediator system is to allow for unified access, though a single user interface, to multiple data storage locations. In a multimedia environment, the typical role of a mediator system to date has been to allow for unified access to archives that are possibly diverse in storage structure; this is often referred to as structural or syntactic integration. With the development of larger, international and multicultural multimedia markets, syntactic integration is no longer sufficient. Emerging multimedia mediator systems take a step further and also aim for semantic integration [17]; this refers to the integration of systems that are not only diverse in structure, but also in nature – and possibly even in annotation language.

The desired goal is to be able to provide a system that is able to connect to any type of multimedia archive, be it a movie archive, news archive, audio archive, multimedia library etc. and allow end users to uniformly query all available data, much like Google for web pages. The difference from Google is that considered data is available in structured annotated forms and the annotation needs to be exploited to the greatest possible extent in order to be able provide optimal access services to end users.

Of course, the uniform exploitation of annotation information that is not available in a uniform format is not trivial. As a result, a main difference of this new breed of mediator systems is in the way they process available information. Most existing mediators rely on methodologies borrowed from textual and multimedia retrieval systems for their retrieval and indexing procedures, respectively. Emerging multimedia mediators rely on methodologies from knowledge engineering and ontologies for retrieval, and multimedia ontologies for indexing purposes, on order to be able to consider this diverse annotation information.

The *Faethon* multimedia mediator system is a characteristic example of semantic oriented multimedia mediators [18]. The abstract architecture of the system is depicted in Figure 2. The *interface to archives* (data-tier) is at the top and the user interface (presentation-tier) is at the bottom – the interaction components are in charge of accepting data and forwarding to the core system while the presentation components are in charge of accumulating data from the core system (application-tier), formatting it and forwarding to the communications component.

The system operates in two distinct and parallel modes: query and update. In *query mode*, the end user is offered the ability through the UI to specify a *syntactic query* – a query based on strict annotation details of the desired documents specified through the MPEG-7 descriptors, e.g. creation information like title or production date – as well as a *semantic query* – a free text query that the mediator will have to process, interpret semantically and match to document semantics. The syntactic query is forwarded to the search engine and handled with traditional multimedia retrieval techniques. Specifically, the search engine forwards the query to the participating archives and merges the responses of those archives that respond in a timely fashion. The semantic query is received, interpreted, analysed and expanded by the query analysis module. Matching to documents is again performed by the search engine, relying on indexing information already stored in the DBMS of the core system during update mode. The overall response – considering both the semantic and metadata part of the user query – is then personalized and re-ranked based on the user profile and returned to the user.

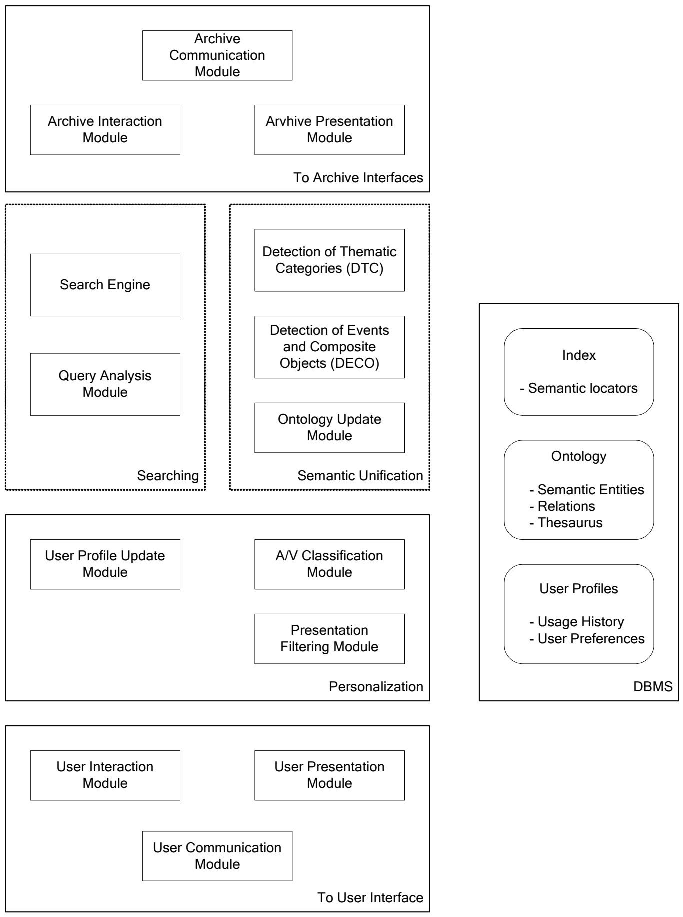

Figure 2. Abstract architecture of the Faethon multimedia mediator system.

Of course, in order for the query analysis module to be able to provide a response, an appropriate indexing of documents needs to have already been performed. This is performed in the offline, *update mode* of operation by the DTC (Detection of Thematic Categories) and DECO (Detection of Events and Composite Objects) modules. It is worth mentioning that the offline mode is not a mode of operation that requires the system to be unavailable; the "offline" term simply refers to the fact that there is no strict time frame within which it is imperative for the participating processes to have provided their output, and thus more sophisticated algorithms may be utilized. In fact the DTC and DECO modules are always active, indexing new documents or refreshing indexing information based on newly acquired knowledge, much like web crawlers of internet search engines that continuously crawl the web indexing new pages or refreshing the indexing information for known ones. In the same mode, accumulated usage information can be processed for the update of the user profiles for personalization purposes.

All of the processes of the core system, both online and offline, are heavily dependent on the knowledge stored in the ontology which allows for intelligent, semantic treatment of data that is independent of the annotation style and language. Other system components include the index, ontology update module, and the personalization subsystem along with the user profiles. These are discussed in [19], [20] and [21].

References

1. Battista S., Casalino F., Lande C., "MPEG-4: A Multimedia Standard for the Third Millenium", Part 2. IEEE Multimedia 7 (1), pages 76-84, 2000
2. ISO/IEC JTC1/SC29/WG1 N1646R: JPEG 2000 Part I Final Committee Draft Version 1.0, 2000
3. Sikora T.., "The MPEG-7 Visual standard for content description - an overview", IEEE Trans. on Circuits and Systems for Video Technology, special issue on MPEG-7, 11(6):696-702, June 2001
4. MPEG-21 Overview v.3, ISO/IEC JTC1/SC29/WG11 N4511, December 2001
5. Synchronized Multimedia Integration Language (SMIL). Recommendation REC-smil-1998 0615, World Wide Web Consortium (W3D), 1998. http://www.w3.org/TR/REC-smil.
6. M. Ramesh Naphade, I.V. Kozintsev, and T.S. Huang, "A factor graph framework for semantic video indexing", IEEE Trans. on Circuits and Systems for Video Technology, Volume 12, Issue 1, Jan. 2002, pp. 40-52.
7. O. Mich R. Brunelli and C.M. Modena, "A survey on video indexing", Journal of Visual Communications and Image Representation, Volume 10, 1999, pp. 78-112.
8. J. Euzenat, D. Fensel, R. Lara and A. Gómez-Pérez, "Knowledge Web: Realising The Semantic Web, all The Way to Knowledge Enhanced Multimedia Documents", Proc. of European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology (EWIMT), London, U.K., November 25-26, 2004.
9. I. Kompatsiaris, Y. Avrithis, P. Hobson and M.G. Strinzis, "Integrating Knowledge, Semantics and Content for User-Centred Intelligent Media Services: the aceMedia Project", Proc. of Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Lisboa, Portugal, April 21-23, 2004.
10. J. Biskup, J. Freitag, Y. Karabulut, and B. Sprick, "A mediator for multimedia systems", Proceedings 3rd International Workshop on Multimedia Information Systems, Como, Italy, Sept. 1997.
11. C. Altenschmidt, J. Biskup, U. Flegel and Y. Karabulut, "Secure mediation: requirements, design, and architecture", Journal of Computer Security, Volume 11, Issue 3, March 2003, pp. 365 – 398.
12. A. Brink, S. Marcus, V. Subrahmanian, "Heterogeneous Multimedia Reasoning", IEEE Computer, Volume 28 , Issue 9, Sept. 1995, pp. 33-39.
13. I. Glöckner, A. Knoll, "Natural Language Navigation in Multimedia Archives: An Integrated Approach", Proceedings of the 7th ACM International Conference on Multimedia, Orlando, FL, pp. 313-322.
14. I. Cruz, K. James, "A user-centered interface for querying distributed multimedia databases", Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data, Philadelphia, PA, 1999, pp. 590-592.
15. Rao B.R., "Making the Most of Middleware", Data Communications International 24, 12 (September 1995):89-96
16. Steinke S., "Middleware Meets the Network", LAN: The Network Solutions Magazine 10 ,13 (December 1995):56
17. M. Wallace, T. Athanasiadis, Y. Avrithis, A. Delopoulos, S. Kollias, "Integrating Multimedia Archives: The Architecture and the Content Layer", Systems Man and Cybernetics 2005.
18. Avrithis Y., Stamou G., Wallace M., Marques F., Salembier P., Giro X., Haas W., Vallant H., Zufferey M., "Unified Access to Heterogeneous Audiovisual Archives", Journal of Universal Computer Science 9(6), pp. 510-519, 2003
19. M. Wallace and G. Stamou, "Towards a Context Aware Mining of User Interests for Consumption of Multimedia Documents", Proceedings of the IEEE International Conference on Multimedia (ICME), Lausanne, Switzerland, August 2002
20. M. Wallace, G. Akrivas, G. Stamou and S. Kollias, "Representation of user preferences and adaptation to context in multimedia content -- based retrieval", Proceedings of the Workshop on Multimedia Semantics, SOFSEM 2002: Theory and Practice of Informatics, Milovy, Czech Republic, November 2002.

21. S. Bloehdorn, N. Simou, V. Tzouvaras, K. Petridis, S. Handschuh, Y. Avrithis, I. Kompatsiaris, S. Staab and M. G. Strintzis, Knowledge Representation for Semantic Multimedia Content Analysis and Reasoning, in Proc. of European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology (EWIMT 04), Lon-don, U.K., November 25-26, 2004.