SEMANTIC UNIFICATION OF HETEROGENOUS MULTIMEDIA ARCHIVES[*]

G. STAMOU, Y. AVRITHIS AND S. KOLLIAS

Institute for Communication and Computer Systems
Department of Electrical and Computer Engineering
National Technical University of Athens, Zographou Campus, Greece
E-mail: gstam@softlab.ntua.gr

F. MARQUES AND P. SALEMBIER

Signal Processing and Communications Department
Technical University of Catalonia (UPC)
Barcelona Spain
E-mail: ferran@gps.tsc.upc.es

Multimedia Content is described via textual, semantic and structural Descriptors and Description Schemes, as introduced in MPEG-7. The semantic part of the description is closer to what the user expects from a multimedia search engine, however it poses difficulties, because of the potential incompatibility of the semantic entities among different archives. In this paper, we present FAETHON, a system that unifies the semantic description of heterogeneous archives, through the use of a semantic encyclopaedia.

1. Introduction

The Description Schemes (DSs) of MPEG-7 [6] are organized, on the basis of their functionality, into groups of DSs, among which the *Content Management* and *Content Description* are the most important in terms of description of multimedia documents. Content Management DSs contain information concerning the *Creation/Production, Usage* and *Media* information of the content, and consist of simple data types (numbers and alphanumerics), which have been being stored in traditional databases. On the other hand, Content Description DSs are classified into *structural* and *conceptual* (semantic) [3,4]. The structural DSs describe the content in terms of segment decompositions, in space (such as moving and still regions) and in time (such as shots), also giving descriptors (such as color and shape). On the other hand, the conceptual DSs describe the content in terms of semantic entities that appear in the multimedia document such as objects, events, concepts and relations among them.

---

[*] Work supported by the EC IST-1999-20502 FAETHON project.

The FAETHON system [1] performs the *semantic unification* of different archives by using a *semantic encyclopaedia* containing definitions of abstract classes. Creation of the encyclopaedia relies both on human experts and existing ontologies. The system correlates the specific semantic entities of the multimedia descriptions of the individual archives to the abstract ones of the encyclopaedia. When a user makes a query, the supplied keywords are translated into the semantic entities of the encyclopaedia. The documents whose descriptions have been correlated to the requested semantic entities are retrieved, and then filtered by taking into account the user's preferences in each semantic entity.

2. The architecture of FAETHON system

From a technical point of view, the overall functionality of the proposed system is to offer a unified interface that will allow end-users to have efficient access to a number of individual audiovisual archives. In this sense FAETHON acts as an intermediate agent that undertakes the tasks of: receiving end-users queries; translating the terms of the query in a set of semantic entities by means of the knowledge of the system; searching the audiovisual archives for the existence of the above semantic entities; receiving the produced responses; presenting the latter to the end-user in order of importance, ranking them using the user profile. The above procedure is the typical information flow in a 3-tier environment. What is innovative in this flow is the semantic level of the resolution of the users' queries. Based on the above process FAETHON's users will be able to issue expressive semantic queries whose answers comprise "understanding" of the involved semantic entities and rapidly converge to the focus, i.e. to what the end-user has in mind understanding the context of the query by also using the information of the user profile [7].

The FAETHON system, as a web service, enables a user to perform a single query on multiple multimedia archives and receive the results in a uniform manner. Its operation has two distinct modes: the query and the update mode (working in parallel). In the query mode FAETHON system serves its end-users by exploiting (a) its already available knowledge, (b) pre-processed information previously extracted from the audiovisual archives and (c) the on-line access to the latter using the user profile and relevance feedback for presenting the information in order of importance to the specific user. Figure 1 presents the system operation in query mode. In its update mode, FAETHON system enhances its knowledge and gathers information from the audiovisual archives, processes this acquired information and stores it for subsequent use. Moreover,

it updates user profiles translating the usage history into user preferences. Figure 2 presents the system operation in update mode.
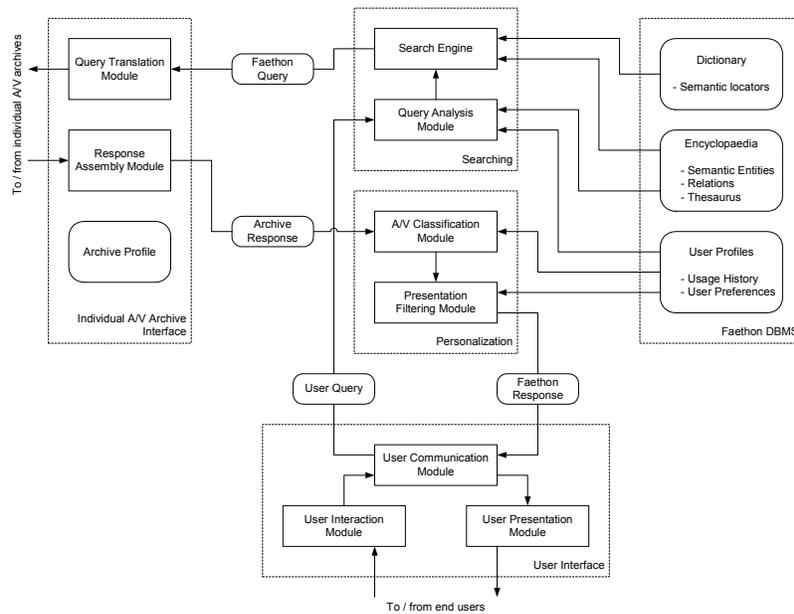

Figure 1. The FAETHON system at query mode of operation.

3.   The knowledge of the system

The knowledge of FAETHON contains the encyclopaedia and the user profiles. The encyclopaedia comprises a set of description schemes for the definition of all semantic entities and their relations. It is, therefore, structured as a graph containing all knowledge of semantic information used in the system. Among other actions, it allows: structured storage of semantic entity descriptions and relations defined by experts to be used for indexing and retrieval purposes; forming complex concepts and events (composite entities) by the combination of simple ones through a set of previously specified relations; expanding the user query by looking for synonyms or related concepts to those words contained in the semantic part of the query. The semantic entities in the encyclopaedia are mostly (media) abstract notions in the MPEG-7 sense. Media abstraction [5] refers to considering a semantic description of a specific instance of multimedia content (e.g., "a soccer player" in a specific image) and

generalizing it to multiple instances of multimedia content (media abstraction, e.g., "a soccer player" for any picture or video). Each semantic entity can contain textual annotation, including keywords defining the entity, sub-entities and their relations and low-level descriptors. The description of all relationships among the SEs in the encyclopaedia, using the semantic relations, forms a graph structure. The graph nodes correspond to the SEs, whereas graph links represent the type of relationship between the nodes connected by them. This graph structure is represented in FAETHON by means of SemanticEntities DS and SemanticRelations DS. All relations are, in principle, fuzzy, and each relation value can be any number between 0 and 1.
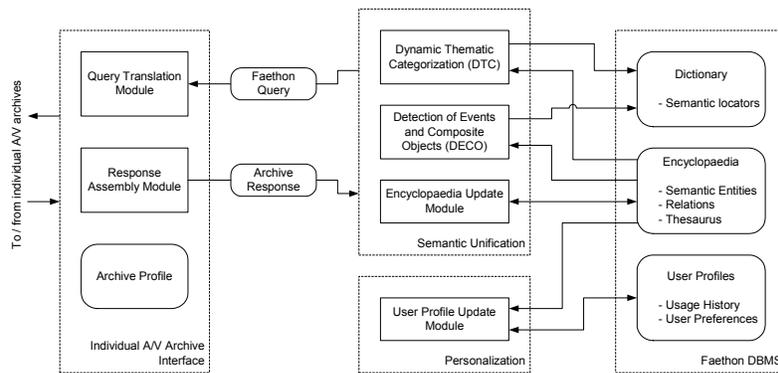

Figure 2. The FAETHON system at update mode of operation.

4. Semantic unification and knowledge-based search

The role of semantic unification is to correlate the multimedia document descriptions provided by the archives with the semantic entities stored in the encyclopaedia. The result is, on one hand, the semantic index, containing the correlations between multimedia documents and semantic entities, and on the other hand, the thesaurus. The semantic unification is performed with the aid of the Detection of Thematic Categories (DTC) module and the Detection of Events and Composite Objects (DECO) (Figure 2). The *DTC module* maps the description of a multimedia document to the set of thematic categories, and stores the relevance values in the semantic index. It makes the correlation by matching the description of the document with the description of the thematic category. The *DECO module* maps the description of a multimedia document to

the set of semantic entities. It is similar in operation to the DTC module, but it is capable of matching composite semantic entities (entities that contain sub-entities) to sets of consecutive DSs found in the description of the document. The *semantic index* produced by DTC and DECO provides a fuzzy mapping between the set of the semantic entities of the encyclopaedia and the set of the document locators of the documents of all the archives. It therefore contains degrees of relevance between semantic entities and documents. It is used to locate documents that match the user query without searching in the archives at query time. The *searching* procedure is based on the Semantic Index. It takes as input the keywords that consist of the semantic part of the user's query and the metadata the user has provided. Its output is a number of document locators, ordered by their relevance value. The processing of the user's query consists of the query interpretation and the query expansion phases. In *query interpretation*, each keyword is transformed into a fuzzy set defined on the set of semantic entities. On the other hand, in *query expansion*, the above sets are expanded using the information of the fuzzy thesaurus. Finally, the *search engine* uses the semantic entities involved in the expanded query and returns the associated document locators based on the information of the semantic index.

References

1. Y. Avrithis and G. Stamou, "FAETHON: Unified Intelligent Access to Heterogeneous Audiovisual Content," in Proc. of Int. Workshop on Very Low Bitrate Video Coding (VLBV), pp. 215-216, Athens, Greece, October 11-12, (2001).
2. Akrivas, G., Stamou, G.: Fuzzy Semantic Association of Audiovisual Document Descriptions. In Proc. of Int. Workshop on Very Low Bitrate Video Coding (VLBV), Athens, Greece, (2001).
3. Nack, F., Lindsay, A.: Everything You Wanted to Know About MPEG-7: Part 1. IEEE Multimedia **6** (3) 65-77 (1999).
4. Nack, F., Lindsay, A.: Everything You Wanted to Know About MPEG-7: Part 2. IEEE Multimedia **6** (4) 64-73 (1999).
5. Special Issue on MPEG-7. IEEE Trans. On Circuits and Systems for Video Technology **11** (6) 685-772 (2001).
6. ISO/IEC JTC1/SC29/WG11 N4032: Introduction to MPEG-7. Singapure (2001).
7. B. Ganter and R. Wille, Formal Concept Analysis: Mathematical Foundations, Springer-Verlag, (1999).