PERSONALIZED CONTENT BROWSING BASED ON NOTIONS OF CONTEXT

Manolis Wallace[1] and Kostas Karpouzis[2]

[1] Department of Computer Science
University of Indianapolis, Athens Campus, Greece
wallace@uindy.gr

[2] Image, Video and Multimedia Laboratory
Department of Electrical and Computer Engineering
National Technical University of Athens – Greece
kkarpou@softlab.ece.ntua.gr

ABSTRACT

In this paper we propose a system that facilitates users to browse through related multimedia documents in an efficient manner. The system utilizes the notion of Electronic Roads to make content browsing more effective; that is, it provides them with intelligent personalized content retrieval services, using relevance feedback together with a novel document clustering algorithm. The above notions are integrated in a system for browsing of culture – related multimedia documents.

INTRODUCTION

When searching for information on a specific subject, most users input a number of textual search terms in a search engine; the search engine then retrieves a number of documents from its index that match the search terms and uses the degree of matching in order to determine their ranking. The individual retrieved documents selected in this way are not necessarily related to each other and, therefore, represent solitary pieces of information, without any notion of semantic continuity.

A newer approach suggests that users are already familiar with the semantic meaning of available information but are interested in only a particular subset of it during each session. As a result, whenever users retrieve a specific information unit, the browsing system suggests that they continue the retrieval process by viewing material that is somewhat semantically related. As browsing goes on, this process forms an abstract chain of documents, which share a common semantic substance. This personal chain of documents can be thought as the users' own *Electronic Road* through a sea of multimedia information, which is characterized by its semantic content.

From the system's point of view, an Electronic Road is defined as the user's navigation path (or session usage history) through the multimedia content [3]. More precisely, it is the series of links to the system's Information Units (IU) that the user follows. IUs are the building blocks of the available multimedia content and consist of the actual data (e.g. video segments, images, sound clips or text documents), along with the attached metadata.

While displaying the requested IU, the system also produces a number of dynamic links that point the user to related Information Units, thus proposing extensions to the current Electronic Road, based on the IU semantic nature as well as the user's profile. In most cases, content-browsing systems suggest individual multimedia documents, related to the one that the user is currently viewing; the extension of the recommendation engine's input to a complete Electronic Road enhances the notion of the semantic context being shared between information units.

In this paper, we describe an integrated web-based system that allows users to browse through multimedia information about the history of the eastern Mediterranean Sea basin in an efficient manner. The system utilizes the notion of "*Electronic Roads*" to make content browsing more effective; in addition, it generates and supports interests and browsing profiles for its users, providing them with personalized enriched content retrieval capabilities. The abstract architecture and data models are illustrated in the subsequent section, followed by a detailed description of the profiling mechanism and the processes behind the filtering and recommendation engines. Finally, the particular implementation is portrayed, along with notions of extension towards the MPEG-7 standard.

THE SYSTEM ARCHITECTURE

The general architecture of the proposed system is provided in Figure 1, where all modules are depicted and the flow of information between them is shown. The multimedia information of the system is stored in the *Digital Library* that includes the *Multimedia Store* and the *Thesaurus and Semantic Index* (Figure 1). The Multimedia Store contains the set of all multimedia documents (of several types like image, video, audio, text etc) indexed by distinct units called *Information Units (IUs)*. The *Semantic Index* contains all the metadata information of the above IUs and provides indexing in order to improve the searching process. It is based on the semantic information (thematic category, keywords, etc) provided by the *Content Expert User* (usually at the same time with the insertion of the raw multimedia data files) through the *Content Expert User Interface*. It is organized in the form of *feature vectors* containing the metadata information of each IU and associating the IUs with the set of all *concepts* used for the semantic description. The *Thesaurus* provides two relations: the *Semantic Taxonomy* associating the concepts to each other and the *IU Similarity* associating the IUs to each other.

Important information stored in the system also concerns its end users (simply called users). Traditionally, information retrieval systems model the user as an entity that has a set of *interests*; a portion of the system's IUs match these interests. The user attempts through various methods of interaction with the system to retrieve the documents that match their interests. In the proposed system, the definition of the *User Interests* and the filtering of the multimedia information according to them is a key point.

In order to be able to determine the user interest at any time for any specific user, three kinds of user information may be considered: the *Usage History* impressing all the past transactions of the user, the *Session Information* describing all the transactions of the user made in the current session (the specific Electronic Road) and the *User Profile* containing the user preferences of the specific user. All these are stored in the proposed system as *User Metadata*.

It is rather obvious that real users, i.e. real people, have multiple interests (preferences) [2]; a single vector is not sufficient for the representation of the user preferences. Therefore, users are modeled as having multiple feature vectors describing their preferences. More specifically, they have a single vector describing their negative preferences and multiple vectors for the positive ones. This definition allows for the participation of a single concept in different interests, and to different degrees.

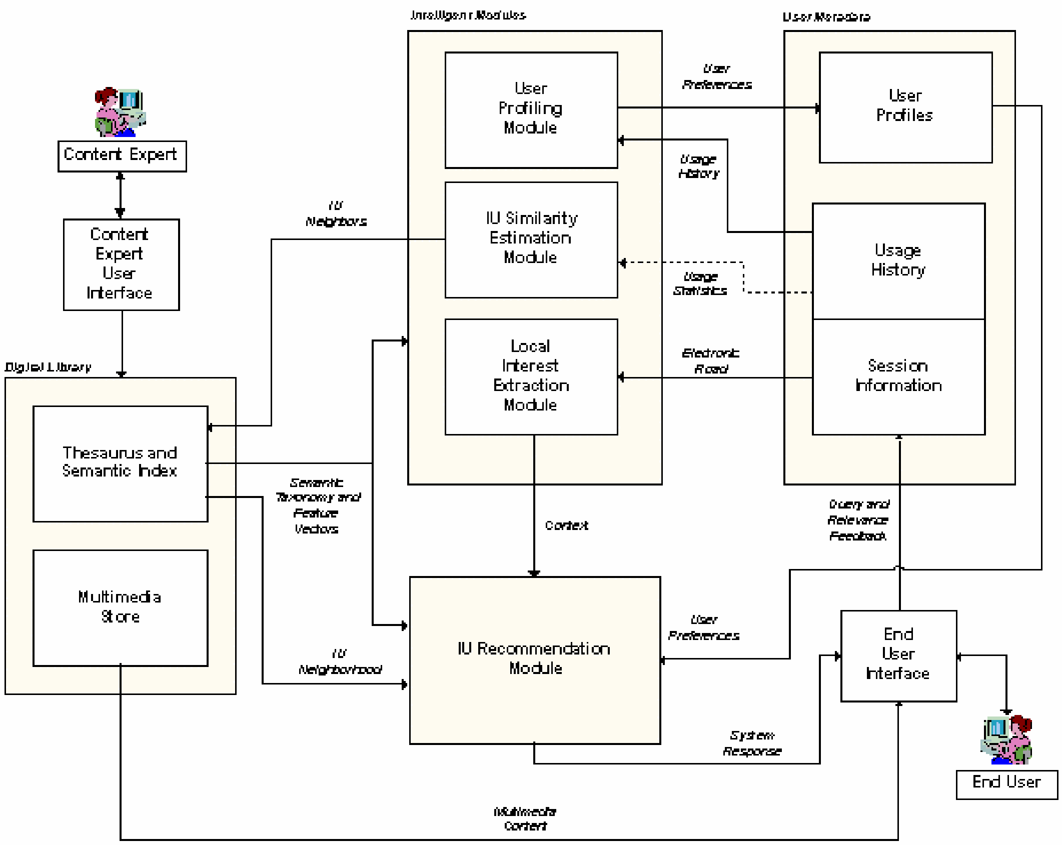

Figure 1. The general architecture of the system

As far as the representation of the usage history is concerned, this is determined to a great extent by the input that is needed by the *User Profiling Module*. In it, the grouping of IUs in Electronic Roads is only used for the running session. Therefore relevance feedback information is divided into two sets of IUs: one referring to the current session, and one containing the remaining IUs. As the user interface does not provide a means for the user to provide an indication of the degree of preference for IUs, the aforementioned sets of IUs are crisp, i.e. they do not support degrees of membership.

Three modules operate into the system in order to provide it with the ability to analyze the user and the IU information. The first one, the *User Profiling Module*, extracts the user preferences out of the usage history and stores it in the user profiles. The second one, the *IU Similarity Estimation Module*, takes the feature vectors (from the semantic index) and assesses the similarity of the IUs. The usage statistics providing the co – existence of IUs in Electronic Roads could also be used to provide more information about IU similarities. The third one, the *Local Interest Extraction Module*, analyses the session information, and more specifically the Electronic Road that the specific user is on, and determines the context of this session.

This context is used by the *IU Recommendation Module*, together with the user preferences and the multimedia indices, in order to provide the user with a personalized suggestion for the continuance of their navigation. These modules are the heart of the system, as they are the ones that allow for intelligent and personalized interaction with the user. They are explained in greater detail in the next section.

THE INTELLIGENT MODULES AND THE IU RECOMMENDATION

The recommendation engine provides users with links to new IUs during their journey through the available multimedia content. The key concept used to form such recommendations is that of the Electronic Road, which traces users’ feedback while forming their current journey and can determine the scope of the user’s current interaction with the system. The feedback provided by the user may, however, contain some degree of uncertainty, resulting in uncertainty regarding the user’s wish. In such cases, we use the user profile, i.e. the user’s known preferences, in order to reduce this uncertainty.

In the following, adapting the terminology of [1], we may refer to the information in the session concerning the user’s current wish as the local interest and to the information in the user profile as the global interests. Extending these, we shall also refer to the portion of the user profile that is related to the user’s current wish as the localized interest.

The proposed system uses as input, for the selection of IUs to suggest to the user, both the Electronic Road and the user profile. Based on these, among all IUs that are related to the one the user is currently viewing, it selects the ones that are most probably related to the user’s wish.

Extraction of the Local Interest: The Electronic Road Context

Each IU may be related to more than one concepts. Therefore, a user may be interested in an IU for more than one reasons. We will refer to the reason that a user selects the IUs that form an Electronic Road as the *context of the Electronic Road*, or context of the session. As the user interacts with the system and forms an Electronic Road, the system can analyze the IUs for which the user has provided positive relevance feedback (by selecting them) and thus estimate the context, as follows..

The Electronic Road metaphor assumes that all the IUs on the path the user is following are related to each other, as well as to the user's wish for other IUs. Therefore, all IUs in an Electronic Road are expected to be related to the same context, i.e. to a common set of concepts. Although they may be related to other concepts as well, in general these concepts will not be common among all these IUs. Based on this notion, the Electronic Road context is determined as the set of concepts that are shared by all IUs in the Electronic Road.

Of course, a user's interaction with the system may gradually alter the user's wish. Therefore, the IUs that appear last in an Electronic Road have a greater probability of being closely related to the user's wish and, as the principles of relevance feedback dictate, are taken into account at a greater extent when extracting the Electronic Road context [5].

The User Profiling Module

Mining the reason a user is interested in a specific IU is an important issue, not only for extracting the user's local interest, but also for detecting the user preferences. Since most users have more than one distinct interests, the IUs viewed by a user in the past cannot generally be related to a unique set of concepts. The user profiling process should be able to automatically group the IUs in a user's usage history and identify the concepts these groups are related to.

To achieve this goal, the proposed system applies a novel context aware hierarchical clustering algorithm. This algorithm, given a set of elements (such as IUs), among which a variety of similarity measures are defined, produces groupings of elements that resemble each other, with respect to one or more of the given measures. Each concept is considered to correspond to a similarity measure; two IUs are considered similar, to the extent that they are both related to this concept. How IUs in a group are similar to each other, i.e. which are the set of concepts that relate them, is referred next as the context of the group and is determined by the proposed algorithm. More on this algorithm can be found in [4].

Modeling user preferences as vectors on the space of concepts facilitates their mining, through the above-mentioned clustering algorithm. Specifically, all the IUs in the usage history for which the user provided positive relevance feedback are supplied to the clustering algorithm as input. For each group of IUs that the algorithm selects as a cluster, the corresponding context is stored in the user profile as a positive preference.

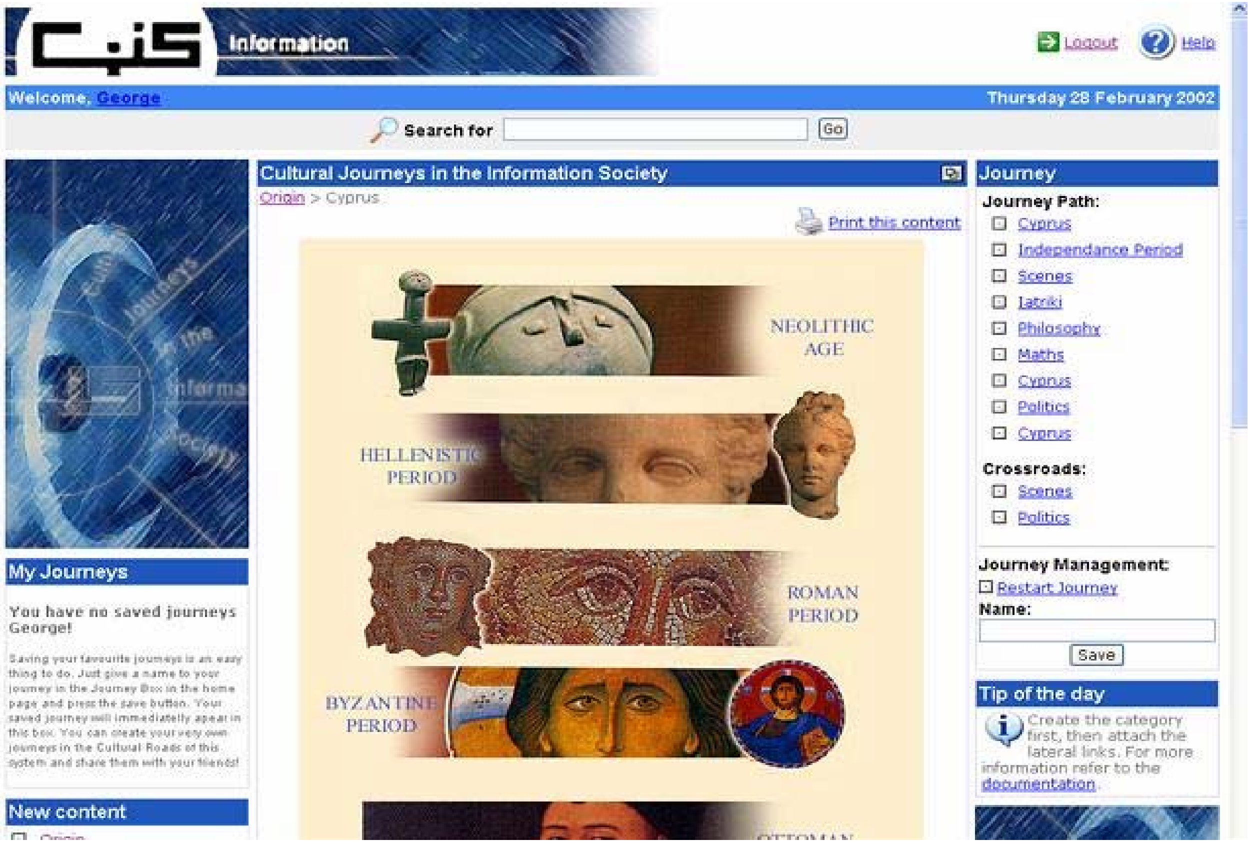

Figure 2. The end user interface

For negative preferences the process is slightly different, as they can all be stored in a single feature vector. IUs for which the user provided at some time negative relevance feedback (and never provided positive feedback) are given as input to the clustering algorithm. The resulting contexts are combined into a single one using a max union.

The IU Similarity Estimation Module

To compute the IUs' to propose to the user in a reasonable amount of time, we can reduce the IU space; this is achieved by considering only the set of IUs that are related to the IU the user is currently viewing. We shall refer to this set as the IU's neighborhood.

Let us first assume that the human expert (i.e., the annotator) has to provide the neighborhood of each IU. Obviously, as the number of IUs increases, this is not a simple task for a human. The proposed system is equipped with a module that is able to automatically estimate IU similarities. In this way, additionally to the neighbors defined by the annotator, the inter – IU relations are also populated based on IU similarities in annotation. Since IUs are represented by feature vectors, the measure of IU similarity is a similarly based on feature vectors.

A widely used practice in the literature is the utilization of clustering techniques for the partitioning of the IUs in clusters; i.e., the neighborhood of each IU is the cluster that contains it. Instead, we propose that each IU is the centre of its own neighborhood. Therefore, instead of pre – detecting clusters of IUs, we simply calculate IU to IU similarities. Once IUs that do not exceed a predefined threshold are ignored, the remaining define, for each IU, its neighbors.

Although the most obvious way to compare two vectors is to assume their normalized *dot product* as a measure of their similarity, this is far from adequate in our case. Two IUs may be considered similar if they are both related to a high degree to a given concept, even if the other categories to which they are related are not common between them. To face this, we select the *max – min product* as a measure of similarity. This assumes high values when the feature vectors indicate high degree of relevance to at least one common concept.

Furthermore, the concepts defined in our system are not independent from each other. This implies that two IUs may be related to each other, even if they are not indexed by the annotator as referring to the same concepts. In order to overcome this problem in the process of calculating the similarity between two IUs, we perform a *semantic expansion* of the feature vectors prior to calculation of inter – IU similarities.

Semantic expansion of feature vectors is based on information concerning the relations among the predefined concepts; this information is stored by the system in the thesaurus. Specifically, the thesaurus represents all concepts as nodes of a weighted ordered graph. A weight approaching one indicates great similarity in meaning, while a weight approaching zero indicates that the two concepts are not highly related semantically.

Utilizing this knowledge, the feature vector of each IU is expanded as to include all concepts that are found in the taxonomy's graph as predecessors of the concepts that have been linked to the IU by the annotator. The degree to which new concepts are related to the IU is analogous to the degree that they are related to their successors and to the degree that these successors are related to the IU.

Applying this module off – line, the proposed system enriches the sets of neighbors defined by human experts, thus generating more relations among IUs.

The IU Recommendation Module

Whenever a user is viewing an IU, they are considered to be inside an Electronic Road. The scope of the proposed system is to effectively propose to the user the IUs that are most suitable for the continuation of this Electronic Road. To make selection of these IUs computationally efficient, only IUs in the current IU's neighborhood are considered. Moreover, the ones that are not related to the context of the Electronic Road or the user profile are ignored.

Since the relation between IUs and concepts is not crisp (taking values that range from zero to one), the context of the session is not a crisp set either. The maximum degree to which any concept participates in this set indicates the degree to which the IUs are related to each other; we refer to this degree as the *intensity of the context* of the session.

When this intensity is great, the IUs of the Electronic Road are closely related to each other. This also indicates that the information contained in the Electronic Road is very specific, and thus helps identify the IUs that the system should propose to the user for the next step of his / her journey. When, on the other hand, this intensity is small, then the session contains a great deal of uncertainty; thus, the user profile needs to be used to a greater extent for the selection of other IUs that may please the user.

It has been mentioned above that neighboring IUs are related to the IU the user is currently viewing. Still, this does not necessarily imply that they are also related to the context of the Electronic Road, or to the user profile. Therefore, these neighbors have to be filtered.

This is not a trivial step, as various open issues are related to it. Some of the most important ones are the following:

1. How to define the exact degree to which the Electronic Road or the user profile has to be used.
2. How to decide which parts of the user profile should be used in this process.
3. How to aggregate the contexts of the user profile with the context of the session.

First of all, the Electronic Road context is always considered to be more important than the user profile, since it is the Electronic Road that describes the reason for the user's current interaction with the system. Therefore, when the intensity of the Electronic Road context is very intense, the user profile is not considered at all. More formally, if k is the intensity of the Electronic Road context, then the Electronic Road is taken into account by a degree equal to k and the profile by a degree equal to $1-k$. Thus, if the session has no context, which will probably correspond to a user who is not browsing for something specific, then the proposed system operates as a filtering engine that is based solely on the user profile. In the opposite case, when $k=1$, the system operates like a query by example engine, where the example is the set of the previously viewed IUs.

The selection of the in – context positive preferences in a user's profile, i.e. the selection of the localized interest, is performed via a direct comparison to the context of the query; since user preferences and contexts are represented in the same way as IUs, this comparison is the same as the one used for the estimation of IU neighborhoods.

Of course, if the intensity of the context of the session is weak, then all positive preferences are considered. In other words, when the context of the session exists, consideration of the user profile may alter it only by a small degree. Negative preferences, as it has already been explained, do not have to be filtered.

The next step is to define the context of the user profile, as far as the specific session is concerned, i.e. the context of the localized interests; the average of all the in – context user preferences is used for this purpose, with negative preferences participating with negative degrees. Finally, the weighted average of this context and the context of the session is calculated; weights are $1-k$ and k respectively. The resulting context will be used for the selection of IUs to be proposed to the user; we will refer to it as the *active interest*.

The active interest describes the concepts an IU should be related to, for being appropriate for continuation of the user's journey. Candidates are documents related to as many as possible of these concepts. Therefore, in the last step, the neighboring IUs are ranked, based on their similarity to the active context. The similarity measure used is the dot product of the feature vectors.

After ranking, the first, e.g. ten, IUs are presented to the user as suggestions. From them, one will eventually be selected, thus becoming part of the Electronic Road and being logged in the positive relevance feedback portion of the user's usage history. The remaining nine are logged in the negative relevance feedback portion of the usage history.

Figure 3. The expert user interface

SYSTEM IMPLEMENTATION

The principles presented herein have been implemented in the CJIS project [6]. This implementation is based on an aggressive object-oriented approach in which IUs are conveniently represented by objects, while taxonomies are represented by graph structures. The traditional Model-View-Controller decomposition is followed and a pure object oriented language, Ruby, is used. Its dynamic and type-less nature, along with its flexible syntax, are particularly important when implementing the object life-cycle manager that is briefly described in the following.

The Model encompasses the raw multimedia data and the metadata that describe them, along with application logic data such as users, preferences, profiles, usage histories, neighborhoods and security constraints. The data that implement the Model are encapsulated in objects and accessed using type-less references. Such references allow the organization of semantically disparate objects in a uniform graph structure, which we call *ObjectGraph*.

The Graph actually integrates multiple overlapping sub-graphs. In other words the objects participate in multiple relation sub-graphs. Some of those sub-graphs represent taxonomies and are unidirectional tree structures, others are bi-directional and represent relations like neighborhood participation. The lifecycle of the objects that participate in the ObjectGraph is managed automatically. A custom developed object manager creates persistent versions of the objects in the underlying DBMS and transparently handles the synchronization of the persistent versions with the live objects.

The Digital Library is the central component of the Model. As described in Section 2, it consists of the Multimedia Store, a low level repository of raw multimedia data, and the Thesaurus and Semantic Index. Our implementation uses an XFS file system on a Linux server being accessed by the administrators and cultural experts using the WebDAV protocol. The Thesaurus and Semantic Index is a high-level organization structure that supports the intelligent algorithms for navigation and recommendation, besides facilitating scoped searching. IU objects are hierarchically organized in multiple sub-graphs attached to the ObjectGraph. The general graph structure used here is capable of modeling the parent-child-siblings relations between the IUs as well as more specific relations, such as hard-lateral links (i.e. links between IU's provided by the content experts to assist the recommendation engine in providing a class of novel suggestions to the user) time period and location cross-links (also provided by content experts to be used by the recommendation algorithms).

The Controller translates interactions with the View into actions to be performed by the Model. In the web-based system, interactions appear as HTTP requests. The actions performed by the Model include querying the indexing structures or changing the state of the Model. Based on the user interactions and the outcome of the Model actions, the Controller responds by selecting an appropriate view. The controller integrates the IU Recommendation Module, the User Profiling Module, the IU Similarity Estimation Module and the Local Interest Extraction Module.

The View renders the contents of the Model. It accesses multimedia data through the model and specifies how that data should be presented. It is the View's responsibility to maintain consistency in its presentation when the indexing structures of the Model change. The View implementation is based on standard web technologies to allow the user to access the system using a common browser. The user interface pages are synthesized using a special purpose XML meta-representation and are transformed to the final HTML documents by applying XSLT stylesheets and localization logic.

The implemented system contained 52 thematic categories under culture, organized in twelve historic periods, each of which was decomposed in the categories "science, architecture, arts, folk arts and history". Each of the latter was decomposed into subcategories at a depth of six levels. The initial Digital Library consisted of approximately 2000 multimedia documents characterized by culture experts from Jordan, Egypt and Cyprus and tested by more than fifty users from these Mediterranean Countries. Figure 2 shows two typical screenshots of the end-user and cultural expert interfaces. The system is currently extended to meet the requirements of the educational system in Cyprus.

The notion of Electronic Roads, being the central notion of this system, is also incorporated in the user interface. Thus, a user is always able to review the last steps in their path, i.e. the last IUs visited, and to return to one of them. In the roads analogy, this might correspond to going back in order to take a different path from a previously encountered road intersection.

Two typical screenshots of the application are presented in the figures. The first one is a screenshot of the end user's graphical interface. In the center one can see the IUs that are proposed to the user, while on the right there are links to the last steps of the current electronic road. The second is a screenshot of the content expert's graphical interface. Through it the expert can insert new IUs, annotate them and relate them to existing information units.

CONCLUSIONS

An integrated system has been presented in this paper, based on an implementation of the Electronic Roads metaphor. This system allows for browsing in related multimedia documents, while mining users' preferences and suggesting similar material to what they are actually viewing. The distributed nature of the system allows for added features that are essential for multi-national and multi-cultural usage, such as localization, distributed content serving and easy localized content updating.

Moreover, the system explores the context of a user's current interaction with it, so as to provide them with semantically related recommendations about the multimedia information. Extension of this exploration in the framework of content-based multimedia analysis, as well as MPEG-7 and MPEG-21 standardization activities is a topic of continuous research and development [7].

ACKNOWLEDGMENTS

This work was partially funded by the EC IST-1999-20502 Project. This work was partially funded by the HERACLETUS fund for basic research.

REFERENCES

1. Barry, C.L, User-Defined Relevance Criteria: An Exploratory Study, *Journal of the American Society for Information Science 45,* (1994) 149-159.
2. Wallace, M., Akrivas, G., Stamou, G. and Kollias, S., "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
3. Wallace, M., Karpouzis, K., Stamou, G. Moschovitis, G. Kollias, S., Schizas, C., "The Electronic Road: Personalised Content Browsing", *IEEE Multimedia* 10(4), pp. 49-59, 2003.
4. Wallace, M., Stamou, G. "Towards a Context Aware Mining of User Interests for Consumption of Multimedia Documents", in *Proceedings of ICME2002* (Lausanne, August 2002).
5. Yong, R., Huang, T.S., Ortega, M. and Mehrotra, S., Relevance feedback: a power tool for interactive content-based image retrieval, in *IEEE Transactions on Circuits and Systems for Video Technology 8* (1998), 644 –655.
6. INCO Project CJIS: Cultural Journeys in the Information Society.
   http://www.cs.ucy.ac.cy/project/cjis/
7. IST Project FAETHON: Unified Intelligent Access to Heterogeneous Audiovisual Content.
   http://www.image.ece.ntua.gr/faethon/