A novel prototype for documentation and retrieval of 3D objects *

Efstathios Onasoglou
Informatics and Telematics Institute
Centre for Research and Technology Hellas
1st km Thermi-Panorama Rd
PO Box 60361, 57001, Greece
onaso (at) iti (dot) gr

Athanasios Mademlis
Information Processing Labaratory
Dep. of Electrical & Computer Engineering
Aristotle University of Thessaloniki
GR 54124 Greece
mademlis (at) iti (dot) gr

Dimitrios Katsikas
Altanet S.A.
7 Nikiforou Ouranou Str.
GR 546 27 Thessaloniki, Greece
d.katsikas (at) altanet (dot) gr

Petros Daras, *Member IEEE*
Informatics and Telematics Institute
Centre for Research and Technology Hellas
1st km Thermi-Panorama Rd
PO Box 60361, 57001, Greece
daras (at) iti (dot) gr

Michael G. Strintzis, *Fellow IEEE*
Information Processing Labaratory
Dep. of Electrical & Computer Engineering
Aristotle University of Thessaloniki
GR 54124 Greece
strintzi (at) eng (dot) auth (dot) gr

Abstract

In this paper a working Multimedia Asset Management system is presented which provides the means for storage, annotation, retrieval and re-use of any multimedia data such as 3D objects, images, sound, video and text. The working prototype is modular, customizable, secure, offers intranet and Internet connectivity and also software inter-operability as it can be compatible with various well-known, established database and ontology standards. The main innovative aspects of the proposed system are the efficient combination of the semantic aspects with multimedia data and the retrieval engine of the system which combines state-of-the-art 3D content-based algorithms and novel semantic driven, relevance feedback methods.

*This work was supported by the AltaB23D project funded by the Greek Secretary of Research and Technology. SUBMITTED TO VMDL07 AND ICIAP07

1. Introduction

Through the last years, with the rapid growth of computer usage, the number of digital assets (documents, images, etc) has been enormously increased. Some companies and organizations have already started adopting Digital Asset Management (DAM) software solutions as a business strategy for secure and efficient storage, search and retrieval of digital assets. The main features of DAM systems are: a) the efficient storage, b) the quick and accurate search and retrieval and c) the reusability of digital assets. Metadata seem to play crucial role in management of non-text-based assets. The DAM systems are relatively new software, however they have been widely accepted in the business community.

Only few DAM systems have been presented so far in the scientific community. In [14], a DAM system is presented, which relies on an XML-based database. The usage of the XML and its native inheritance support allows the system to efficiently store the metadata of an object along with the object and increases the search speed. However, major drawbacks concerning this approach can be identified in the limited search capabilities and the poor performance of the system concerning procedures of updating, inserting and deleting objects. In [16], a media asset management system is proposed, which is oriented to video objects. In this system the MPEG-7 standard is adopted for the metadata definition and description. The metadata context is manually inserted, using a suitable tool, which assists the video annotation, by automatically detecting key changes in a video sequence (e.g. scene transition, zoom, pan etc). The tool has been extended in order to allow collaborative work on the same video. The search module, which is text-based, is based on the metadata context. A very interesting feature of the system is the utilization of a relevance-feedback technique, which provides suggestions to the final user, based on the history of similar searches. Finally, Kallioja [10] presents a DAM approach dedicated to television production companies. This work focuses on the metadata annotation, the retrieval and the exchange of them through a research for effective media asset management practices and a study about the possibilities and requirements for using a specific file format for media exchange.

There are many commercial tools available for DAM, for almost every need ([2, 3, 6, 1, 4, 5]). The main features included in every commercial DAM system are: (a) efficient storage capabilities, where every digital asset is stored along with its metadata into an appropriate database, (b) a search and retrieval module, and (c) a security module, for protecting unauthorized access to the digital assets. The digital assets can be converted to any compatible format. The majority of the DAMs, includes a version control system, in order to track the changes and keep historical data for the digital assets. Most of the DAM systems are general and, thus, are able to handle almost every digital asset. There are also more specialized solutions, limited to assets of specific type ([5]) and others that integrate project management and control tools ([2]).

The rest of this paper is organized as follows: in the following Section the added value of the proposed system is described. In Sections 3,4 the parts of the system are presented and in Section 5 the description of the retrieval engine is given. Finally, conclusions are drawn in Section 6.

2. Innovation of the system

The proposed working prototype harmonically combines classical database concepts with ontology compatible and state-of-the-art content-based retrieval (CBR) methods, while supporting novel relevance feedback (RF) algorithms. Further, there is given the option of exploiting an innovative content-free retrieval algorithm. Content-free retrieval (CFR) algorithms are still in the cradle, however they are expected to be the status quo in information retrieval algorithms. By combining CBR, RF and CFR algorithms the proposed system is able to objectively retrieve geometrically similar 3D objects and also refine the results according to slight variations of each user's subjective notion of similarity. Moreover, apart from the aforementioned techniques (CBR, RF, CFR), the proposed system combines thesauri, text and web search techniques in order to ameliorate irrelevant results in a query and to modify the results according to the needs of each user.

Furthermore, the idea of considering a reposited object as an artifact (a general semantic concept) is introduced and a discriminating line is drawn between two different, cooperating though, annotation modules. The first annotation module is located to the Documentation layer (Figure 1) and corresponds to data that characterize the artifact as a physical, real-world entity. Such data can be variously the name of a statue, the name of its creator, the date of creation, or even bookkeeping information such as the method of 3D reconstruction used to create the statue's 3D model. The second annotation module, the media manipulation module (Figure 1, Core Layer), correlates an artifact with various multimedia data such as 3D representations in various levels-of-detail (LODs), images, audio and possibly text and video. A winning advantage of the proposed system is that information retrieval can be accomplished by means of comparing multimedia data (e.g. what other reposited objects have the same 3D representations?), by means of text/thesauri and Documentation data query, or even better by a combination of them.

The proposed system can handle every known type of multimedia (although CBR and CFR methods are only 3D-oriented) in order to store, annotate and retrieve the data relevant to them. The whole framework is highly customizable in order to suit to the needs of diverse application fields. To the extend of the writers' knowledge, it has not been presented till day another working multimedia asset management system that is so generic, customizable and with added specialized, expert support for 3D data retrieval as the presented framework.

3. Description of the Framework

The media database framework that is proposed in this paper, is a fully-fledged working prototype. It is modular, easily-customized, secure with various notions of security restrictions for various users according to their needs and their rights. The proposed working prototype supports multimedia storage, annotation and retrieval through classical text-based search and thesaurus search but more importantly it provides advanced, state-of-the-art retrieval algorithms for 3D object retrieval, such as content-based, semantic-based methods and relevance feedback algorithms. The working prototype has web and network support, based on a client-server model which allows users to work with the databasethrough a network. That is to say a remote, possibly mobile user can work with the database system with the same ease as a local user. Therefore, the connectivity and inter-operability problems that are present in many storage system are ameliorated. A block diagram of the architecture layers is illustrated in Figure 1. In order to illustrate the usability and working abilities of the proposed application an example of its use is presented in the sequel.

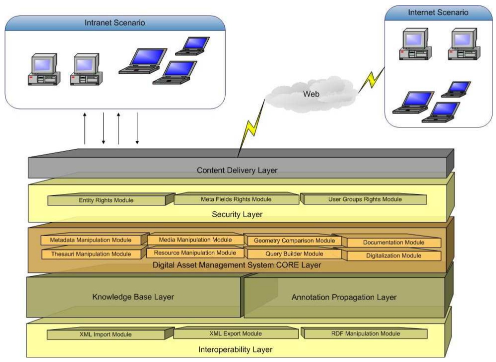

Figure 1. Architecture of the presented working prototype

4. Working description of the System

4.1. Administrative Tasks

Let us assume that a database administrator (dba) has just logged in the system and that it has not undergone any customization till moment. The first task of the dba is to describe what types of multimedia are to be used in the database. This is illustrated in Figure 2(a) where there have been added the fields Photo, Video, Sound, 3D Objects and text (Files). The very next step is to define external players for the media types that have been inserted. In Figure 2(b) five such players have been defined for various types of Photo objects. Also, there is support for language parametrization by relating general database keywords to user-defined translated terms.

Afterwards, the conceptual entities that are to be annotated must be defined. These entities are referred to as artifacts or collections. In Figure 3(a) such artifacts (General Artifacts, Art Works, Books, Files) are depicted. The inserted artifacts are application depended. That is to say, a dba can (and certainly will) insert different artifact collections for different working areas; for example the defined collections in Figure 3(a) can be defined for the needs of a museum or a cultural institute.

Each artifact collection that has been added is just a conceptual description of the qualities that an organization

Figure 2. (a) Adding new media types, (b) Adding new media players

needs to describe, store, retrieve and reuse. These general concepts should be described by various metadata. Figure 3(b) illustrates the entities that contain the metadata, which are associated with the annotation of the exemplary Art Works collection. The dba can select with a tick mark which entities are taking part in the description of each collection. In Figure 3(c) are depicted the fields' definitions for the Acquisition Card of an Art Work. Also, there is an option for each field, whether it can be multilingual or not. If it can, then the annotation will be concurrently present in various languages.

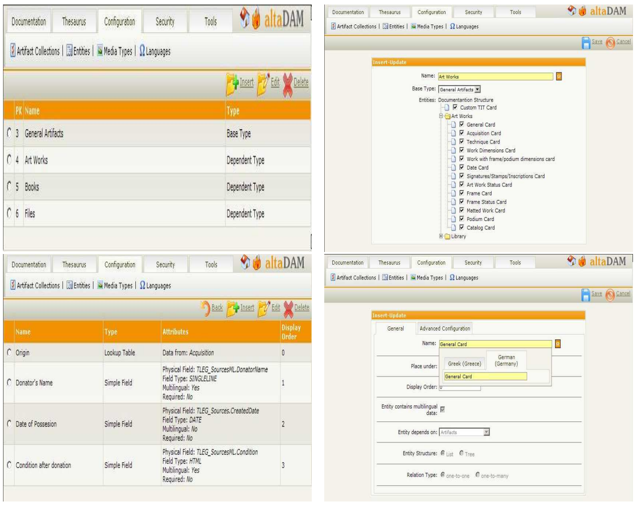

Figure 3. (a) Determining the artifacts, (b) Assigning metadata, (c) Field definitions, (d) Adding new entities

In the proposed system it is possible to create and customize new metadata entities and metadata fields. Figure 3(d) illustrates the capability of inserting a new entity and customizing its fields.

4.2. Annotation

After defining the database, it is time for inserting metadata information and annotating the stored objects in the database. The annotation of an artifact is two-pronged.

Firstly, an object is seen as physical entity that needs to be documented. In Figure 4(a) the documentation of a work art as a physical entity is illustrated. The metadata fields and the entity cards are exactly those that have been added by the dba in Section 4.1. In this example, various issues such as the name of the artist and the art category are concerned. A striking feature of the proposed system is its capability to provide pre-filled fields with suggested values, after it has learned internal associations with annotation propagation algorithms.

The second aspect of annotation is the media annotation. A stored object can be associated with various multimedia types, e.g. an art work can be associated with some 3D representations of the work, photos and possible sounds, videos or texts. Figure 4(b) illustrates the capability of loading and associating a media entity with a stored object. The player of the media file is the one defined by the dba as in Section 4.1.

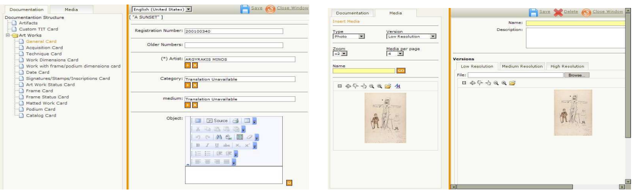

Figure 4. (a) Documentation, (b) Media Annotation

4.3. Querying the System

The vast majority of users will not be occupied with the creation of the entities and the metadata fields, but with querying the system. The search can be either simple, involving minimal description of the desired data, e.g. artifact name and type, or advanced (Figure 5(a)). In the latter case great flexibility is given as multiple queries can been combined with binary operators and certain fields can be set to the desired. Further, a keyword type search is supported (Figure 5(b)), which takes into consideration also the thesaurus features of the proposed system (Section 4.4).

However, the great advantage of the proposed working prototype is its ability to perform 3D content-based multimedia search (Section 5.1). A user can insert a 3D object as a query and the system is able to retrieve all stored objects

Figure 5. (a) Advanced search, (b) Keyword search

that have 3D representations similar to the given one. Furthermore, the retrieval system has refinement capabilities that are based on state-of-the-art relevance feedback methods (Section 5.2). During a refinement cycle the user marks the retrieved results as relevant or not-relevant to her/his query and feeds the system with this data. The relevance feedback algorithm utilizes those data and helps the system to retrieve more relevant objects.

4.4. Thesauri and Web Search

The proposed working system apart from the use of annotation and relevance feedback algorithms, also summons various thesaurus structures in order to capture the semantics of the users' queries. In the current implementation a geographic thesaurus has been created. In Figure 6(a) the capability of creating semantic relations in a thesaurus is illustrated. The user picks a thesaurus term (e.g. Europe in the figure) and correlates it with another term, in a language of her/his choice, by defining the type of correlation (Synonym, Opposite, Homophone, Paronym, etc). There is also the option (under Thesaurus/Thesauri Relations) to correlate thesauri to each other and to define the type of correlation (e.g. Logical, Geographical etc), thus to develop a powerful semantic network of thesauri data. As the most aspects of the proposed system, the thesauri are fully customizable. An (expert) user can construct a thesaurus, define new types of semantic correlations between terms and also new types of relations between thesauri.

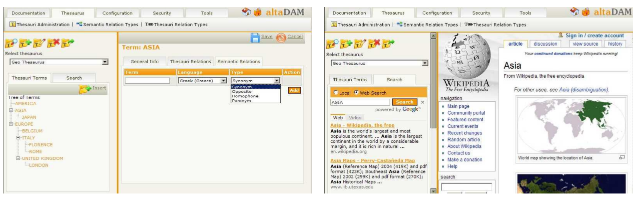

Figure 6. (a) Defining semantic relations, (b) Web-based search

A thesaurus query can be either local or web-based (Figure 6(b)). The search is performed through user-defined aspects of the constructed semantic network and thus provides excellent retrieval performance. The current working prototype has embedded support of searching through the Google [7] search engine, which provides unlimited power in the search capabilities.

4.5. Interoperability

As it has already been denoted, the proposed system is modular, customizable and highly versatile. The intranet and Internet connectivity capabilities allow the system to interoperate through a wide spectrum of client machines that can even be mobile or wireless. Another issue that is dealt with, is the software interoperability. The proposed system provides users with the ability to input a database or ontology XML-like schema and thus to create the entities of the database in a batch. Furthermore, the schema of the entities can be exported in a XML-like format. The exact form of the import/export XML-like file is customizable. Striking advantages of the latter is the compatibility of the proposed system with standards as CIDOC [9] and the re-usability of the exported data by other ontology development tools, since the proposed system can be customized to export RDF [8] or other ontology-based file formats.

4.6. Security

The proposed system puts great effort in the security domain. By emphasizing the security issues the system keeps the availability and quality of service in high percentages and also enhances the usability and the manipulation according to the rights and the responsibilities of different users. The security model the system follows is a Unix-like model that has proven its stability, confidentiality and security power over a wide range of applications. The dba can define security groups and then add privileges to each group. That is to say, the dba can define which metadata, annotation or administrative aspects of the system will be visible to each group. After the definition of the groups and their privileges, the dba can correlate users with groups. Thus an expert annotator in the field of painting will be provided only with the annotation forms, whereas a DBMS administrator will have a view of the whole system.

5. The innermosts of the retrieval system

5.1. Geometric Descriptor Extraction and Comparison Module

The geometric comparison module of the proposed system relies on the comparison of geometric descriptors ex-

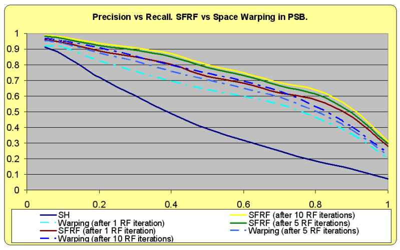

Figure 7. Comparison between SFRF and feature space warping using the PSB.

tracted by utilizing the Spherical Trace Transform (STT) [17]. In the STT framework a set of functionals is applied to the 3D objects producing a descriptor vector which is completely invariant to affine transformations and thus suitable for 3D object matching. Further, weights are assigned to each descriptor, so as to significantly improve the retrieved results. The weights are computed based on the statistical behavior of every descriptor. The geometric dissimilarity metric between two objects is based on a weighted Minkowski L1 distance.

5.2. Relevance Feedback

The most powerful technique in improving retrieval adequacy, apart from devising better low-level (geometric) descriptors, is the use of Relevance Feedback (RF) algorithms. In RF schemes, the user becomes an active part of the retrieval system. S/He supplies the system with a query and some scores for relevant and non-relevant examples. Then, the system refines the search and responses with a new list of objects that contain more similar to the user query, objects.

In the proposed framework, a Semantic Force Relevance Feedback (SFRF) algorithm, originally presented in [15] for 3D object retrieval, is used. In this algorithm the relative scores of various users, which express users subjectivity, are kept accumulatively as additional descriptors for each object. Each object is interpreted as a charged particle, whose relative scores represent the value of the charge. Based on these charges, semantic, newton-like forces are calculated between the 3D objects, which are repelled or attracted properly in the feature space. The forces are of dual nature, semantic and geometric. This combination of low and high-level features in one formula, is one step towards bridging the semantic gap as much as possible till time being. The SFRF algorithm has illustrated excellent results in experiments performed using the Princeton Shape Benchmark (PSB) [17], outperforming geometry-only based algorithms, as well as other state-of-the-art RF systems. In Figure 7, is illustrated the performance of SFRF (continu-

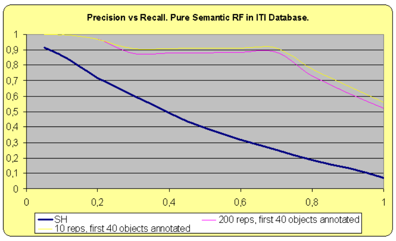

Figure 8. Comparison between CFR and Spherical Harmonics using the ITI database.

ous lines) and Space Warping [13] (dotted lines) for 1, 5 and 10 RF iterations, with 20 objects receiving feedback in each repetition. The line labelled SH is the precision-recall curve for the Spherical Harmonics [12] algorithm. It is obvious that the use of the SFRF algorithm improves significantly the retrieved results.

Furthermore, except for the SFRF method, which utilizes both geometric and semantic data, there is an option of exploiting another novel, purely content-free (CFR) RF method. Only few CFR methods have been developed till day [11, 13]. In the method originally introduced in this paper, the descriptor vector of each database object is constructed from the relevancy votes that have been given by users. For example, let us assume that a user queries the system with $\mathcal{O}_1$ and the system retrieves a list of objects, say 7 objects. Let us, also, assume that the user judges as highly relevant the objects $\mathcal{O}_3$ and $\mathcal{O}_4$ giving them scores "2". As relevant the objects $\mathcal{O}_2$ and $\mathcal{O}_5$ with scores "1" and as non-relevant the objects $\mathcal{O}_6$ and $\mathcal{O}_7$ with scores "-2". Assuming that $N = 7$, the descriptor vector for this query and for one user is: $\mathbf{q}_1 = [x, 1, 2, 2, 1, -2, -2]$, where x denotes the self relevance. In case that another user queries the database with the same object $\mathcal{O}_1$ and his feedback is "2" and "1" for objects $\mathcal{O}_3$ and $\mathcal{O}_4$, respectively, then the final result for the descriptor vector $\mathbf{q}_i$ is the vector sum of the old and the new descriptor vector: $\mathbf{q}_1 + \mathbf{q}_2 = [x, 1, 2, 2, 1, -2, -2] + [x, 0, 2, 1, 0, 0, 0] = [x, 1, 4, 3, 1, -2, -2]$. It is this semantic vector that forms the semantic feature space and defines the position of its object in this space. The objects are manipulated according to further RF iterations by pure semantic forces. In the long run, CFR methods illustrate excellent performance as indicated in Figure 8, where the precision-recall curves for the CFR algorithm are depicted for 10 and 100 RF iterations in the ITI database [17] in comparison with the pure geometric method of Spherical Harmonics [12].

6. Conclusions

The proposed Digital Asset Management working prototype provides modularity, safety and customization ability while it offers excellent retrieval accuracy of the reposited objects as it recruits state-of-the-art CBR methods and novel RF algorithms. The presented generic framework is able to store and retrieve any kind of multimedia data as 3D objects, images, sound, video and text. Furthermore, it can be compatible with various well-known, established database and ontology standards.

References

[1] Artesia. http://www.artesia.com/ html/ artesia_for_dam.html.

[2] Canto and software, cumulus dam, http://www.canto.com/ main_menu/ products/ enterprise/ enterprise.php.

[3] Clearstory, dam solutions, http://www.clearstorysystems.com/ solutions/ digital-asset-management.html.

[4] D-bam team, kbam dam, http://www.d-bam.com/ kbam_platform.asp.

[5] Image server, http://www.braintransfer.de/ en/ products/ image_server.asp.

[6] Interwoven, mediabin, http://www.interwoven.com/ products/ dam/.

[7] www.google.com.

[8] www.w3.org/rdf.

[9] M. Doerr. The cidoc crm, an ontological approach to schema heterogeneity. In *Semantic Interoperability and Integration*, 2005.

[10] M. Kallioja. Media asset management approach to metadat in television production. Master's thesis, February 2006.

[11] T. Kanade and S. Uchihashi. User-powered content-free approach to image retrieval. In *International Symposium on Digital Libraries and Knowledge Communities in Networked Information Society*.

[12] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz. Rotation invariant spherical harmonic representation of 3d shape descriptors. In *Eurographics/ACM SIGGRAPH symposium on Geometry processing*, 2003.

[13] D. Liu and T. Chen. Content-free image retrieval using bayesian product rule. In *International Conference on Multimedia and Expo*, 2006.

[14] S. Natu and J. Mendonca. Digital asset management using a native xml database implementation. In *Proceedings of the 4th conference on Information technology curriculum*, 2003.

[15] E. Onasoglou and P. Daras. Bridging the semantic gap with semantic force relevance feedback. In *International Conference of Image Processing IEEE, 2007*. Under Revision.

[16] M. Tsunakawa, F. Konishi, and T. Nakanishi. Media asset management (mam) system for efficient content management using metadata. In *NTT Technical Review*, 2004.

[17] D. Zarpalas, P. Daras, A. Axenopoulos, D. Tzovaras, and M. G. Strintzis. 3d model search and retrieval using the spherical trace transform. *EURASIP Journal on Advances in Signal Processing*, 2007.