Semantic enrichment for 3D documents: Techniques and open problems

Torsten Ullrich[1]; Volker Settgast[1]; René Berndt[2]

1 Fraunhofer Austria Research GmbH, Visual Computing
Inffeldgasse 16c, A-8010 Graz, Austria,
{torsten.ullrich, volker.settgast}@fraunhofer.at
2 Institut f. ComputerGraphik und WissensVisualisierung,
Technische Universität Graz, Inffeldgasse 16, A-8010 Graz, Austria,
r.berndt@cgv.tugraz.at

Abstract

With increasing knowledge the process of knowledge management and engineering becomes more and more important. Enriching documents by using markup techniques and by supporting semantic annotations is a major technique for knowledge management. This invaluable information is of extreme importance in the context of civil engineering, product life cycle management, virtual archival storage, and preservation. In these fields of applications annotation techniques for 3D documents are a vital part. They provide semantic information that makes up the basis for digital library services: retrieval, indexing, archival, and searching. Furthermore, metadata are of significant importance as they set the stage for data re-use and they provide documentation of data sources and quality, which is vital for every engineering department. Using metadata helps the user to understand data. Additional information allows focusing on key elements of data that help to determine the data's fitness for a particular use and may provide consistency in terminology. In this paper we give an overview on state-of-the-art annotation techniques focussed on 3D data.

Keywords: annotation techniques; semantic enrichment; geometry processing

1. Introduction

In 1998 William J. Clinton announced at the 150th Anniversary of the American Association for the Advancement of Science that "the store of human knowledge doubles every five years". With increasing knowledge the process of knowledge management and engineering becomes more and more important. Enriching documents by using markup techniques and by supporting semantic annotations is a major technique for knowledge management. It allows an expert to establish an interrelationship between a document, its content and its context.

Annotations made by groups or individuals in the context of teamwork or individual work allow to capture contextual information, which can improve and support cooperative knowledge management policies; i.e. annotations can be considered under the perspective of documentation. In fact, tracking the changes and focal points of annotations implies tracing the underlying reasoning process.

This invaluable information is of extreme importance in the context of civil engineering, product life cycle management, virtual archival storage, and preservation. In these fields of applications annotation techniques for 3D documents are a vital part. In this paper we give an overview on state-of-the-art annotation techniques focussed on 3D data.

2. Terms and Definitions

Different documentation standards and annotation processes are used in various fields of applications. Unfortunately, each branch of science has slightly different definitions of bibliographical terms. To clarify these terms and to avoid misunderstandings and misconceptions we present the relevant definitions of terms used in this article.

A document is any object, "preserved or recorded, intended to represent, to reconstruct, or to demonstrate a physical or conceptual phenomenon". This definition has first been verbalized by Suzanne Briet in her manifesto on the nature of Documentation: Qu'est-ce que la documentation? [1]. In Michael K. Buckland's article "What is a document?" various document definitions are given and compared to each other [2]. As we will concentrate on 3D data sets, the "physical or conceptual phenomenon" will always be a three-dimensional phenomenon.

A distinct, separate subpart of a document is called entity. Other authors refer to a subpart as segment. Metadata about documents or parts of documents are defined as "structured, encoded data that describe characteristics of information-bearing entities to aid in the identification, discovery, assessment, and management of the described entities." The American Library Association formalized this definition in its Task Force on Metadata Summary Report [3]. According to this definition metadata is always structured. Unstructured, encoded data, such as comments and free texts, are hereinafter called annotations. As metadata is always structured, it can be specified in a formal, explicit way: an ontology is a "formal, explicit specification of a shared conceptualisation". It provides a shared vocabulary, which can be used to model a domain; i.e. the type of objects and/or concepts that exist, and their properties and relations. Tom Gruber established this definition in his article "A translation approach to portable ontology specifications" [4]. The connections between a document and its metadata or annotations are called markup instructions. They provide local or global reference points in a document.

In the context of computer-aided design, reconstruction, and archival storage, a document is very often the result of a process chain. Data describing a single processing step or a document's process chain is termed paradata.

3. Classification

Metadata and annotations – semantic information in general – can be classified in several ways. Depending on the field of application, they can be classified according to the following criteria.

3.1 Document data type

Semantic information enriches a document. As documents can be grouped according to their type, these categories can be transferred to metadata and annotations as well. We concentrate on 3D data in this article, which can be subdivided further into different kinds of 3D representations.

Boundary Representation. Boundary and surface representations are the most common kind of data representations in computer graphics. These representations comprehend lists of triangles, polygonal meshes, spline surfaces, and subdivision surfaces, etc. to name a few. Several annotation systems allow users to leave text messages on the surface of a 3D model [5] or to draw annotations on the surface and in free space of a virtual scene [6]. The main field of applications is architectural model annotation. Adobe embeds 3D models into PDF documents [7] and combines this technique with its annotation system.

Point Clouds. In a digital documentation process, the input data set is very often a simple point cloud – measured points in space without any additional structure – recorded by e.g. a laser scanner. Although many tools exist (for example: geomagic[1], Leica Cyclone[2]) to annotate and markup point clouds, the automatic documentation of the recording process has many gaps.

Volume Data. Computer tomography and similar acquisition techniques generate volumes of measured points consisting of many layers of high-resolution 2D images. This is the predominant acquisition technique in biomedical sciences, in which documentation, markup and annotation has always been put into practice. Consequently, many established annotation and markup systems exist; for example the volume data annotation tool called VANO by Peng et al. [8] and the Annot3D Project by Balling et al. [9].

Miscellaneous. Besides these "main data types" numerous data representations specialized in its field of application are available. As annotation is a key activity of data analysis, many visualization systems offer annotation capabilities [10], [11].

3.2 Scale of Semantic Information

Semantic information can be added for the entire data set or only for a fragment of the object. For some metadata like "author" it can be sufficient to mark the entire document. But 3D data creation is often a collaborative task with many people working on one complex object. For comments and detailed descriptions a specific place within the 3D data set is needed. Also to communicate suggestions for improvement during an evaluation process it is necessary to make comments on some parts of an object but independent from the entities of the object. This can be done by defining an anchor, like a point, a surface or a volume in addition to the actual data set. It is essential that an anchor can be defined independently from the object. In this way it is also possible to annotate something which is missing in a specific place. Often also the viewer's parameters are stored together with the added information to make it easier to read.

[1] http://www.geomagic.com

[2] http://leica.loyola.com

3.3 Type of Semantic Information

The "Metadata Encoding & Transmission Standard"[3] defines the following types of metadata and annotation:

Descriptive Information. Descriptive information describe the content of an object and comprehend amongst others the Dublin Core metadata set [12].

Administrative Metadata. Administrative metadata provide information regarding how a document was created and stored, as well as intellectual property rights, information regarding the provenance of a document, transformation/conversion information, etc.

Structural Metadata. A structural map and structural links outline a hierarchical structure for a digital library object and embed a document into a context.

3.4 Type of creation

The creation of semantic enrichment of 3D documents fall basically in two categories: manual or automatic. Most of the metadata (especially administrative and descriptive metadata) can be generated automatically, but depending on the domain certain fields need support from an expert. Especially categorizing 3D models can be a difficult task for automatic indexing, e.g. classify buildings according to architectural theory or genres (e.g. the Getty Art & Architecture Thesaurus (AAT)).

Annotations, in terms of free text like comments and remarks, are usually entered manually (the translation of a comment using an automatic translation service would be an example for automatic created annotations). While the manual processing of structured metadata is done by experts, comments or remarks can also come from non-expert users. This method (social tagging) has become very popular within Web 2.0.

3.5 Data organization

The data organization is an important aspect thinking of the sustainability of the annotation. There are two basic concepts how programs can store annotations.

The truth is in the file. The metadata and annotations are stored within the original documents. EXIF or XMP are good examples for that strategy. The main drawback is that the file format must support the possibility to add such arbitrary data. While modern 3D formats like Collada offer this functionality (e.g. the extra tag), others do not. For these a sidecar file can be an appropriate location for storing the annotation. Putting the original document and the sidecar file(s) in a container (portmanteau) is a very popular technique. In

[3] http://www.loc.gov/standards/mets/

most cases the container is a simple zip archive, but with a different and unique extension (examples are the Open Document Format (ODF) or Microsoft Office Open XML).

The truth is in the database. In this concept the metadata and annotations are stored within a database system. This guarantees that the user will access up-to-date data, but he needs to retrieve the associated annotations separately. In addition the application must be able to access the database.

3.6 Information comprehensiveness

Semantic enrichment can be further classified by the comprehensiveness of the information. The amount of comprehensiveness can vary from low to high in any gradation. An example for a low comprehensiveness would be the Dublin Core metadata set [13]. It allows 15 properties to be added as semantic information. In contrast, for example the CIDOC Conceptual Reference Model (CRM) is a scheme with a high comprehensiveness. It is a framework for the definition of relationship networks of semantic information in the context of cultural heritage [14]. CIDOC CRM offers a formal ontology with 90 object classes and 148 properties (in version 5.0.1) to describe all possible kinds of relations between objects.

4. Standards and File Formats

An important aspect of semantic enrichment is to agree on standards. For semantic information it is more a question of organizing documents in a standardized way. In the area of 3D data representations an important aspect is the file format and its ability to support semantic enrichment. Many concepts for encoding semantic information can be applied to 3D data but only a few 3D data formats support semantic markup.

4.1 Semantic Information

The definition and storage of semantic information develops mainly around textual documents. Established standard schemes like Dublin Core and CIDOC CRM (see Section 3) can also be used for 3D content. The Moving Picture Experts Group (MPEG) defined a standard for the annotation of media resources. MPEG-7 is a scheme to describe metadata for audio and visual contents. 3D models however are not part of the scheme. In [15] Bilsco et al. propose an extension to the standard for 3D data.

4.2 Three-Dimensional Content

For 3D data there is not one single standard format like for example JPEG can be seen as standard for digital photos. Many different formats with a large variation of shape descriptions and features are in practical use. For semantic information on 3D data it is even harder to identify one or a few standard formats because most of the commonly used 3D formats do rarely support semantic enrichment.

On the one hand it is possible to extend existing 3D file formats to support metadata and annotations. Especially XML based formats are well suited to be extended while still being readable by existing applications. Some extensions have been proposed for Extensible 3D (X3D) and for Collada. On the other hand there are document formats which have been extended to support 3D content, like PDF 3D.

Collada. The XML-based Collada format was initiated by Sony Entertainment mainly to establish a standard way of data exchange between different creation tools for 3D content. It is hosted by the Khronos group as an open standard. The Collada description allows storing metadata like title, author, revision etc. not only on a global scale but also for parts of the scene like nodes and geometry. Custom extensions to the format are possible as part of the XML scheme. The Collada format can be found in Google Warehouse and the Google Earth application. Metadata like the location of the 3D data on the virtual earth however is stored in separate files.

PDF 3D. Since version 1.6 the Adobe PDF format supports the inclusion of 3D content. For presentations the complete set of data can be exported to PDF 3D. 3D data has to be converted to the Universal 3D (U3D) format or to the Product Representation Compact (PRC) format to be used inside a PDF file. PDF 3D allows to store annotations (logically) separated from the 3D data of the annotated object. Typically the exported 3D data is modified to get a smaller file size. But only if the model is stored without lossy compression it is possible to extract the original data and use the PDF 3D as an exchange format. The PRC format supports product manufacturing information and allows to use the geometry data to be used as input for computer aided manufacturing. U3D is mainly used as a visualization and publication format. Also the PDF format allows adding additional annotations to the model and even annotating the annotations. 3D PDF has the potential to become a standard for 3D models with annotations. The viewer application is widely spread and PDF documents are the quasi standard for textual documents.

The export of 3D content containing metadata or annotations to another file format often leads to information loss. To minimize data loss some semantics, for example labels and measurements, can be integrated into the

3D data as geometry. But afterwards depending on the format it may be hardly possible to distinguish between the metadata and the geometry.

Some software companies offer solutions for their own product lines. For examples Dassault Systems introduced an XML-based format called 3D XML. It is supported by all of their applications as an exchange format. There are many of those solutions but none of them is a standard.

5. Examples

To illustrate semantic information processing, we present results from three on-going research projects: CityFIT, PROBADO and 3D-COFORM[4].

Currently the state-of-the-art for automatically generated city models are basically just extruded ground polygons with roofs. Still missing are detailed 3D models of facades. The goal of the CityFIT project is to reconstruct these facades automatically using the example of Graz, Austria. The main idea of CityFIT [16] is to turn the general implicit architectural knowledge about facades into explicit knowledge, based on both architectural theory and empirical evidence. To achieve this, it combines inductive reasoning with statistical inference. In this example semantic knowledge is encoded in a facade library in form of algorithms. Each execution of such an algorithm generates a (part of) 3D model. During the reconstruction process, the textured point cloud is matched by algorithm instances of the facade library and its instantiation parameters are added as metadata to the point cloud. Therefore, this reconstruction process identifies architectural patterns and enriches the data set semantically.

This metadata is essential in digital libraries. Hence, considering content-based retrieval tasks, multimedia documents are not analyzed and indexed sufficiently. To facilitate content-based retrieval and browsing, it is necessary to introduce recent techniques for multimedia document processing into the workflow of nowadays digital libraries. The PROBADO-framework will integrate different types of content-repositories – each one specialized for a specific multimedia domain – into one seamless system, and will add features available in text-based digital libraries (such as automatic annotation, full-text retrieval, or recommender services) to non-textual documents [17].

[4] http://www.3d-coform.eu/

The third example is a project in the context of cultural heritage (CH). The context of cultural heritage distinguishes itself by model complexity ("masterpiece of human creative genius" [5]), model size (archaeological excavation on the scale of kilometers with a richness of detail on the scale of millimeters), and imperfection (natural wear and tear effects). In this context, the interplay of content and metadata as well as paradata is extremely complex and difficult to model. The aim of the 3D-COFORM project is to establish 3D documentation as an affordable, practical and effective mechanism for long term documentation of tangible cultural heritage. In order to make this happen the consortium is highly conscious that both the state of the art in 3D digitization and the practical aspects of deployment in the sector must be addressed.

6. Recommendations and Open Problems

In this section we discuss some good practices (and also problems) for handling semantic information. Special attention is drawn to aspects which influence the sustainability of the semantic information (e.g. integrity of information, long-term preservation).

Techniques as presented in [18] offer a great flexibility in terms of annotating 3D formats which do not offer a built-in mechanism. One problem of this approach is that modifying the model afterwards can break the integrity of the semantic information. Any 3D authoring operation might change the geometry in a way that the referenced part of the model either no longer exists or has changed its meaning. A possible solution to detect such conflicts is to add a checksum to the annotations. But without deep knowledge of the semantic data structure it is impossible to preserve the semantic information through a processing pipeline. While for special domains a number of complex and expensive product lifecycle management (PLM) solutions exists, other domains, e.g. cultural heritage still lack such a common infrastructure.

For the issue of long-term preservation the use of open standards is a very essential point. This addresses both the 3D file format and the format for the semantic information (if not already included in the 3D format). Especially the next version of Adobe's PDF standard for long-term archiving (PDF/A-2) will play a major role. PDF/A-2 as PDF/A-1[19] will use the "truth is in the file"

---

[5] http://whc.unesco.org/en/criteria

strategy. This guaranties that all information will be available even in an offline scenario. One drawback is that PDF 3D relies on either U3D or PRC. This requires in most cases a conversion step, which inevitably leads to loss of information. Because of the wide distribution of the Adobe Reader, more and more domains start to use PDF 3D as a format for their visualizations. Some examples are described in [20].

Which approach fits best depends on the project constraints. In many cases the 3D file format cannot easily be replaced by another; even worse, most tools still build on their own proprietary formats.

References

[1] S. Briet, Qu'est-ce que la documentation? EDIT -´editions documentaires industrielles et techniques, 1951.

[2] M.K. Buckland, "What is a "document"?," Journal of American Society of Information Science, vol. 48, no. 9, pp. 804–809, 1997.

[3] ALCTS Committee on Cataloging: Description and Access. "Task Force on Metadata -Summary Report," American Library Association, vol. 4, pp. 1–16, 1999.

[4] T.R. Gruber, "A translation approach to portable ontology specifications," Knowledge Acquisition -Special issue: Current issues in knowledge modeling, vol. 5, no. 2, pp. 199–220, 1993.

[5] T. Jung, M.D. Gross, and E. Y.-L. Do, "Annotating and Sketching on 3D Web models," Proceedings of the International Conference on Intelligent User Interfaces, vol. 7, pp. 95–102, 2002.

[6] K. Osman, F. Malric, and S. Shirmohammadi, "A 3D Annotation Interface Using DIVINE Visual Display," Proceeding of the IEEE International Workshop on Haptic Audio Visual Environments and their Applications, vol. 5, pp. 5–9, 2006.

[7] I. Adobe Systems, "3D Annotations Tutorial," Adobe Solutions Network, vol. 7, pp. 1–16, 2005.

[8] H. Peng, F. Long, and E. W. Myers, "VANO: a volume-object image annotation system," Bioinformatics, vol. 25, no. 5, pp. 695–697, 2009.

[9] J. S. Balling, "Visualization, Annotation, and Exploration of Three Dimensional Datasets Using an Existing 3D Rendering System," BioEngineering Senior Design, vol. 1, pp. 1–6, 2004.

[10] M. M. Loughlin and J. F. Hughes, "An annotation system for 3D fluid flow visualization," Proceedings of the IEEE Conference on Visualization, vol. 5, pp. 273 – 279, 94.

[11] L. Offen and D. W. Fellner, "BioBrowser – Visualization of and Access to Macro-Molecular Structures," Mathematics and Visualization – Visualization in Medicine and Life Sciences, vol. 5, pp. 257–273, 2008.

[12] S. Sugimoto, T. Baker, and S. L. Weibel, "Dublin Core: Process and Principles," Lecture Notes in Computer Science – Digital Libraries: People, Knowledge, and Technology, vol. 2555/2010, pp. 25–35, 2002.

[13] "Dublin Core Metadata Initiative." http://dublincore.org/, 1995.

[14] C. C. S. I. Group, Definition of the CIDOC Conceptual Reference Model. ICOM/CIDOC Documentation Standards Group, 2003.

[15] I. M. Bilasco, J. Gensel, M. Villanove-Oliver, and H. Martin, "An MPEG-7 framework enhancing the reuse of 3D models," Proceedings of the International Conference on 3D web technology, vol. 11, pp. 65–74, 2006.

[16] B. Hohmann, U. Krispel, S. Havemann, and D. W. Fellner, "Cityfit: High-Quality Urban Reconstructions by Fitting Shape Grammars to Images and Derived Textured Point Clouds," Proceedings of the ISPRS International Workshop 3D-ARCH, vol. 3, pp. 61–68, 2009.

[17] R. Berndt, H. Krottmaier, S. Havemann, and T. Schreck, "The PROBADO-Framework: Content-Based Queries for non-textual Documents," Proceeding of the Conference on Electronic Publishing (ELPUB), vol. 13, pp. 485–500, 2009.

[18] S. Havemann, V. Settgast, R. Berndt, and Ø. Eide, "The Arrigo Showcase Reloaded – towards a sustainable link between 3D and semantics," Proceedings of the 9th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST), vol. 9, pp. 125–132, 2008.

[19] International Organization for Standardization (ISO), "ISO 19005-1:2005 (Document management – Electronic document file format for long-term preservation)," 2005.

[20] M. Strobl, R. Berndt, V. Settgast, S. Havemann, and D. W. Fellner, "Publishing 3D Content as PDF in Cultural Heritage," Proceedings of the 10th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage (VAST), vol. 6, pp. 117–124, 2009.