PROBADO3D – Towards an automatic multimedia indexing workflow for architectural 3D models

René Berndt[1]*; Ina Blümel*[2]*; Raoul Wessel*[3]

1 Computer Science, Graz University of Technology
Inffeldgasse 16c, A-8010 Graz, Austria
r.berndt@cgv.tugraz.at;
2 German National Library of Science and Technology TIB
D-30167 Hannover, Germany
ina.bluemel@tib.uni-hannover.de
3 Institute of Computer Science II Computer Graphics, University of Bonn
D-53117 Bonn, Germany
wesselr@cs.uni-bonn.de

Abstract

In this paper, we describe a repository for architectural 3D-CAD models which is currently set up at the German National Library of Science and Technology (TIB), Hannover, as part of the larger German PROBADO digital library initiative: The proposed PROBADO-framework is integrating different types of multimedia content-repositories and adding features available in text-based digital libraries. A workflow for automated content-based data analysis and indexing is proposed.

Keywords: Digital libraries, multimedia indexing, content-based retrieval

1. Motivation

The amount of newly generated multimedia content increases year by year and the use of this complex, non-textual data are becoming more and more important. However, this data is not analyzed and indexed sufficiently within the workflow of today's digital libraries, which are focusing on textual documents. Even though a lot of research has been done on how to manage, search, retrieve and present multimedia documents, there is still the need for integrating multimedia documents in existing library workflows. User-friendly tools must be developed so that both the management of multimedia documents for librarians and the user access to these documents (both content-based and in the conventional way of searching metadata) become possible.

Manual indexing multimedia content with keywords often results in a loss of information. To give an example: if a complex 3D model of a roof is just marked with the keyword "roof", it is not identifiable as a flat roof, a pitched roof or a cupola. In current document interpreting processes two different persons have to use the same keywords to make the document detectable: the person interpreting the document and, to detect the document, the person who is searching. Another issue is that authors are not motivated to extend their documents with metadata, even in the presence of suitable tools. Furthermore, it is also nearly impossible to interpret all existing data manually. The result is that in most cases multimedia documents are “black boxes” whose content could not be made accessible individually.

2. Introduction

As a concrete step into this direction, the ongoing cooperative German digital library project PROBADO [1] aims for setting up a framework for integrating repositories containing multimedia documents. The project targets to develop an integrated workflow for both document handling and cataloguing according to the classical library workflow and content-based document processing, i.e., making the collection accessible through content-based retrieval, the latter involving automatic content-based document analysis and indexing.

To achieve these goals, project partners from the University of Bonn and Graz University of Technology, each having expertise in distinct areas of multimedia document analysis and retrieval, are cooperating with partners of two large German libraries, the German National Library of Science and Technology in Hannover (TIB) and the Bavarian State Library in Munich (BSB).

Rather than being a pure research project, it is a special focus of PROBADO to achieve long-term usage of the developed systems and workflows at the cooperating libraries.

Two multimedia repositories are currently set up: one for architectural 3D models at the TIB and one for music at the BSB. As key contributions, we

- describe methods to support automatic processing of general documents in the library processing chain of document acquisition, annotation, search, delivery, and storage,

- develop and implement a common PROBADO platform serving as a web-based access point for searching and accessing general document types stored in the connected repositories. A service-oriented framework allows easy integration of new multimedia type repositories,
- develop and implement PROBADO-enabled multimedia repositories which are located at particular libraries and that are suitable for both conventional and content-based access.

3. Related Work

Within this field there are related scientific initiatives for 3D search engines. There are the Princeton Shape Retrieval [2] group with content-based search engines and Aim@Shape [3] with content-based and metadata based search engines.

And there exists the former EU-project MACE[1], that aims to connect various repositories of architectural knowledge and enrich their contents with metadata. Searching and browsing are very much based on architects needs, e.g. by conceptual connection, geography, language, competence. The search engine is only able to process metadata. There are no 3D models integrated.

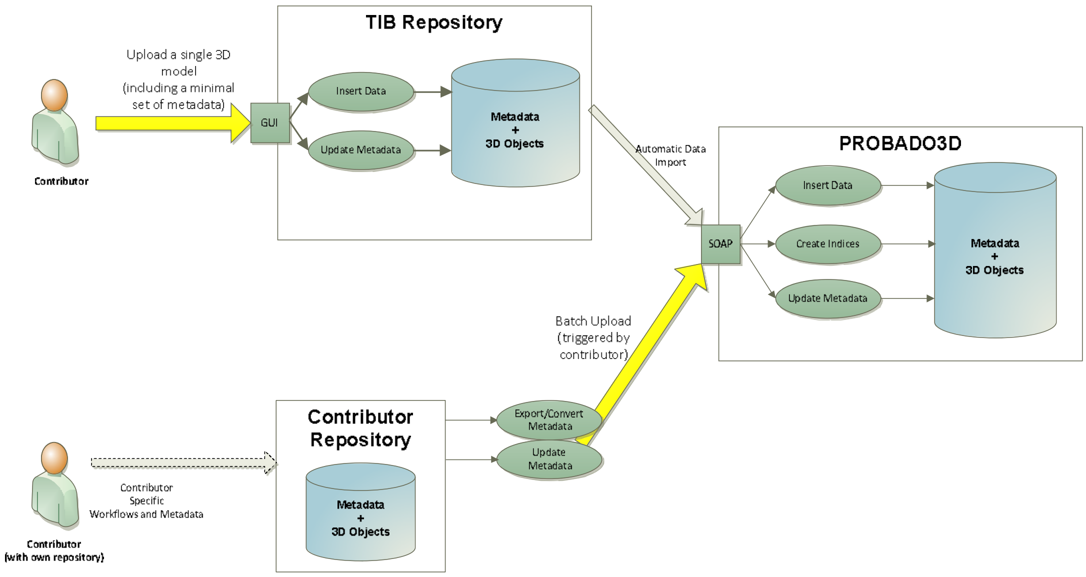

Figure 1. Use-case for integrating 3D models into PROBADO3D

[1] http:// http://portal.mace-project.eu/

4. Use-case: Import of 3D models into PROBADO3D

A repository for architectural 3D models is currently set up at the TIB in Hannover. As being the German National Library of Science and Technology, the TIB references relevant scientific material for superregional supply of technical literature and data.

Momentarily, the repository contains about 7,000 building, construction unit and object models which are converted, indexed and described with metadata as well as approximately 13,000 models that are to be analyzed and indexed.

The source 3D model can be categorized into two major groups:

- Models which are hosted by the contributor itself (Type-A Model) – usually provided by architectural practices
- Models which cannot be accessed through the internet (Type-B Model) – e.g. a submission from a student's master thesis.

Especially Type-A models can be subject to access restriction by the contributor, e.g. pay-per-view or IP-based access for certain groups. Usually the contributor has already implemented the technical details for access or payment.

In the paper we will present the integration workflow (see Figure 1) for models of Type-A and Type-B in detail.

5. Processing Pipeline

One major goal of the PROBADO project is to minimize the manual cataloguing work and to automatically generate the appropriate metadata wherever possible. As a 3D model normally does not bring along any describing data (if it is not catalogued in a database beforehand), the main source for first metadata is the automatic deduction.

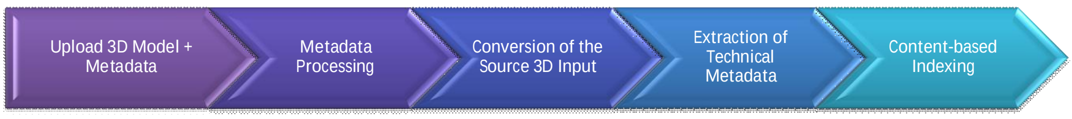

Figure 2. Schematic overview of the PROBAD3D processing pipeline

The following steps are a short excerpt of the different stages of the processing pipeline for the input data (see Figure 2):

Upload 3D Model and Metadata

The first step in the processing pipeline is the upload of the original 3D model and optional metadata by the contributor. This functionality is provided by a SOAP-based web service. The data is submitted as a single archive (see Figure 3), containing one XML file (METADATA.XML) and the 3D model file (plus additional files e.g. materials, etc. belonging to the model).

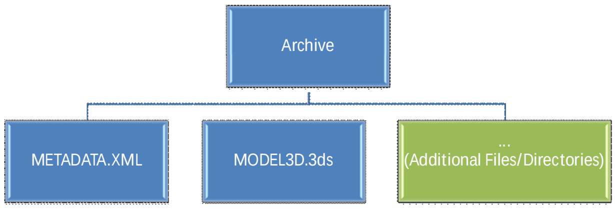

Figure 3. Archive file structure for the SOAP-based web service

The archive is then extracted to the file system for further processing. Beside the upload of new data the web service also supports modifying the 3D model and/or metadata of already uploaded data.

Metadata Processing

The metadata provided by the contributor (METADATA.XML) are stored in the metadata database of PROBADO3D. Since the PROBADO project concentrates on content-based indexing, the only required metadata is information about the contributor. An excerpt of the optional metadata is listed in Table 1.

Conversion of the Source 3D Input

For easier indexing, searching and viewing, a copy of each model is normalized and automatically converted into a uniform format for indexing and different formats for preview and delivery.

Due to the number of input formats it is not feasible for the content-based indexer to implement support for all these formats. The OBJ format from Alias Wavefront was chosen as the uniform format for indexing. For previewing, the PDF format from Adobe was selected, because of the wide distribution of the Adobe Reader.

Table 1. Optional information included in the contributors' metadata

Name	Description
Event	Contains data about events (e.g. competition, seminar, presentation)
Title	The title of the 3D model (not the filename)

Description	A textual description of the 3D model
Subject	Keywords or classification of the 3D model
Location	A geographic reference to a real building
Object	A reference to a real building
ExternalInfo	Contains additional information about external providers, events
Relation	To define relations between models

Most of the tools used for this task during the design phase supported only semi-automatic conversion or had license restriction in terms of usage as a service. In order to provide a fully automatic conversion for the workflow, DeepServer from Right Hemisphere was chosen. The following requirements for the conversion module were verified:

- Automatic conversion By using a watched folder DeepServer can execute a custom defined workflow on adding files to this folder. For the evaluation the input formats were converted to OBJ, PDF, and PNG (thumbnail).
- Support of all input formats DeepServer supports a large number of input formats. Even proprietary formats like the MAX format can be used if the appropriate software is installed.
- Licence The licence of DeepServer allows usage and deployment as a service.

Extraction of Technical Metadata

Technical metadata of the 3D model are automatically extracted during the conversion process. These include number of vertices, polygons, textures, etc. which can provide an approximate estimate about the model complexity. This task is also performed by DeepServer (see Figure 4); the resulting technical metadata are stored in the metadata database of PROBADO3D.

<keyframeanimated>False</keyframeanimated>	<TextureLinks>True</TextureLinks>
<NumObjects>39</NumObjects>	<UVMapped>True</UVMapped>
<NumPolygon>17819</NumPolygon>	<XMin>-434.3582</XMin>
<NumTextures>2</NumTextures>	<XMax>3345.902</XMax>
<SoftbodyAnimated>False</SoftbodyAnimated>	<YMin>-724.8393</YMin>
<AnimationLength>0</AnimationLength>	<YMax>2269.562</YMax>
<NumVertices>10171</NumVertices>	<ZMin>-0.04620994</ZMin>
	<ZMax>722.0472</ZMax>

Figure 4. Technical metadata of a 3D model extracted by DeepServer

Content-based Indexing

Indexing of architectural 3D models is a prerequisite for content-based query by example and document browsing. By creating a concise object description, the similarity between two 3D models can be computed. In the last years, automatic content-based indexing research resulted in the detection of many indexing characteristics especially in the lower layers of semantic, i.e. characteristics relying on rather pure geometrical shape content. For a detailed overview of these algorithms we refer to [8]. We compute global and local low-level semantic features for characterizing architectural components. In a query-by-example scenario where the user either uploads an existing 3D component or uses a sketch-based PROBADO3D interface to generate one, the search for similar objects in the database is conducted using global shape features based on spin-images [9] which are easy to compute and guarantee fast response times of the query engine. For browsing the PROBADO3D repository based on shape similarity, we comprehensively characterize the components using high quality local shape descriptors and special distance measures tailored to the requirements of architectural 3D models [6].

While low-level semantic features are an effective means to characterize the geometric shape of an object, they are not well-suited to describe the structure of building models which is mainly defined by the topology of rooms and floors. To overcome this drawback, we introduced the concept of Room Connectivity Graphs [4]. These graphs are especially designed to capture the topology of buildings. Rooms are represented by vertices, and connections between rooms like doors, windows, stairs, etc. are represented by edges. The graph is additionally enriched by semantic attributes like the dimension of rooms or the type of the connection. By that, users can search for building models that contain a certain spatial configuration of rooms. For the definition of such configurations we provide a graphical user interface. Additionally, the search can be further constrained regarding e.g. the area of certain rooms.

Content-based indexing of components using global and local low-level features works completely automatic and does not require manual data processing. The extraction of Room Connectivity Graphs however requires the building models to be oriented in a consistent way, i.e. the object's positive Z-axis must point towards the virtual sky. Additionally, the scale of the object must be known. In our experience so far it shows that this requirement does not lead to an increased amount of manual preprocessing interaction. In many cases, the scale is contained as metadata in the underlying 3D model data. Additionally, when presented a new charge of models from a content-provider, orientation and scale is usually consistent within one batch and therefore only requires minimal manual interaction.

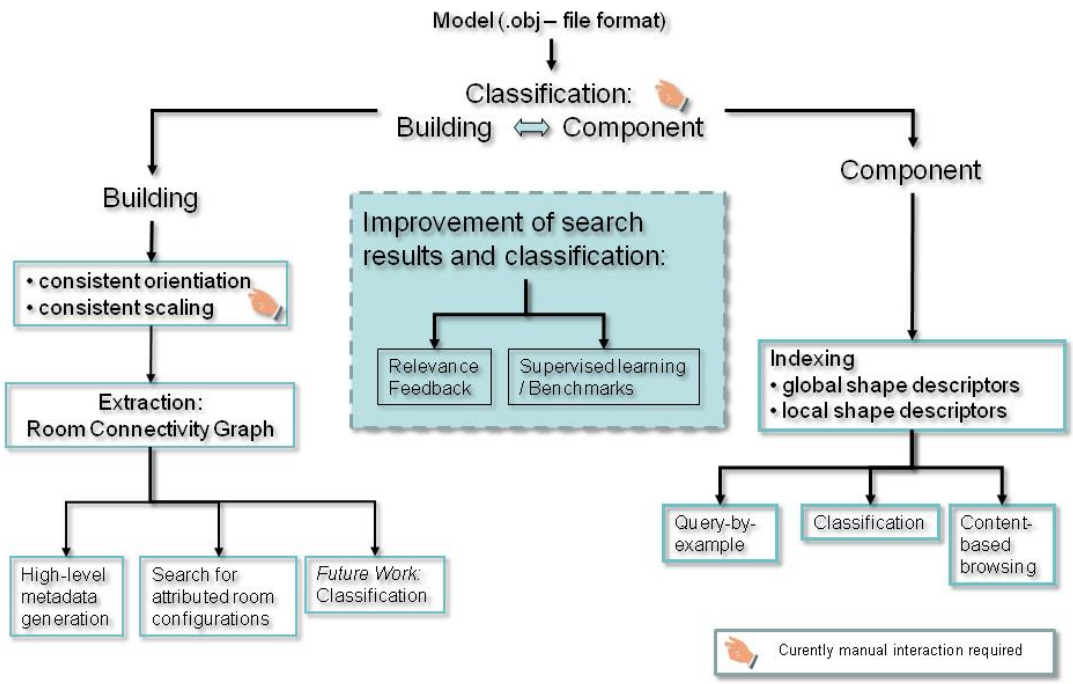

Figure 5. Content-based indexing workflow

Generation of High-Level Metadata

The afore described indexing techniques involving local shape descriptors and Room Connectivity Graphs serve as a starting point for fully automatic generation of high-level 3D object metadata. For predicting the object category of architectural components, we developed a supervised learning framework [6]][5][7] that classifies components according to their associated local shape descriptors. To this end it incorporates shape knowledge about a large number of manually classified architectural components contained in the Architecture Shape Benchmark [7]. The shape classification in this benchmark was created according to common architectural shape taxonomies.

The extracted Room Connectivity Graphs provide a large amount of information about building models that is important to architects. For example, we automatically extract the number of building floors, room areas, gross floor area, window areas per room and per floor, number of rooms per floor etc.

The resulting high-level metadata is finally stored in the PROBADO3D metadata database, allowing the user to textually search for components belonging to certain object categories as well as to search for building models fulfilling certain spatial specifications, e.g. concerning the number of floors or the gross floor area.

6. Conclusion and Future Work

Most parts of the processing pipeline have already been implemented. The integration of DeepServer is planned for the next project phase. The proposed workflow will then be established as a service at the German National Library of Science and Technology.

Future work will concentrate on improving the automatic classification and processing of the 3D models. Similar to the above described categorization of components, we will additionally examine how building models can be automatically classified according to their Room Connectivity Graph.

One additional content-based indexer using semantic enrichment methods based on procedural shape representations [10] is currently implemented and integrated into the PROBADO3D system. By fitting a procedural description to the target model the semantic information carried with the generative description can then also be applied to the target model (e.g. number of columns, stairs, etc).

Acknowledgments

PROBADO is a joint research project supported by the German Research Foundation DFG under the LIS program. PROBADO started in February 2006 with a tentative duration of five years. Partners are the University of Bonn, Technische Universitaet Darmstadt, Graz University of Technology, the German National Library of Science and Technology in Hannover, and the Bavarian State Library in Munich.

The work presented in this paper was partially supported under grants INST 9055/1-1, 1647/14-1, and 3299/1-1. For further information, please visit the project website at http://www.probado.de/.

References

[1] BERNDT, R.; KROTTMAIER, H.; HAVEMANN, S. and SCHRECK, T. *The PROBADO-Framework: Content-based Queries for Non-Textual Documents*. In Proceedings 12th International Conference on Electronic Publishing (ELPUB 2009), 2009,pp. 485-500.

[2] SHILANE, P.; MIN, P.; KAZHDAN, M.; FUNKHOUSER, T. *The Princeton Shape Benchmark*. In Shape Modeling International (June 2004).

[3] FALCIDIENO, B. *Aim@Shape Project Presentation.* In Proceedings of Int. Conference on Shape Modeling and Applications, 2004, pp. 329.

[4] WESSEL, R.; BLÜMEL, I.; KLEIN, R. *The Room Connectivity Graph: Shape Retrieval in the Architectural Domain*. In Proceedings of The 16-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision'2008, UNION Agency-Science Press, February 2008.

[5] WESSEL, R.; KLEIN, R. *Learning the Compositional Structure of Man-Made Objects for 3D Shape Retrieval*. To appear in Proceedings of 3rd EUROGRAPHICS Workshop on 3D Object Retrieval, May 2010.

[6] WESSEL, R.; BARANOWSKI, R.; KLEIN, R. *Learning Distinctive Local Object Characteristics for 3D Shape Retrieval* In Proceedings of Vision, Modeling, and Visualization 2008 (VMV 2008), pp. 167-178, Akademische Verlagsgesellschaft Aka GmbH, Heidelberg, October 2008.

[7] WESSEL, R.; BLÜMEL, I; KLEIN, R. *A 3D Shape Benchmark for Retrieval and Automatic Classification of Architectural Data* In Proceedings of Eurographics 2009 Workshop on 3D Object Retrieval, pp. 53-56, March 2009.

[8] TANGELDER, J. W. H.; VELTKAMP, R.C. *A survey of content based 3D shape retrieval methods.* Multimedia Tools and Applications, volume 39, 2008, pp. 441-471.

[9] JOHNSON, A. *Spin-Images: A Representation for 3-D Surface Matching.* PhD thesis, Robotics Institute, Carnegie Mellon University, August, 1997.

[10] ULLRICH, T.; SETTGAST, V. and FELLNER, D.W. *Semantic Fitting and Reconstruction*. In: Journal on Computing and Cultural Heritage 2008(1), pp. 1201-1220.