Foundations of 3D Digital Libraries: Current Approaches and Urgent Research Challenges

Benjamin Bustos[*]
University of Chile

Dieter W. Fellner[†]
Technische Universitaet Darmstadt and Fraunhofer Computer Graphics Institute, Germany

Sven Havemann[‡]
Graz University of Technology, Austria

Daniel A. Keim[§]
University of Konstanz, Germany

Dietmar Saupe[¶]
University of Konstanz, Germany

Tobias Schreck[‖]
Technische Universitaet Darmstadt, Germany

ABSTRACT

3D documents are an indispensable data type in many important application domains such as Computer Aided Design, Simulation and Visualization, and Cultural Heritage, to name a few. The 3D document type can represent arbitrarily complex information by composing geometrical, topological, structural, or material properties, among others. It often is integrated with meta data and annotation by the various application systems that produce, process, or consume 3D documents.

We argue that due to the inherent complexity of the 3D data type in conjunction with and imminent pervasive usage and explosion of available content, there is pressing need to address key problems of the 3D data type. These problems need to be tackled before the 3D data type can be fully supported by Digital Library technology in the sense of a *generalized document*, unlocking its full potential. If the problems are addressed appropriately, the expected benefits are manifold and may lead to radically improved production, processing, and consumption of 3D content.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval; I.3.5 [**Computer Graphics**]: Computational Geometry and Object Modeling—*Curve, surface, solid, and object representations*

[*]bebustos@dcc.uchile.cl
[†]d.fellner@igd.fraunhofer.de
[‡]s.haveman@cgv.tugraz.at
[§]keim@dbvis.inf.uni-konstanz.de
[¶]dietmar.saupe@uni-konstanz.de
[‖]tschreck@gris.informatik.tu-darmstadt.de

First International Workshop on Digital Libraries Foundations In conjunction with ACM IEEE Joint Conference on Digital Libraries (JCDL 2007), Vancouver, British Columbia, Canada, June 23, 2007

Keywords

Digital libraries, generalized 3D documents, 3D representation, content-based retrieval.

1. INTRODUCTION

The rapid evolution of information and communication technology has always been a source for challenging new research questions in computer science. What happens regularly is that a new generation of technology makes it suddenly possible to process, store, and/or transmit much larger amounts of information. Thus, a gradual *quantitative* increase can turn into a sudden *qualitative* leap, simply because things become possible that were not possible before. The nightmare from a computer science point of view is the *data grave*: Information that is physically present is "lost" for usage because it is simply not accessible with reasonable user efforts.

Digital Library technology aims to revert the data grave problem into a situation where the stored content is brought to its full productive potential by solving the storage, organization, and content-based access problems. For textual documents, retrieval services attacking the data grave problem are widely available, e.g., in form of desktop search engines. But what is the analogue of full text search in a repository of 3D content?

Content-based Digital Library support for 3D data is highly desirable, as the sources for producing digital 3D content are gaining momentum. We argue that the next major technological revolution will be triggered by massive 3D data sets that we will be generated in the near future. The *modeling bottleneck*, the fact that the creation of digital 3D objects was long much too expensive, is overcome by new technologies. Sources generating massive amounts of 3D data will be *3D scanning* (using scanner devices, see Figure 1 for an example), *photogrammetry* (reconstructing 3D data from 2D images), and *procedural/parametric shape design* (creating new shapes from existing similar, parameterized shapes).

Not only is it easier to *produce* digital shape, also the possibilities to *utilize* and take benefit of the created 3D data sets are increasing. A large shift towards 3D is obvious. On the *PC desktop*, computer games have helped 3D graphics hardware becoming the standard, and after Apple's

Figure 1: Creation and markup of a 3D model from range-map input. Left: Original input data, 3 out of 20 range maps taken from a statue are shown un-textured. Top right: Simplified versions of the range maps, textured and un-textured. The gravestone and the statue's cheek were manually segmented for semantic markup. Bottom right: Several range maps were integrated and smoothed.

pioneering work in MacOS X, with Microsoft Vista, 3D will also be integral part of the Windows desktop. Instead of being optional, 3D on the desktop will actually become a standard. In industry, the prospect of *mass customization* is a driving force behind the digitalization of the whole production chain, relying heavily on 3D models to represent process information.

We argue that in the near future, we will be confronted with massive amounts of 3D content, and that novel 3D Digital Library support will be crucial in making the best possible use of these data amounts. In Sections 2 and 3, we outline the current state of the art in Digital Library support for 3D documents, and identify critical research problems. We argue that if these problems are addressed appropriately, a significant leap ahead in the effective use of massive 3D content will be possible. The efforts required to this end are expected to pay off, as illustrated by potential future 3D applications envisioned in Section 4. Section 5 concludes the paper.

2. 3D DATA AND ITS REPRESENTATION

The 3D data type is a very powerful means of capturing and communicating information. Due to the nature of the data type and complexities involved in acquisition, production and processing of 3D data, a number of serious problems in 3D data representation, encoding, content markup, and data history management exist. To date, these problems have not been sufficiently solved, and they are a major obstacle to a full integration of the 3D data type into Digital Libraries. In this Section, we discuss some of the most important research questions of 3D data management in Digital Libraries, according to our view.

2.1 Understanding 3D shape representations

A fundamental difference between 3D and other media types is that there is no canonical 3D representation. While e.g., the image data type can be seen as a set of color samples organized on a regular grid, representing 3D data is more complicated. Existing approaches can be roughly divided into *surface-* and *volume-based* representations, which in turn can be given in discrete, parametric, or implicit form. E.g., 3D objects can be specified by a set of parameterized surface patches based on splines, or a grid of voxels (a discrete volumetric representation). On top of these two broad categories, *structural* information can represent the relationship between models parts in form of scene graphs or boolean set operations which are highly useful for certain shape modeling or manipulation tasks.

These shape representations are *not* all equivalent, because they differ in their expressiveness (the types of forms they can encode) and consequently, in their semantics (content). E.g., a closed surface bounds a volume, but a volumetric data set contains many surfaces at the same time (e.g., iso-surfaces). They also differ regarding the way we can process and analyze them. Discrete representations relate to sampling theory and may exhibit aliasing effects, while continuous representations are noiseless and usually better suited for analytic processing.

An encompassing 3D shape representation taxonomy covering *all* known 3D representations is needed, to better understand the relationships between the existing representations. This should allow a better tackling of the difficult problem to analyze and relate the content of 3D models, irrespective of the given representation, to support common Digital Library tasks such as organizing objects by similarity, deducing hierarchical catalogue orderings, etc.

2.2 Generic 3D file format

Unfortunately, to date there is no single commonly accepted, comprehensive 3D file format, but application-dependent, proprietary file formats are prevailing. In practice, it is usually impossible to convert losslessly between the different established file formats, which is a fundamental problem for importing content from heterogeneous sources into 3D Digital Libraries. In the CAD domain, where almost exclusively NURBS-based model representations are employed, several long-standing, mature exchange standards such as STEP and IGES exist. The problem with these formats is that over time they have become extremely comprehensive and elaborate, so that implementing converter programs for these formats constitutes a challenge on its own.

From the research viewpoint, the file format problem has been completely ignored up to now, although it is apparently a significant problem which requires fundamental efforts. The focus in 3D modeling research up to now has been to further extend the set of shape representations, rather than to work on a powerful yet transparent canonical 3D file format. The existence of such a format would not only allow the easy integration of 3D content from heterogeneous sources, but could also support the adoption of advanced 3D representations by real-world 3D applications.

2.3 Stable 3D markups

Another important concern from the Digital Library perspective refers to stable markup methods for 3D content. Given a 3D model or scene, reliable methods are needed which allow the stable identification (markup) of parts of the 3D content for attaching annotations, hyperlinks, cross-references, etc. 3D content is often preprocessed or edited along the 3D application pipeline, which usually significantly affects the 3D content representation. E.g., consider a lossy compression performed on a 3D mesh model prior to its transmission over a network. Mesh simplification (decimation) methods affect the number, position, and connectivity of mesh vertices. Any 3D markup method based directly on the mesh index, and which is not explicitly known to the mesh compressor, must then be considered unstable.

So, the research problem to be addressed is to define *generic, stable 3D markup* methods, by designing methods to robustly reference portions of a 3D model. The markup methods should be independent of the 3D content representation, and robust with respect to certain shape editing and processing operations which might be needed by the applications.

A solution to the problem of updating shape markups during shape processing operations is that the processing algorithm is (a) aware of the markup and (b) keeps track of appropriately defined geometric primitives that are affected by the processing operation. Then, after the processing has taken place, these primitives can be converted back to a markup of the initial type. The crucial point here is that the shape representation must be able to enumerate shape components in the reference. Such shape component enumeration can be regarded as *spatial queries* and take the form of closeness to a point, containment in a frustum, or ray intersection. Identification of an efficient set of shape queries which allow implementation of robust 3D markup remains an important research challenge. Figure 1 (right) illustrates a 3D scene with markups.

2.4 Data origin and processing history

During 3D acquisition, production, and processing, the 3D content is often composed from different, heterogeneous sources, and manually or automatically processed by different users and applications. If in a given 3D model, some local shape detail becomes of specific interest to an analyst, it is a vital feature that it is possible to trace back the origin of the specific detail, its degree of authenticity, and the kind of processing applied on it. To this end (a) suitable standards for describing the *provenance* of 3D content, and (b) a general scheme for capturing the *data processing history* applied on the content needs to be developed. Regarding (b), ideally, the captured information should allow a complete replay of the processing the data has undergone.

The enormous complexity of this problem may not be apparent immediately. First, we have to cope with two levels of heterogeneity, namely, the various shape representations, and the various processing operations possible on these representations. Both are not canonical along the different 3D creation and processing tools available. Second, capturing model editing operations must take place at the right granularity. Practically, it is neither possible not useful to capture each manual editing step individually, but an appropriate level of aggregation has to be chosen. Third, to actually replay the processing operations is extremely difficult: It requires that all tools used in the 3D production pipeline support the processing history and add to it. Practical experience regarding software versions, operating systems, undocumented ad-hoc scripting by users etc. suggests this is a tremendous task. More subtle problems in this context are reported in [10].

3. ORGANIZING AND SEARCHING

The previous Section discussed urgent research problems relating to the representation, storage, and processing of 3D content. Assume those problems were already solved. Technically, it would then be easy to build large repositories of 3D content from heterogeneous sources using crawlers, converters, and storage systems. The second major challenge is then to provide effective *content-based organizing* and *searching* functionality for making use of the resulting large 3D repositories.

One way for organizing and searching 3D repositories is to make use of mark-up, authoring, and editing information, or other meta data associated with the models. Unfortunately, the availability and comparability of such information cannot be assumed for content integrated from heterogeneous sources. Instead, *analysis algorithms* are needed to automatically generate suitable meta data information from the repository. The output of the content-based automatic analysis can then be used for organization and retrieval of the content, as a replacement for or addition to object meta data, as far as such is available. We next discuss key challenges in content-based 3D organization and retrieval.

3.1 Need for a dictionary of 3D features

A fundamental library service is *content organization* in the sense of giving structure which helps the user navigate the repository, and to formulate queries which allow to retrieve content of interest. This structuring needs to be based on attributes or features of the data itself. E.g., in case of text documents, features such as title, author, or the main topics addressed by the text are candidates for structuring of text collections. For the 3D data type, attributes such as author or producer might be consistently specified in a generic format. But what constitutes the actual *content* in a 3D data set, and how can appropriate descriptors be automatically generated from the models?

Conceptually, a suitable content description is expected to be determined by the application domain the content is used in. E.g., for a 3D CAD model, the features relevant to a given 3D model can be expected to depend on the engineering context associated with that model. It can be assumed that a set of model features which are relevant in a given engineering context are not necessarily useful to organize, say, a repository of models representing historic buildings, as both object types are made use of context of their own conceptual background.

What is missing is an encompassing definition of features (aspects, properties) which are relevant to organize and distinguish collections of arbitrary 3D content. A general *taxonomy of 3D features* in the sense of a 3D dictionary needs to be defined, where the dictionary entries

(1) allow to meaningfully describe any type of 3D content,

(2) are descriptive and discriminating in nature, and

(3) can be robustly extracted (detected) by appropriate automatic analysis algorithms.

Unfortunately, to date no such taxonomy exists, so it is problematic to speculate whether and which 3D analysis algorithms would be capable to robustly detect such features, or how such algorithms should be designed. The problem of defining a 3D feature dictionary is complicated by the fact that it is not clear (a) on which conceptual level the features should be defined, i.e., on the statistical, syntactical, or semantical level, and (b) how the features will relate to the 3D shape representation problem, e.g., if they should be defined based on surfaces, on volumes, or on structural properties. The next Section relies on 3D features to introduce a model of the 3D similarity space useful for designing 3D retrieval systems.

3.2 A model for the 3D similarity space

A most fundamental task in Digital Libraries refers to searching for similar content: The user issues a query to the system, and receives a sorted list of answers. A popular searching paradigm is *query-by-example*, where an exemplary object is provided, and the system returns the most similar elements from the repository. However, the notion of similarity per se is under specified. Like for other data types, for 3D objects many different similarity notions are possible, and the Digital 3D Library should offer support for searching along each of those notions. We propose to organize the space of 3D similarity notions along the three dimensions *similarity type*, *addressed feature*, and *invariance properties*. In the following, we discuss each of these dimensions.

Similarity type

Global similarity considers the similarity between complete 3D object instances, and is used to retrieve whole objects. **Partial (local) similarity** on the other hand bases similarity relationships on correspondences of object parts, not necessarily the objects as a whole. This notion is useful e.g., for retrieving scene models, where similarity may be given by correspondence of individual objects in the scene, not necessarily at the same positions. To his end, the partial similarity makes use of the global similarity notion, applied on individual scene elements. A third type of similarity relates to **functional correspondences**, and can be globally or locally defined. Here, similarity relationships are established between objects or object parts based on application-dependent, functional correspondences. E.g., in a CAD context, complementarity between machining parts could establish a functional correspondence.

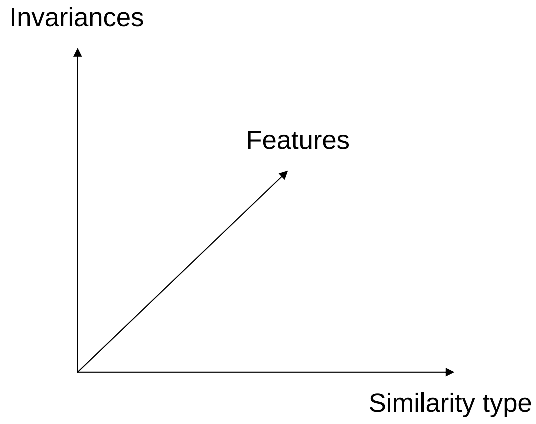

Figure 2: The 3D similarity space model: A combination of similarity type, 3D feature, and invariance setting constitutes a similarity notion.

Addressed features

The similarity types given above can rely on different types of features defined for 3D content description. Important classes of features are based on **geometrical**, **topological**, or **structural** properties of the models. Also, **volumetric** features, or features based on **surface properties** are candidates. It is also possible to consider **annotation** and **markup** information, e.g., processing history or cross-reference information, which might be associated with a given 3D object. Local markups are suited to implement the functional similarity type defined above. However, considering annotation information for 3D similarity evaluation requires a standardized annotation scheme, which allows to compare the individual annotation entries. A major problem in this context is that the types of possible features depend on the 3D representations available, as not all representations allow analysis of all of these features, cf. the discussion in Section 2.

Invariances

The similarity and feature type dimensions are complemented by addition of certain invariance modifiers to the similarity notion. Typical invariance modifiers specify e.g., whether or not **position**, **scale**, or **orientation** of the 3D content in their respective coordinate systems is to be considered when evaluating similarity. Also, invariance regarding the **level of detail** of the 3D content can be desired. Many more invariance modifiers are possible, and in part depend on the type of similarity and feature specified. Integrating such invariance requirements into analysis algorithms is not trivial, but a problem in its own for many existing 3D analysis algorithms.

Figure 2 illustrates the space of 3D similarity notions spanned by these dimensions. The space of possible 3D similarity notions is huge. This model is useful for identifying important similarity notions as well as “blind spots” in this space, which have not been appropriately addressed by research yet (cf. also the next Section). An implication of this large similarity space is that any 3D Digital Library, in which at least some 3D similarity notions are to be sup-

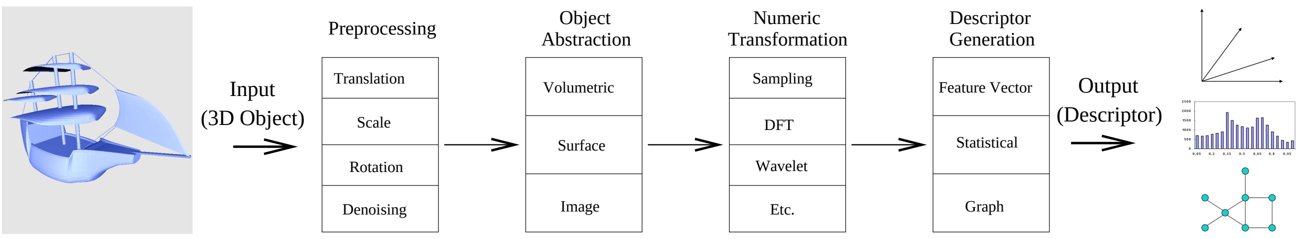

Figure 3: 3D FV extraction process model.

ported, requires significant efforts regarding implementation of 3D analysis algorithms, the output of which is used to quantify the degree of similarity. Also required are then efforts towards the *user interface* side, where the user is to be supported in specifying similarity queries.

3.3 Implementing the similarity notions

Although the 3D similarity notion space is conceptually rich, current methods for retrieval of 3D content mostly focus on the *global geometric similarity* notion. The *transformation* approach determines the similarity between two 3D objects under concern by the cost associated with efficiently transforming (morphing) the global geometry of on object into the other. A simpler, yet efficient approach relies on shape *descriptors*, which are calculated offline for the 3D content. At query time, not the objects themselves, but their descriptors are used for similarity evaluation.

Due to its simplicity and generality, *feature vectors* [4] are often employed as efficient model descriptors. The basic idea is to encode the output of certain shape analysis algorithms in form of vectors of real-valued numbers, effectively representing the 3D content by points in a high-dimensional feature vector space. Distances between the point representations can be calculated, and used as a measure for the (dis)similarity of the underlying objects.

In [3], a process model for the generation of global shape descriptors was presented. Figure 3 illustrates the model which was introduced to capture the essential processing pipeline of most of the current retrieval-oriented shape descriptor algorithms. Briefly, a 3D model is first preprocessed to achieve desired invariance properties. Then, the basis for feature extraction is selected by considering the model as a volume, or by abstracting to its surface or a projection of the model. From this abstraction, low-level features such as the distribution of surface curvature, or shape features calculated from rendered 2D object images, can be captured. From the outcome of this analysis descriptors are formed, with the basic forms being vectors, histograms, or graphs.

To date, a magnitude of low-level 3D analysis algorithms have been proposed, as surveys indicate [2, 12, 9]. Most of them were *heuristically* introduced and motivated by techniques from geometry and image processing. Their suitability for solving the retrieval problem cannot be analytically decided, but needs to be experimentally evaluated by benchmarks [11]. Figure 4 illustrates the evaluation of a number of different low-level descriptors on an exemplary query for a 3D model. As can be seen, each descriptor (one row per query) yields another set of answer objects.

Low-level features are usually efficient to extract and store, and can be quickly evaluated at query time. Besides global shape description, low-level features have recently been used in approaches attacking the partial-similarity problem. These first identify a set of "salient" or "interesting" local features, which are then matched against each other in a second step [7, 6]. Feature vector descriptors can also readily be used together with relevance feedback and machine learning techniques to improve retrieval effectiveness.

The most important drawback, however, is that low-level features are not aware of higher-level *semantic concepts* underlying the objects or object parts, and that the correspondence between low-level features and high-level semantic concepts is not clear for most of the low-level features. A prerequisite before retrieval in 3D Digital Libraries can take place on the semantic level could be the definition of a catalog of semantic shape features, followed by the development of low-level analysis algorithms which can detect, describe, and compare the identified semantic features. Obviously, this is a most challenging problem in 3D content-based retrieval for the following years and beyond.

4. THE 3D DIGITAL LIBRARY VISION

In Sections 2 and 3, we discussed the current situation in representing, describing, and retrieving 3D content. The results achieved so far are remarkable, yet they raise further research challenges which have to be solved to unlock the full functional potential of 3D Digital Libraries. If the problems in representation and content-based organization of 3D content are solved, new and highly productive 3D applications will emerge. Semantically enriched markup, indexing, and retrieval will allow the deep integration of 3D content into Digital Libraries, and fascinating new applications can be envisioned. We sketch some of them in the following.

Intelligent 3D data acquisition

Intelligent 3D scene acquisition will consist of fully automatic segmentation and interpretation of any scanned scene in such a way that each contained object is recognized, its degrees of freedom are identified, and it becomes readily editable in a way that respects its inherent structure and semantics. Appropriate *shape templates* will be associated with the elements in the acquired scene, semantically enriching the data.

Semantic editing and modeling

Based on the recognition not only of low-level shape features, but also of structure and semantics, new possibilities to work with 3D data will emerge. The separation of function and shape will allow for highly efficient editing operations, where the shape of a model can be instantaneously edited by manipulating a few core high-level parameters. Also, the composition of new 3D objects and scenes based

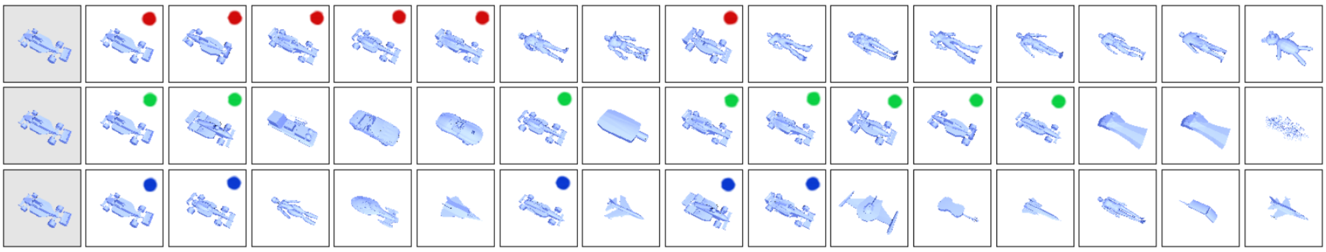

Figure 4: Query-by-example for a Formula-1 racing car model in a 3D repository. Different automatically-extracted low-level object descriptions were used in executing the query, producing different result sets.

on existing content will be greatly simplified, once 3D editors are made aware of semantic properties of the models. A first approach of *modeling by example* [5] illustrates the potential of this paradigm.

Intelligent content-based access

3D search engines will become highly intelligent tools once semantic shape analysis methods are available. If confronted with a user query, the search system will evaluate many different similarity notions on all conceptual levels. From that evaluation, the system will determine the most appropriate similarity notion, and then present the user with the most promising search results. The user will be offloaded from the difficulties in 3D search as currently given, e.g., manual feature selection or supplying much explicit relevance feedback.

Automatic analysis of large 3D collections

Once the representation problems are solved and semantic shape analysis algorithms are available, the fully automatic population of huge 3D Digital Libraries can take place. Content from many heterogeneous sources will be integrated into a decentralized, unified repository. The repository structure will be analyzed, and a suitable organization will be automatically learned from the data.

5. CONCLUSIONS

In this paper, we discussed fundamental aspects and urgent research challenges in 3D Digital Library technology. We argued that based on the technological effects on the production and consumption side, in the near future massive amounts of 3D content will become available. For Digital Library support of these massive data amounts to become effective, a couple of key problems regarding the 3D data level have to be addressed, e.g., in data representation, file format, and stable markup. Furthermore, shape analysis algorithms need to become aware of 3D semantics, to be able to implement advanced automatic organization and retrieval capabilities, and to create large libraries of 3D content that can be effectively searched and accessed. Specifically, low-level features alone are not enough to this end, but defining a catalog of semantic 3D features, and designing algorithms for their robust detection in 3D content are a promising starting point to this end. Once these research challenges are appropriately addressed, Digital 3D Libraries will offer new, highly productive applications in intelligent content acquisition, editing, organization, and accessing.

Acknowledgments

This paper extends previous work by the authors that appeared in IEEE Computer Graphics & Applications in 2007 [8, 1]. The work presented here was supported by the project Probado (www.probado.de) funded by the German Research Foundation (DFG), and by the EPOCH (www.epoch-net.org) and DELOS (www.delos.info) Networks of Excellence funded by the European Commission. It was also partially supported by the Millennium Nucleus Center for Web Research, Grant P04-067-F, Mideplan, Chile (first author).

6. REFERENCES

[1] B. Bustos, D. Keim, D. Saupe, and T. Schreck. Content-based 3d object retrieval. *IEEE ComputerGraphics & Applications*, 2007. To appear.

[2] B. Bustos, D. Keim, D. Saupe, T. Schreck, and D. Vranić. Feature-based similarity search in 3D object databases. *ACM Computing Surveys (CSUR)*, 37:345–387, 2005.

[3] B. Bustos, D. Keim, D. Saupe, T. Schreck, and D. Vranic. An experimental effectiveness comparison of methods for 3d similarity search. *International Journal on Digital Libraries, Special Issue on Multimedia Content and Management*, 6(1):39–54, 2006.

[4] R. Duda, P. Hart, and D. Stork. *Pattern Classification*. Wiley-Interscience, New York, 2nd edition, 2001.

[5] T. Funkhouser, M. Kazhdan, P. Shilane, P. Min, W. Kiefer, A. Tal, S. Rusinkiewicz, and D. Dobkin. Modeling by example. *ACM Trans. Graph.*, 23(3):652–663, 2004.

[6] T. Funkhouser and P. Shilane. Partial matching of 3D shapes with priority-driven search. In *Symposium on Geometry Processing*, June 2006.

[7] R. Gal and D. Cohen-Or. Salient geometric features for partial shape matching and similarity. *ACM Trans. Graph.*, 25(1):130–150, 2006.

[8] S. Havemann and D. Fellner. Seven research challenges of generalized 3d documents. *IEEE ComputerGraphics & Applications, Special Issue on 3D Documents*, 27(3):70–76, May/Jun 2007.

[9] N. Iyer, S. Jayanti, K. Lou, Y. Kalyanaraman, and K. Ramani. Three-dimensional shape searching: state-of-the-art review and future trends. *Computer-Aided Design*, 37:509–530, 2005.

[10] M. Pratt. Extension of iso 10303, the step standard, for the exchange of procedural shape models. In *Proc. International Conference on Shape Modeling and Applications (SMI04)*, pages 317–326, June 2004.

[11] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The princeton shape benchmark. In *SMI '04: Proceedings of the Shape Modeling International 2004*, 2004.

[12] J. Tangelder and R. Veltkamp. A survey of content based 3D shape retrieval methods. In *Proc. International Conference on Shape Modeling and Applications (SMI'04)*, pages 145–156. IEEE CS Press, 2004.