Using Ontologies and STEP Standards for the Semantic Simplification of CAD Models in Different Engineering Domains

J. Posada[1], *C. Toro*[1], *S. Wundrak*[2], *A. Stork*[2] .
[1]. *VICOMTech Research Centre, Donostia-San Sebastian, Spain,*
[2]. *Fraunhofer Institute for Computer Graphics, Darmstadt, Germany*

Abstract
We present in this article an Ontology based compression system that uses STEP compliant standards for the compression and design review visualization of large CAD data sets. Our approach is orthogonal to the traditional techniques applied in the field as we complement previous works introducing semantic criteria along with algorithms for the categorization, simplification and user-oriented adaptation of the engineering components described by domain specific standards. As an example we have implemented two test cases in two specific domains -ISO-STEP 13013-AP227 in the case of industrial Plant Design and CIS/2 in the case of Steel Detailing Design (Structures Design).

Keywords: Interactive Semantic Compression, Walkthroughs, Large Model Visualization, Large Data Sets, Design Review, ISO-STEP.

1. Introduction

The possibility to have interactive walkthroughs for the very large geometric datasets offers clear benefits as it reduces Design times and allows the engineers and Designers to detect early potential construction problems that may appear **([6], [11], [3])**. The main approaches presented in the literature in the field of Large Model Visualization (LMV) for Design review purposes are mainly related to algorithms and compression methods to be applied to the geometric entities that compose the Computer Aided Design (CAD) model **([11], [3])**. However no special attention is given to the fact that these models contain mainly well known engineering parts arranged in concrete shapes, The consequence is to have models of millions (or even billions) of triangles when converted from CAD to Virtual Reality (VR) rendering environments, whose redundancy could be more intelligently exploited, taking into account the nature of the models and the semantics explicitly contained in the CAD model. Many legacy systems, as well as only 3D CAD representations of integrated Plant Information Managers (PIM) systems, are often the only basis for Design Review walkthroughs. Unfortunately, high-level semantics are not fully exploited for visualization in CAD systems although some semantic information is implicit in the geometry. We present in this article a system for the Design review visualization of large data sets that can be applied in various engineering domains dealing with larges amount of data. We base our approach in the categorization, simplification and semantic compression techniques for the engineering parts contained in the model. The supporting models are based on international standards for product data in their specific domains -ISO-STEP 13013-AP227 **[8]** in the case of industrial Plant Design and CIS/2 **[5]** in the case of Steel Detailing Design (Design of structures for roofs, etc). Our approach complements and enhances other efforts of the research community by adding semantic criteria to the simplification techniques. This paper is arranged as follows: In section 2 some background is presented. Chapter 3 explains our proposed semantic compression technique. Chapter 4 presents two case studies with statistics and results. In Chapter 5 we formulate our conclusions.

2. Background

There are four families of techniques used commonly in interactive walkthroughs of large databases **[10]**: (*i*) rendering acceleration techniques, (*ii*) database management, (*iii*) interactive collision detection, and (*iv*) system integration. As seen in **[3]** the main acceleration techniques used are basically visibility culling, object simplification and image-based representations. Geometric simplification techniques e.g. Levels of Detail (LOD), Hierarchical Levels of Detail (HLOD) **[9]** give good results in handling massive data sets; the integration of LOD and good occlusion culling techniques are usually the key factors to achieve interactive rates in walkthrough systems **[1]**. On the other hand, there are emerging commercial applications (e.g. *NavisWorks* **[13]**, *Mantra4D* **[12]**) that incorporate the latest graphics hardware accelerations as well as many of the classical culling and simplification techniques with good results. In a direction similar to the first part of our work, Shikhare **[17]** describes an algorithm for automatic discovery of repeating geometric features. The collected information is then used for compressing the model by removing the redundancies on its representation, although it doesn't take into account the semantic significance of the repeating structures found. In general, we found that in most approaches the semantic implicit in the geometric representation is not considered for the model, but assumed, no matter what the user's motivation or background is.

3. An architecture for the Semantic compression of CAD models

In this chapter we present our Design review walkthrough system. We use semantic compression added to simplification techniques of the geometrical data to increase the efficiency and complement the traditional Computer Graphics methods in the field. The architecture can be seen in figure 1.

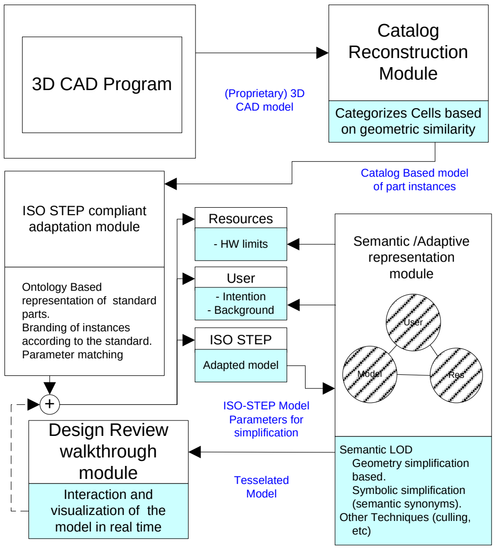


Figure 1: Our semantic walkthrough architecture

We take as a starting point any proprietary geometric 3D CAD representation belonging to the Industrial plant Design or Steel Detailing Industry. We deliberately assume that no other information is available (e.g. from a modern PIM system) since for many reasons –legacy data, databased model exchange between companies, etc. - this is the general case. We then reconstruct automatically the families of engineering parts in the model; associate those families to the standard, introduce both geometric and semantic object simplification techniques, and present the adapted model in an interactive system for Design review walkthroughs. The work of this paper is strongly based on our previous work **[15]**. However, **[15]** still left open the following topics open:

- Searching and classification of instances (cell matching algorithm).
- The standards adaptative/semantic module.
- An interaction with the Ontology model.

3.1 The Catalogue Reconstruction Module

This module traverses the 3D CAD model identifying groups of geometric primitives (we call these groups/families *cells*) automatically, and categorizes them in groups based on geometric similarity. The 3D CAD model creation in the domain of Plant Design or Steel Detailing is based in the parametric definition and selection of appropriate engineering parts from specific catalogues. However, the resulting CAD models usually do not contain any explicit instancing information, and the first step towards an increased semantic representation of the model is to group these cells using the cell-matching algorithm.

3.1.1 Searching and classification of repeating elements in a CAD model

General methods for searching repeating structures in unorganized sets of geometric primitives exist, but are usually slow on large models **[17]**. The estimated runtime for models of our size may easily exceed a full day. We have focused on a fast algorithm for finding instances (repeated cells no matter their orientation or position in space not sorted, as in a soup of elements), which is a reasonable approach since the engineering parts rarely correspond to exactly one geometric primitive. The real-world models we have studied from different systems preserve this grouping structure. However, no assumption is made regarding the internal order of the primitives inside a cell. For the automatic classification of instances an algorithm is needed that decides whether two cells Ci and Cj match each other, this is, if Cj can be considered as a spatial instance of Ci. In a real world model, several thousands of cells will be matched against each other in the catalogue reconstruction module. To preserve the near real-time nature of our approach, one important requirement for the algorithm is to be fast.

Topology based matching

We considered a first approach matching two cells solely based on the cell's topology. The cell topologies of both cells were traversed recursively and the geometric primitives were compared one-by-one. A difference in their topological structure or the kind of geometric primitives was interpreted as a do-not-match result. Only if the topology matching was successful, a more detailed geometric matching was performed. This approach proved to be very fast, but not feasible, since visually identical elements often had variations in their cell structure. Furthermore, we had an initial assumption that proved to be wrong: that cells were stored in a normalized orientation – since they originally came from a catalogue – and were then positioned in the model using a transformation matrix. Instead, the cell's geometry is often relocated by direct redefinition of the parameters of the geometric primitives.

Point-clouds based matching

There are several methods proposed in the literature to match point clouds representing 3D surfaces **[3]**, since the problem of registration of point clouds is very relevant in several fields (3D model acquisition, reverse engineering, quality control, etc.). Some of the most popular algorithms are the Iterative Closest Point (ICP) algorithm **[3]**, with several variations, the Least Square Surface Matching (LS3D) method, the spin images method **[7]** and the Iterative Closest Points using Invariant Features (ICPIF) **[18]**. We have considered the possibility to apply some of these algorithms to our cell matching problem, generating point sets from the cells. In the generic registration problem the correspondences between the point sets are unknown a priori **[3]** and no one-to-one matching between the clouds points can be assumed (since they come usually from scanners). In our case, however, given the special conditions of our problem, this one-to-one correspondence exists, making the task easier.

Thus, we developed a simplified algorithm (somehow similar in the approach to ICPIF, although much simpler and restricted) that could be applied successfully for the cell matching and classification.

3.1.2 The Cell matching algorithm

We can reformulate the cell-matching problem as follows:

(i) Given two cells Ci and Cj, each composed by an unordered set of geometric primitives, Cj matches or is an instance of Ci if a rigid transformation (rotation/ translation) matrix T exists that transforms Cj into Ci (see figure 2). The cell-matching algorithm must:

Decide if Cj matches Ci within a given tolerance.
(ii) Obtain the transformation matrix T.

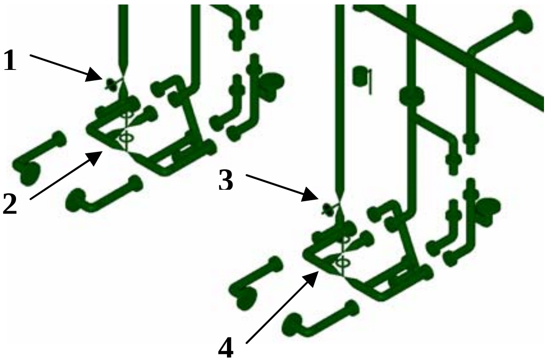


Figure 2: Sample of non-explicit instances in the 3D CAD model

A cell Ci is composed by an unordered set of geometric primitives GCij:

Ci = {GCi1 , GCi2 , … , GCim)

Each G is characterised by:

1) A *characteristic point set* (CPS_G)*:* A set of points associated univocally with the spatial position of G. Notice that *CPS* is not an exhaustive set of surface points.
2) A *characteristic scalar set* (CSS_G): A set of scalar values associated univocally with the dimensions of G.
3) A *characteristic type* (CT_G)*:* A code correlative with the geometric primitive type.

CPS_{Ci}, CSS_{Ci}, CT_{Ci} are the sets formed with the *CPS, CSS* and *CT* of all primitives G_{Ci}.
For example:

$$CPS_{cyl} = \{P_{origin}, P_{end}\} \quad CSS_{cyl} = \{r\} \quad CT_{cyl} = \{\texttt{CYL}\}$$

Where the points are the centres of the covers, r the radius and `CYL` the type of the `cylinder` primitive. Notice that *CPS* varies between instances of the same cell but *CSS* and *CT* don't.

The *point cloud* PC_{Ci} is the point cloud formed by all the *CPS* of the primitives *Gci.* Each *CPS-point* keeps a link with its corresponding *CSS* and *CT.*

$$PC_{Ci}=\{CPS_{G1(P0)},\ldots,CPS_{G1(Pn)}, CPS_{G2(P0)},\ldots, CPS_{Gn(Pm)}\}$$

The *ordered point cloud* OPC_{Ci} is the ordered set of all points in PC_{Ci} with respect to the squared Euclidean distance d^2_i from each point to the barycenter (geometric centroid) of PC_{Ci}.

$$OPC_{Ci} = \{\{CPS_{Ga(Pb)}, d^2_{ab}\}, \{CPS_{Gc(Pd)}, d^2_{cd}\}, \{CPS_{Ge(Pf)}, d^2_{ef}\}\}$$

With $d^2_{ab} \leq d^2_{cd} \leq d^2_{ef}$

The core of our matching algorithm is based on the geometric comparison between PC_{Ci} and PC_{Cj} as follows (if *Ci* and *Cj* have both *n* geometric primitives):

1) Discard obvious non-matching cells (different count of primitive types *CT*).

2) Get OPC_{Ci} and OPC_{Cj} ordered with respect to the distance of each point to the respective barycentre.

3) Check that the ordered vector V_{di} with these distances in OPC_{Ci} is equal within a given tolerance to the the corresponding ordered vector V_{dj} of distances in OPC_{Cj}. If not, *Ci* and *Cj* don´t match.

4) Compare also the respective CSS_{Ci} and CSS_{Cj} (following the order given by V_{di}). If not equal, it means that although the point clouds coincide, the invariant scalar values don't; therefore *Ci* and *Cj* don´t match.

5) Get 3 non-collinear points P_{i1}, P_{i2}, P_{i3}, that have unique values of squared Euclidean distance to the barycenter, this is, and no other points in OPC_{Ci} have the same distance to the barycenter. As OPC_{Ci} is already ordered by this distance, this is fast.

6) Get the corresponding 3 points P_{j1}, P_{j2}, P_{j3} in OPC_{Cj} such that $d^2_{j1} = d^2_{i1}$, $d^2_{j2} = d^2_{i2}$, $d^2_{j3} = d^2_{i3}$. As OPC_{Cj} is ordered too, this is straightforward.

7) Calculate the rigid transformation *T* that transforms P_{j1}, P_{j2}, P_{j3} into P_{i1}, P_{i2}, P_{i3}. If *T* exists, *Ci* and *Cj* match.

In the rare case that the 3 points of step (5) cannot be obtained, a more general algorithm (e.g. ICP) could be executed. Our clouds contain typically under 100 points, and we have several thousands cells, classified in tenths to hundredths of groups. In the practice we have always been able to get these points.

3.2 ISO-STEP Compliant Adaptation Module

A 3D model of an Industrial Plant or a Steel Detailing model of a structure typically has representations of pre-defined engineering parts. These elements are described by an ISO standard, STEP-10303-227 **[8]** in the domain of Plant Design, and CIS/2 **[5]** in the domain of Steel Detailing. We integrated a module to explicitly associate this semantics to the geometric parts from the reconstruction described in section 3.1.

3.2.1 The ISO STEP 10303-227 Standard

ISO STEP-10303-227 **[8]** is part of an international Standard for the computer interpretable representation and exchange of product data. Product data represents information in formal manner suitable for communication, interpretation, or processing by human beings or computers. The objective of STEP is to provide a neutral mechanism capable of describing product data throughout the life cycle of a product independent from any particular system. The nature of this description makes it suitable not only for neutral file exchange, but also as a basis for implementing and sharing product databases and archiving. The core of STEP consists of a collection of conceptual models, which describe the content, and structure of product data items. These data models, also called information models, are formally specified in the modelling language EXPRESS **[8].** The Application protocol 227 describes the specifics for plant spatial configuration; Figure 3 shows an excerpt of the Express diagram with the elbow and flange elements (Plant Design domain).

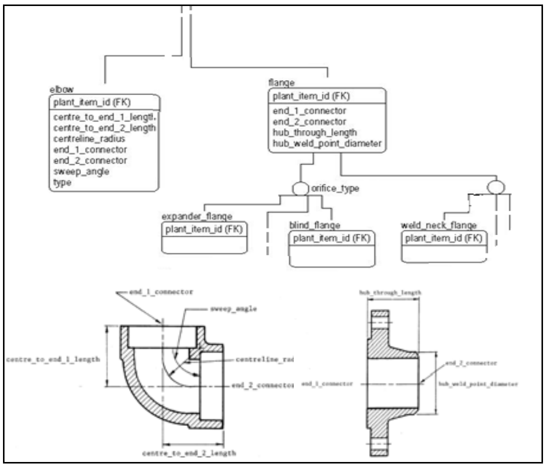


Figure 3: ISO-STEP 10303-227 Express detail.

3.2.2 The CIS/2 Standard

CIS/2 is based on deliverables of the Eureka EU130 CIMSteel Project and is an extension to the general STEP model for the specific case of the Steel Detailing industry **[5].** CIS/2 is an extended release of the CIMSteel Integration Standards (CIS), a set of formal computing specifications that allow software vendors to make their engineering applications mutually compatible. The CIS/2 documentation specifies what information may be transferred between software applications, and how that information must be structured in a repository or data exchange file. CIS/2 substantially extends the engineering scope of CIS/1, and introduces advanced data management capabilities to enable data sharing.
At its simplest, the CIS provides specifications and guidelines for the development and implementation of translators that enable the users of engineering software to export data from one application into another. However, CIS/2 also allows software vendors to support concurrent engineering via more direct mechanisms for information sharing and management. Thus, CIS/2 also provides specifications and guidelines for the development and implementation of Database Management Systems (DBMSs) built around the CIS and its related technology. Such a DBMS is known as a Product Model Repository (PMR). There are several complex structures that are used repeatedly throughout CIS/2. Most of these structures come from the integrated resources that are common to all ISO-STEP product models.

3.2.3 Motivation for an Ontology Support

We have modelled a full Ontology related to the ISO-STEP standards in both case studies (Industrial Plant and Steel Detailing) because our ultimate objective is to have a system where the concepts and relationships of the domain could be modeled and queried using semantic criteria **[2]**, beyond the mere data modeling structures of the norm. This Ontology modelling also allows a more transparent interrogation of the user task/profile that can also modelled as Ontologies (see next chapter). We use ISO-STEP models as a basis to develop this module (Figure 4 b). The main reason to use this approach is related to the fact that STEP is only a data exchange format, but our requirements for semantic simplification required higher capabilities to express relationships and concepts.

3.2.4 Construction of the Ontologies

The Ontologies are modelled using Protégé 2000 **[16]** ,Other software for Ontology modelling were surveyed but since the comparison with other Ontology tools is outside the scope of the paper, we will report our results using this editor. We have modelled the Ontologies adapting the tags and relationships (to be more suitable for a knowledge representation model) presented in the ISO STEP **[8].** This serves as an important contribution to the model part of the semantic triangle described in **[15]**. The current Ontology of the domain model in the case of industrial Plant Design, has a total of 298 classes, 143 slots and 451 frames, and currently represents the 60% of the ISO application protocol 227. In the case of Steel Detailing model, have a total of 186 classes, 87 slots and 297 frames, and currently represents the 40% of the CIS/2 Standard. For the User and Task parts of the semantic triangle, we based our implementation in similar concepts implemented by our group in the European Project WIDE (IST-2001-34417). See figure 4a.

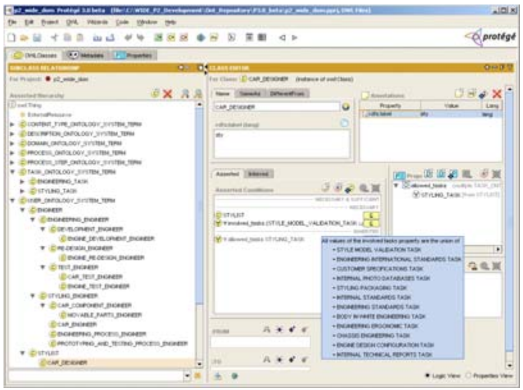


Figure 4 a: User and Task Ontologies based on the WIDE approach.

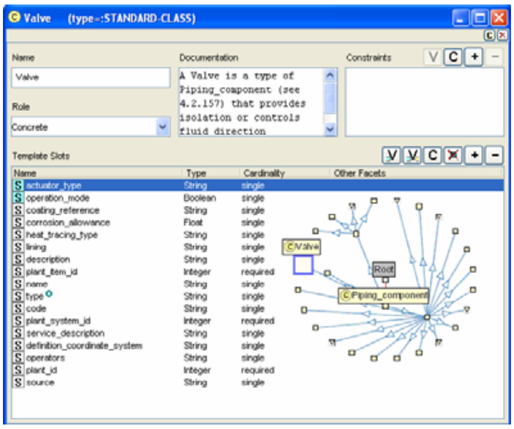


Figure 4 b: Model STEP Compliant Ontology

For visualization and interaction purposes we have tested OntoViz and TGViz **[16]** plug-ins, and the queries are made through an RDF – OWL compliant parser that interacts with the adaptive visualization module.

3.2.5 Interaction with the STEP-based Ontology

In order to select an adaptive representation of the model we query the Ontologies giving a user task/profile (manager, engineer, etc) the available computer resources and the model (three Ontologies in total). The model Ontology is filled with the real parameters of the CAD model (this is an automatic process), and then a semantic association followed by a semantic adaptation allows the visualization enhancement by producing an output that has embedded juts the needed information for each user/task profile and available computer resources.

3.2.6 Semantic association of parts with the standard

In order to add the semantic information we follow a two stages approach.

1) Name each group of cells after an ISO – STEP compliant concept. We call this process "*Branding*". The user visualizes one representative part of the cell group and matches it with a concept of the Ontology in a graphical concept tree (see Figure 5).

2) Once the cell group is associated with a concept in the Ontology domain, the user matches semi-automatically the cell parameters (geometric features) with those parameters specified in the ISO-STEP standard. We call this process "*Matching*".

3.3 The Semantic/ Adaptative Representation Module

This module takes as input the adapted 3D CAD model in which the families of cells identified in the Catalog Reconstruction module already correspond to ISO-STEP parts. However, this is only one of the aspects we consider in order to make a good semantic compression of the model for a Design Review walkthrough scenario.

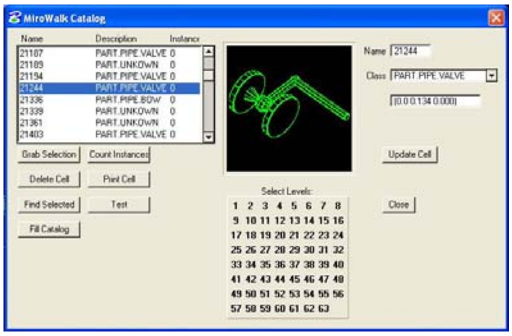


Figure 5: Cell group branding

As explained in (**[14] [15]**), we have defined a framework in which three factors influence the final adaptation of a 3D CAD model for Design Review walkthroughs: (*i*) *The user intention and background*, (*ii*) *the available resources*, and (*iii*) *the model characteristics* (already processed in the previous module, see part 3.1). In many walkthrough applications, there is very little or no consideration of the profile and motivation of the user of the system. We introduce explicitly the concepts of *user profile* and *user task*, which influence the final output model in this semantic adaptation module. Thus, the parameters used by different Computer Graphics techniques (such as LOD, culling, etc.) inside the Adaptive Representation module are defined (with a rule-based approach) according to the user needs. In a similar way we take into considerations the available resources (e.g. clusters of PCs vs. single PC, available RAM, etc.) to prepare the walkthrough experience, creating different adapted representations in each case. We are now moving from the rule-based adaptation system of *user profile* and *available resources* towards a deeper integration with the *model characteristics,* by modelling those two aspects in especial Ontologies that can be integrated with the model Ontology described in 3.2

3.4 Adaptative Representation overview

With the adapted model, as well as the parameters for graphical optimization of the final tessellated model an output model is produced to be displayed in the walkthrough viewer. We have implemented several CG techniques in the mentioned tool, although in this article we focus in a special use of the LOD technique that has reported substantial improvement in the walkthrough performance.

Geometric LOD vs. 3D Semantic Symbols

As explained in section 2, LOD techniques are based on a varying accuracy in the representation of a 3D object. Usually LODs are either automatically generated from the geometric definition of the object or they are modelled ad-hoc. In both cases the geometric similarity between the LOD and the object is preserved as much as possible. On the other hand, the use of 2D symbols is a widespread engineering practice that is slowly moving also to the 3D representation **[4].** Once we have the ISO-STEP adapted model, we generate alternative representations according to the parameters given by the previous module: *(i)* We use parametric *geometric LOD* for those components of the model that have the largest influence in the number of triangles generated. These geometric LOD are based on the standard parametric parts of the ISO-STEP compliant representation (instead of basing the LODs on the original 3D objects in the CAD model). *(ii)* On the other hand, we generate in parallel alternative 3D semantic symbols for all components, which gives a much higher semantic compression ratio (better compression) without semantic loss for special user profiles and tasks. This of course depends on specific configuration of users/tasks, models and resources. In Figure 6 some ISO-STEP elements (Plant Design domain) are selected to show the adapted representation and the elements to me matched (branding and matching as explained in section 3.2.3).

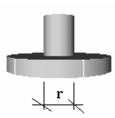
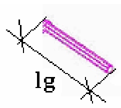
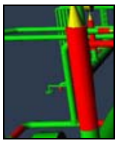
3D CAD (Geometric LOD)	**Selected Semantic representation (parametric)**	**ISO-STEP Matched parameters**
ISO-STEP 10303-AP227 FLANGE (COD. 4.2.84)		
	 **r** = radius **s** = side **XYZ** = Coord. System **P** = position (px,py,pz)	**STEP 4.2.84.3** Hub through length = s **STEP 4.2.84.4** hub weld point diameter = 2*r **STEP piping** connectors : give XYZ, P.
ISO-STEP 10303-AP227 STRAIGHT_PIPE (COD. 4.2.232)		
	 **c** = complexity **r** = radius of joint **XYZ** = Coord. System **P** = position (px,py,pz) **lg** = length	**STEP4.2.52.3** end to end length = lg **STEP piping** connectors : give r, XYZ, P.

Figure 6 – Adapted representation of some components in the Plant Design Domain: geometric LODs and semantic symbols

3.5 Design Review walkthrough module.

Once the semantic data is added and used to simplify the elements via the semantic synonyms, the elements are ready for visualization and walkthrough evaluation. In the Design Review Walkthrough module, we implemented not only the traditional LMV techniques presented in part 2, but also the semantic compression module presented in this article. The visualization tool is a stand-alone application and includes tools for pan, zoom, and navigation in real time. The collection of cells in a tree hierarchically structured appears at the left side of the scene, allowing the selection of a given cell or group. Other possibilities present in the viewer are the per-part identification; the seek function and the possibility to manipulate the parts (move spatial position, scale, etc).

4. Case studies.

4.1 A chemical plant, statistics and results

We present in this chapter the results of using our framework in a real-world chemical plant model (of a well known brand whose name we must keep confidential). The model was generated in a professional Plant Design system, whose 3D CAD geometric representation was used as the basis for the Design Review walkthroughs. It is a large three-story building with a structural skeleton as main supporting construction. The building halls are filled with a complex piping system spanning the three stories and also reaching into the outside environment. Attached to the piping system are numerous flanges, boilers, valves, tanks, fittings, pressure gauges, etc. Especially the piping system and its attached parts contain a lot of curved elements that are very costly to render. After a thorough analysis of the model we found out that more than 65% of the triangles (even after the use of pure geometric LOD) were produced in the piping system substructure. This high proportion (60-80%) is preserved also in other models we have tested. We have therefore concentrated our efforts in this subsystem, and the techniques used for semiautomatic detection, adaptation to the ISO-STEP 10303-AP227 standard, and semantic compression representation, are focused on the typical components of this subsystem: valves, flanges, elbows, pipe sections, piping clamps, T-adaptors, sprinkler heads, etc.With regard to the Catalog Reconstruction Module, it is interesting to see how the elements in this concrete model were grouped.

Total Elements	99799 (100%)
Elements within cells	64520 (65%)
Number of Cells	13147
Number of Cell families	1104
Instantiated Cells	12043

Table 1. Distribution of Elements in the Model

Table 1 shows that a high proportion of the total of primitive elements in the model are indeed grouped in cells (65%). This accounts also for a high proportion of the total number of triangles rendered (about 87% of the triangles, even using geometric LOD with complexity = 0.3). From the geometry not organized into cells, another 10% of the triangles come from about 100 complex objects -boilers and tanks- and 3% of the triangles are part of other repeating element like columns, windows, square pipes, etc. The Catalog Reconstruction Module (3.2.) was able to classify the 13147 cells in 1104 families with the Cell Matching algorithm. Table 2 shows that actually a large number of cells belongs to a few cell families of the ISO-STEP 10303-AP227 standard; the rest (unknown) are relatively sparse but are not very relevant in relative weight for the final result. This means that the ISO-STEP 10303-AP227 Adaptation Module (3.3.) was able to classify 82% of the total cell families and relate them to the standard. It is evident that a brute-force, blind conversion with very high quality from the original CAD geometry would create an untraceable model in the practice for Design review.

Plant Design Component	Number of Instances	% of total instances
Valves	867	7 %
Elbows	2064	16 %
Flanges	3663	28 %
Pipe Section	3509	27%
T-adaptors	425	3%
Clamps	191	1%
Unknown	2428	18%
Total Cells	13147	100%

Table 2. Categorization in cell families

Just the valves would create several million triangles. Therefore we take as basis for our comparisons a model already including several simplifications, especially the use of geometric LOD on the original CAD geometry with a complexity of 0.3. This complexity factor in our system is a parameter between 0.0 and 1.0, where 1.0 is the highest accuracy representation (for instance, for NURBS primitives tessellation, it specifies the relative deviation with respect to a predefined threshold). We estimate that a value under 0.3 would create distortions on the tessellated model too evident for the user. It can be seen that the use of this technique (geometric LODs) brings a high reduction on the number of triangles generated. This and other computer graphics techniques are the basis of common walkthrough systems (culling, pre-fetching, impostors, scheduling, etc.). We bring an additional improvement to these traditional techniques by introducing semantic parametric representations, based on the knowledge of the domain and the related standards, bringing an even better improvement factor, as shown in the table 3. The element with the highest reduction (valve), for example, is represented semantically with just 7,68% of the best geometric LOD simplified object. In the case of the clamps, however, the semantic criterion gives an even better hint: the clamps are just *not shown* (drop culling technique) for this specific task and user. The semantic compression improves in several cases more than 80%-90% the purely geometric simplification approach and this especially in those components with highest weight in tessellated model. The tessellated model using only geometric LOD plus some culling / fetching techniques gave an average number of triangles of 3450 Ktris, with a complexity of 0.3 (this is already a very good simplification factor). However, applying the semantic compression model, we could reduce the model in additional 1659 Ktris, for a net reduction of 51% in the total number of triangles between the semantically compressed model with respect to the geometric LOD simplified model.

4.2 An Steel Detailing Structure, statistics and results

We present in this chapter the results of using our framework in a real-world Steel Detailing Structure. The model was generated in a professional CAD system intended for the Design of spatial structures. It represents the roof of a sport facility and is constituted by Standard elements such as profiles, tubes, spheres and joints (see figure 5).

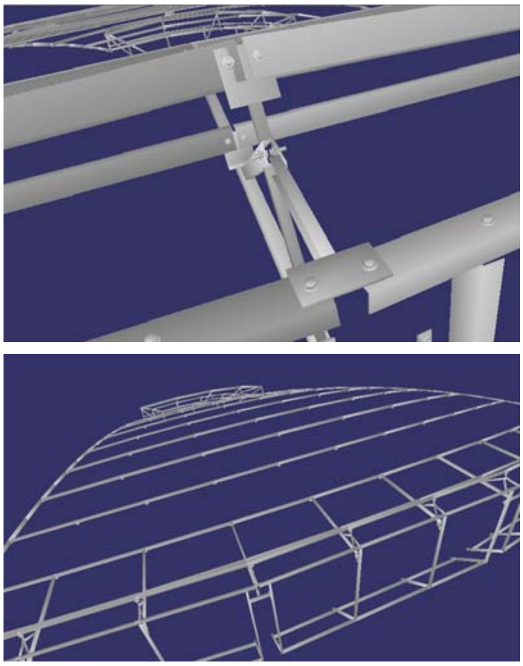


Figure 5 : Steel Detailing Structure (upper image, detail, lower image the whole structure)

It is interested to se how the elements of the model are grouped (see Table 3).

CIS/2 compliant Steel Detailing Element	Number of instances	% of total instances
Profile (CIS/2 Structural Frame Item - Profile)	3000	15%
Joint Tube-Sphere (CIS/2 Joint)	3000	15%
Screw (CIS/2 Fastener Simple Bolt)	6000	30%
Tube (CIS/2 Structural Frame Item - Profile)	6010	30%
Sphere (CIS/2 Node)	1500	7%
Unclassified	743	4%
Total Cells	19510	100%

Table 3. Categorization of elements in the Steel Detailing model

The elements in this model are less complex geometrically speaking compared with the Industrial Plant model elements; the immense majority are sharp and rectilinear (with the exception of the spherical joints/nodes). This means two things (*i*) they can represented by a lower number of triangles and (*ii*) they can be matched faster by the algorithm presented in section 3.1.2 (a comparison between the two models concerning the speed of the algorithm is out of the scope of this paper). In this model the Catalog Reconstruction Module missed to recognize a total of 743 out of 19510 (4%) of the elements contained in the model which means that 96% of the elements were successfully catalogued. After the application of traditional LMV techniques, this model showed that the more complex elements are the Spherical Joints (represented by 2018 triangles) this kind of cell after semantic simplification was represented by 242 triangles which mean a compression rate of 88%.

In Table 4 the resulting compression is shown for the various elements of the model, the comparison between no semantic simplification and the semantic simplification is shown both at a level of unitary elements and per family (e.g. a family of Profiles means all the profiles in the model). Since a given instance can differ to another in dimensions but not in the geometry (e.g. the case of a profile), is possible to have an instance with a few more triangles. Therefore the statistics shown are calculated with a typical instance. It is interesting to point that the overall compression gained by using the semantic approach is about 78%.

Steel Detailing Element	**Without Semantic Simp.**		**With Semantic Simp**		**% Reduction**
	Tris per cell	**Tris per family**	**Tris per cell**	**Tris per family**	
Profile	28	84000	18	54000	36%
Joint	20	60000	12	36000	40%
Screw	156	936000	46	276000	71%
Tube	68	408680	42	252420	38%
Sphere	2018	3027000	242	363000	88%
	Total	**4515680**	**Total**	**981420**	**78%**

Table 4. Number of triangles reduction using semantic criteria.

5. Conclusions

We have presented a semantic compression system for Design review interactive walkthroughs in two different design domains. The use of the semantics implicit in the geometric model of the plant (especially, the fact that it is composed by standard engineering parts), and in the user intention and background, have given a sensible improvement in the application of standard computer graphic techniques -in this article we presented mainly the influence on LODs-. We improved previous works with new modules and algorithms for automatic categorization, simplification, semantic compression and walkthrough adaptation of a complex plant, and tested our system on a real-world model. In order to achieve generality, we based our work on special algorithms and the use of Ontologies related to international standards (ISO-STEP 13013-AP227 and CIS/2). At present time the process data is not yet handled by our approach, but in a future work this information will be modelled as a separate Ontology to be used as a Design advisor to the user.

The degree of knowledge of the user in our approach is related directly to the knowledge of the domain, however no special familiarity with Ontology modelling and queering is needed as this module is not visible to the end user.

6. Acknowledgements

The authors want to express their gratitude to the German Academic Exchange Service (DAAD) and the Spanish Ministry of Education and Science for their support for the exchange of researchers between the participating institutions. This work has been partially supported by the European Network of Excellence AIM@SHAPE, IST project 506766 and the Basque Government, INTEK Project MiroView.

7. References

[1] Andujar C. / Saona-Vazquez C./. Navazo I /. Brunet P. :Integrating occlusion culling and levels of detail through hardly-visibly sets. In Proceedings of Eurographics, 2000.

[2] Borst, W.N: Construction of Engineering Ontologies, PhD Thesis, University of Twente, Enschede, 1997.

[3] Besl, P.J. / Mckay, N.D.: A method for registration of 3D shapes . IEEE Pattern and Machine Intelligence, 14 (2), pp239-256, 1992

[4] CAD SYMBOL Web page, http://www.cadsymbol.net.. Last access on: 5th of December 2004

[5] CIS/2 – CIMSteel integration Standard, WebPage: http://www.cis2.org/ Last access 11 May 2005.

[6] Funkhouser T.A./ Khorramabadi, D./ Sequin, C.H./ Teller, S. : The UCB system for interactive visualization of large architectural models. Presence, 5(1):13–44, 1996.

[7] Jhonson, A.E / Hebert, M.: Using Spin Images for efficient object recognition in cluttered 3D scenes. IEEE Pattern Analysis and Machine Intelligence, 21 (5), pp 433-449, 1999.

[8] International Standard ISO 10303, Industrial Application Systems and Integration – Product Data Representation, Application protocol 227, Plant Spatial Configuration., First edition, Geneva, Switzerland 2001

[9] Luebke, D, Reddy, M., Cohen, J., Varshney, A., Watson, B., and Huebner, R : Level of Detail for 3D Graphics:.; published by Morgan Kaufmann as part of their series in Computer Graphics and Geometric Modeling. ISBN 1-55860-838-9. 2002.

[10] Manocha, D.: Interactive Walkthroughs of Large Geometric Datasets. In Proc. of ACM Siggraph, Los Angeles – USA, 1999.

[11] Manocha ,D.: Interactive Walkthroughs of Large Geometric Datasets. In Proc. of AMC Siggraph, USA, 2000.

[12] Homepage: MANTRA4D. LMV system http://www.mantra4d.com/ Last access on: 2th of May 2005

[13] Homepage:NavisWorks. http://www.navisworks.com/ Last access on: 2th of May 2005

[14] Posada, J; Larzabal, A; Stork, A: Semantic-Based Parametric Control of CAD Model Conversion for Large Model. Cruz-Neira, Carolina (Ed.) u.a.; Ecole Supérieure des Technologies.Industrielles Avancées u.a.: Virtual Concept 2002. Proceedings. 2002, pp. 38-43.

[15] Posada, J; Wundrak, S; Stork, A.; Toro, C: Semantically controlled LMV techniques for plant Design review, Proceedings of DETC/CIE 2004 ASME 2004 Design Engineering Technical Conferences 24th Computers and Information in Engineering (CIE) Conference September 28-October 3, 2004 Salt Lake City, Utah

[16] Protégé Ontology editor WebPage: http://protege.stanford.edu/, last Last access on: 15th of March 2004

[17] Shikhare, D. / Bhakar, S. / Mudur S. P.: Compression of Large 3D Engineering Models using Automatic Discovery of Repeating Geometric Features. VMV 2001 Stuttgart, Germany, November 2001.

[18] Sharp, GC. / Lee, S.W. / Wehe D. K.: ICP Registration Using Invariant Features. IEEE Pattern Analysis and Machine Intelligence, 24 (1), pp 90-102, 2002