Chapter 7
GEODIM: A Semantic Model-Based System for 3D Recognition of Industrial Scenes

**Yuliana Perez-Gallardo, Jose Luis López Cuadrado,
Ángel García Crespo and Cynthya García de Jesús**

Abstract Keeping an inventory of the facilities within a factory implies high costs in terms of time, effort, and knowledge, since it demands the detailed, orderly, and valued description of the items within the plant. One way to accomplish this task within scanned industrial scenes is through the combination of an object recognition algorithm with semantic technology. This research therefore introduces GEODIM, a semantic model-based system for recognition of 3D scenes of indoor spaces in factories. The system relies on the two aforementioned technologies to describe industrial digital scenes with logical, physical, and semantic information. GEODIM extends the functionality of traditional object recognition algorithms by incorporating semantics in order to identify and characterize recognized geometric primitives along with rules for the composition of real objects. This research also describes a real case where GEODIM processes were applied and presents its qualitative evaluation.

Keywords Semantic rules · Recognized geometric primitives · Industrial scenes

7.1 Introduction

Describing in digital scenes what humans perceive by themselves is challenging, due to the large amount of information that must be handled in different contexts. Although several studies such as [1–10] have addressed 3D object recognition with

Y. Perez-Gallardo (✉) · J.L.L. Cuadrado · Á.G. Crespo · C.G. de Jesús
Computer Science Department, Universidad Carlos III de Madrid,
Av. Universidad 30, 28911 Leganés, Madrid, Spain
e-mail: yuliana.perez@alumnos.uc3m.es

J.L.L. Cuadrado
e-mail: jllopez@inf.uc3m.es

Á.G. Crespo
e-mail: acrespo@ia.uc3m.es

C.G. de Jesús
e-mail: cgarcia@alumnos.uc3m.es

excellent results, they may lack from semantic information to describe these objects.

According to [11], an inventory is a detailed, orderly, and valued relationship between elements that make up the assets of a company or person at a given time. It is detailed because the characteristics of each of the elements that integrate the patrimony are specified. It is considered orderly since it groups elements in their respective accounts. Finally, an inventory is valued because the value of every asset is expressed in units.

This research paper was carried out from an industrial approach. It focuses on the description of industrial scenes to create inventories of the elements that compose them. However, the analysis of such scenes may face certain difficulties, such as the size of the plant, the diversity of elements to recognize, or their amount. Also, in the real world, creating an inventory of the facilities of a factory implies several visits to the plant in order to identify and verify all the elements. Therefore, GEODIM aims to create a digital 3D mockup from a scanned 3D point cloud of a medium-large size industrial facility. This model includes all the information required to be familiar with all the types of objects involved and their features.

As proof of concept, GEODIM is applied to the 3D point cloud of the facility, and an enriched digital mockup with logical, physical, and semantic information is obtained as a result. The real case for this study involves the real indoor scene of an actual factory, and two evaluations are presented to assess the quality of classification of GEODIM.

The remainder of this research is thus structured as follows: Sect. 7.2 describes recent advances in the state of the art on object recognition and semantics in object recognition. Then, GEODIM is described in Sect. 7.3, while the real use case and the evaluation of the system are described in Sects. 7.4 and 7.5, respectively. Finally conclusions and future work are addressed in the Sect. 7.6.

7.2 State of the Art

Scenes of indoor factory facilities, building scenes, or even generation of product models, are examples of digital models in the industrial sector. Many efforts have been currently made to create these digital mockups in different contexts and with different objectives. Nevertheless, although their generation is a significant advance, they have failed to describe scenes from a semantic sense. The challenge now is therefore to describe the environment seen in a virtual scene in such a way that its physical and semantic properties are detailed.

Due to the nature of this research, the state of the art has been divided into a couple of subsections to describe the two main processes involved in GEODIM: object recognition and semantics in object recognition.

7.2.1 Object Recognition

Every method for 3D object recognition has special features depending on the domain or field where it is used. For instance, authors in [1] presented a technique for the recognition and reconstruction of surfaces from 3D data by applying line element geometry. Also, researchers in [2] modeled a gesturing hand through the use of key geometrical features of a hand and by constructing a skeletal hand model. Similarly, the work of [3] described a biometric-based security system using hand recognition. In general, the system relied on abductive learning and hand geometric features.

Other examples are the study presented by [4], where authors made use of color information to improve content-based retrieval of 3D models, and the work of [5], who incorporated highly discriminative affine invariant 3D information much earlier in the process of matching. Also, an approach for recognizing 3D objects was described by [6], where authors employed model synthesis to define a large number of possible geometric interpretations of images.

In addition, system identification was used for an emotion recognition model in [7]. The model included an extended Kohonen self-organizing map created by using a 26-dimensional facial geometric feature vector. On the other hand, researchers in [8] analyzed the detection effect of classic edge detection operators in infrared images. In the same year, a features recognition system was also proposed by [10], where the object-oriented structure was used for the generation of a geometric database. Finally, the work of [9] introduced a hybrid system that combined probabilistic graphical models with semantic knowledge for objects recognition in order to identify object in scenes for intelligent mobile robots.

Table 7.1 compares the different object recognition algorithms present in the literature. Although works [1, 8, 10] analyzed scenes and images by considering geometrical features of the objects and also detected existing objects, these studies failed to recognize such items. Similarly, in the area of Biometrics, despite positive

Table 7.1 Comparison of object recognition algorithms

Authors	Type of analysis			Semantics
	Reconstruction	Detection	Recognition	
[1]	✓		✓	✗
[2]			✓	✗
[3]			✓	✗
[4]			✓	✗
[5]			✓	✗
[6]			✓	✗
[7]			✓	✗
[8]		✓		✗
[9]			✓	✓
[10]		✓		✗

results obtained in the contexts of hand and emotion analysis, works [2, 3, 7] limited to object recognition without a semantic sense. They could have thus extended their functionality to meet extra information of the people identified based on semantic models. Meanwhile, works by [5, 6] considered the geometric features of objects as discriminant classification, while [4] relied on their color. However, none of these studies managed to provide a semantic meaning to those elements. Therefore, it is crucial to go beyond object recognition and expand the functionality of models, so they can describe a digital scene just as people perceive it. Similar to GEODIM, research by [9] relied on object recognition and used ontologies. However, the work made inferences by analyzing the context of the scenes supported by a knowledge base, such as: How long should a table measure so it can be considered as such? This type of knowledge is not applicable to GEODIM system, since tubes and objects from industrial scenes largely vary in size, although their shape is consistent. Therefore, identifying cylinders, tori, and spheres is valuable. GEODIM actually applies semantic rules to validate and correct the classification of objects and creates their topology. Furthermore, it is possible to extend its semantic meaning to connect to external ontologies.

7.2.2 Semantic on Object Recognition

As previously mentioned, GEODIM seeks to extend the object recognition process to enrich data obtained semantically in order to ease the inventory making process in factories. From this perspective, ten relevant works were found in the literature review within the scope of semantic technologies applied to 3D recognition.

First, authors in [12] proposed a feedback algorithm based on supervised feature extraction techniques; the algorithm used relevant feedback to retrieve semantically-similar objects. Also, a visual system was introduced by [13] to identify objects based on their functionality in an unknown terrain. Moreover, within the area of urban environments, a system for recognizing objects in 3D point clouds was described by [14]. The system recognized small objects in city scans.

A system for building object maps of indoor household environments was also developed by authors in [15], and the system relied on techniques for statistical analysis and feature extraction. Similarly, authors in [16] considered the spatial relationships between geometric primitives as part of the definition of the object and used ontological reasoning to solve the classification through SWRL rules. Likewise, a dataset called IAIR-CarPed was introduced by [17], which is the fine-grained and layered object recognition dataset with human annotations.

Recently, researchers in [18] developed a knowledge-based detection approach of 3D objects using the OWL ontology language known as WiDOP, which used VRML language to define the ontology of an indexed scene. In addition [19], introduced a framework that recognized materials on the surface of an object. The framework aimed to resolve the multi-label material recognition issue by exploiting object information. In that same year, a semantic mapping approach called

ASCCbot was also proposed by [20]. It relied on human activity recognition in a human–robot coexisting environment and enabled to create metric maps. Finally, researchers in [21] focused on the development of a framework that used top-down information to estimate the 3D indoor layout from a single image. The framework employed a semantic segmentation feature and an orientation map.

Table 7.2 compares algorithms of those studies that used semantics in object recognition. Although papers mentioned relied on semantics in different contexts, they limited themselves to labeling. On the other hand, GEODIM has the ability to provide semantic meaning to objects through the use of ontologies, so it is also possible to make inferences, obtain valuable information from the proposed model, or even go further, since the ontology of GEODIM can be linked to external sources of semantic information. On the one hand, Table 7.2 shows that, while studies from [12, 14, 15, 19] managed to identify elements within scanned scenes, they did not provide a semantic meaning to the segments. Consequently, they were unable to make inferences between elements or additional relevant information. On the other hand, not only does GEODIM make annotations to objects by means of tags, but it also provides these objects with a semantic meaning and describes their main features according to their type. From another perspective, research in [13, 20] created representation of spatial relations and maps by human-user and human–robot interactions, respectively. Likewise, GEODIM creates spatial relations of the objects by using Jena rules in order to define the topology of the elements of the scene. Moreover, its functionality is extended when the system adds extra information to objects in order to enrich the scanned scenes. The study by [21] was a much closer approach to our work, although the difference is that GEODIM analyzes scenes formed by clouds of points and does not train the model to classify. GEODIM also calculates the geometric characteristics of the objects, which allows it to migrate to other scenarios without needing so many changes in the core. The semantics is applied to infer topological position of the elements and correct classification problems. Moreover, the system provides the user with a better idea of the actual scene by having it fully available, instead of projecting a simple

Table 7.2 Comparison of semantic recognition algorithms

Authors	Type of analysis		Semantic by Tagging
	Recognition	Detect	
[12]	✓		✓
[13]	✗	✓	✓
[14]	✓		✓
[15]	✓		✓
[16]	✓		✓
[17]	✓		✓
[18]	✗	✓	✓
[19]	✓		✓
[20]	✗	✓	✓
[21]	✓		✓

incomplete image. With all this in mind, GEODIM innovates in terms of extending the operation of recognition algorithms reported in the literature by using the semantics of objects to improve their classification within a 3D point cloud.

This paper presents an extension of a conceptual model for the representation of digital mockups. The proposed model serves as a basis for the exchange of logical, physical, and semantic information of objects through a method of inverse engineering and by applying semantic technology and calculating spatial relationships. The model is able to recognize complex shapes from a 3D point cloud obtained from real objects in factories.

7.3 GEODIM Overview

GEODIM is an algorithm able to enrich models of indoor scenes of factories with logical, physical, and semantic information of the objects involved and by means of two processes: recognition of geometric primitives and semantic enrichment. These processes depicted in Fig. 7.1 are supported by a semantic model that contains all the information obtained through the execution of the algorithm. These processes and the model are described in the sections below.

Figure 7.2 depicts the workflow of GEODIM system, which starts by (1) receiving a point cloud of a real scene from a laser scanner. The process of geometric primitives recognition (:PrimitivesRecognition) analyzes the point cloud and then segments it in order to generate a simple classification of its elements (without semantic sense) and create a list of segments (2). From that point on, the semantic enrichment process calculates (3) specific properties according to their shape in (:GeometricFeatures), such as the trajectory, length, and diameter. Then, the list of (4) segments and their geometric features is sent to (:Topology) and (5) the spatial relationship of the segments are calculated by applying sematic rules. Afterward, the segments with their topology calculated by semantic rules are (6) sent to (:SemanticValidation). This task avoids issues of over- and under-segmentation of building primitives by (7) joining some segments or deleting others, so it is necessary to complete (8) the list of segments with spatial relationships and (9) recalculate in (:GeometricFeatures) the geometric features of the new segments formed. The result (10) is validated in (:ValidationByExperts), where expert users can (11) modify relationships as they deem appropriate. The list of modified-by-the-expert-user segments (12) is sent back to (:ValidationByExperts). Finally, all this calculated data are handled by (13) the Ontology Manager, which populates (14) the proposed ontology. The result would thus be a logical, physical, and semantic representation of a digital mockup obtained from real objects within an industrial environment. A more comprehensive description of the processes of the GEODIM algorithm is provided in the following section.
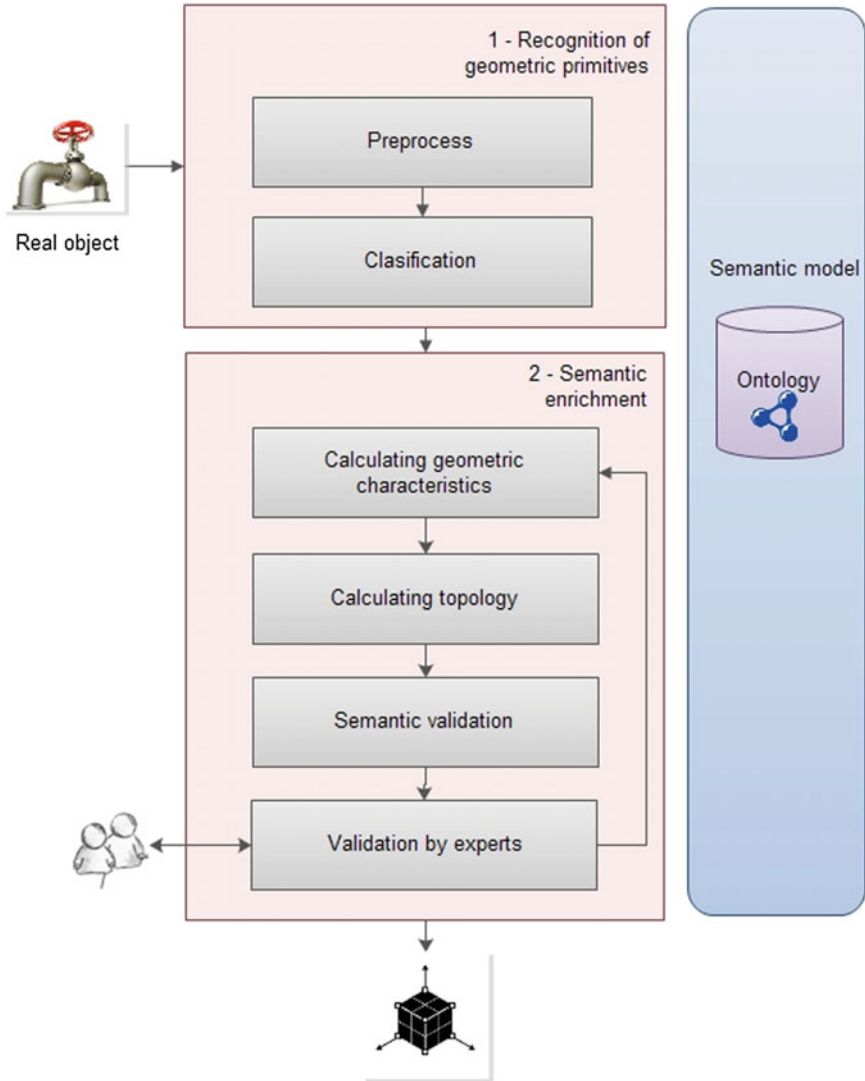
Fig. 7.1 Architecture of GEODIM

7.3.1   Process of Geometric Primitives Recognition

This process segments the point cloud according to its geometric characteristics. The entry of this process is the complete point cloud and the output is a list of objects segmented according to their geometric features. The geometric primitives analyzed by this process are cylinders, tori, and spheres.
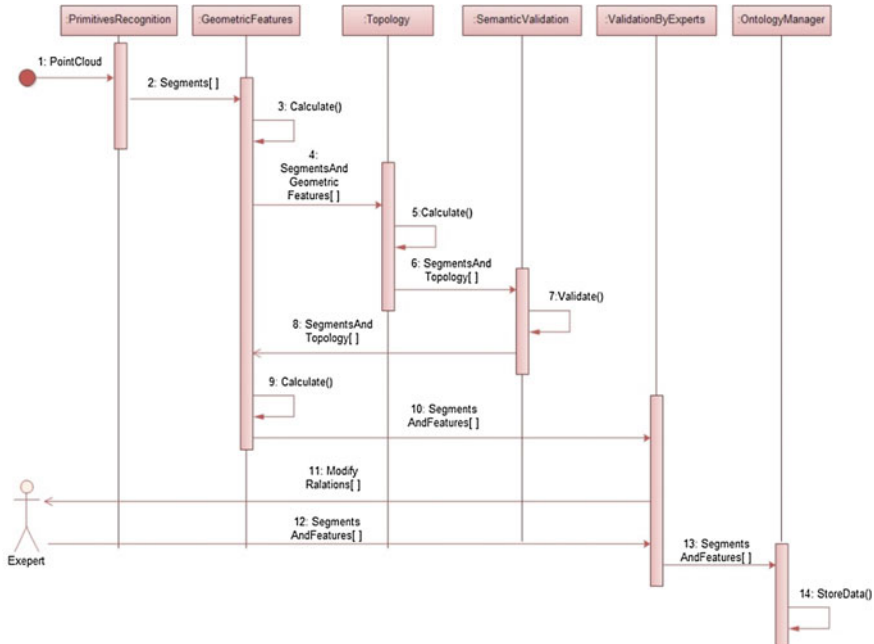
Fig. 7.2 Workflow of GEODIM

The semi-automatic process of recognizing geometric primitives in GEODIM was inspired by the work of authors in [22]. However, certain characteristics were adapted to the needs of this new system. GEODIM is categorized as a semi-automatic model, because it suggests a classification of objects into the point cloud, but this classification can be modified according to the knowledge of experts. The process of geometric primitives recognition is therefore composed of two sub-processes, Preprocess and Classification, which are described below.

a. **Preprocess**. It seeks to reduce the information to improve the quality of the original scene [23]. In GEODIM duplicated information was removed and points were ordered with a space-partitioning data structure, i.e., a Kd-tree structure. Normal values were also calculated to support the Classification sub-process.

b. **Classification**. It converts raw data into meaningful, useful, and understandable information [24]. Every object in GEODIM is classified with its corresponding geometric primitive according to its features, which are in turn obtained by calculating the percentage of fit of each primitive (cylinder, tori, sphere, and plane). These percentages are evaluated and the primitive type with the highest percentage of fit is selected and assigned to its corresponding object. The combination of geometric primitives allows for the generation of complex geometric shapes that describe real-world objects.

Since this research centers on industrial environments, every primitive has been assigned to a real element according to their similarity by semantic rules, which are described in the next section. Up to this point, GEODIM has a cloud of segmented points, and each segment is assigned to a class of primitive. In other words, GEODIM has created the logical representation of the scene. This list of segmented objects is the entry for the next process.

7.3.2 Semantic Enrichment Process

As entry, this process counts on a list of objects segmented according to their geometric characteristics, and as output, it provides with a logical, physical, and semantic representation of a digital mockup. To achieve this, GEODIM calculates the geometric properties of objects and their topologies, but it also semantically verifies the possible issues with over- and under-segmentation of building primitives by applying semantic rules.

Calculating geometric features and spatial relationships allows for the description of objects in a logical, physical, and semantic form. On the one hand, geometric features such as height, width, perimeter, and radius are considered logical information to GEODIM. On the other hand, spatial relationships and properties such as position, size, or number of points are viewed as physical information. Finally, semantic information for GEODIM involves matching every object with its corresponding element in the ontology proposed into the semantic model. However, merely the first two (logical and physical) tasks are calculated in the semantic enrichment process, while semantic information is described in "The topology of objects" [26].

a. **Calculating geometric characteristics**. This sub-process calculates specific geometric properties (trajectory, length, diameter, height, radius, etc.) according to the type of geometric primitive (cylinder, plane, torus, and sphere). All this information is stored based on the definition of the proposed ontology. Table 7.3 shows the geometric properties selected for every type of geometric primitive.
b. **Calculating topology**. Topology is a mathematical model used to define the location of and relationships between entities [25]. A topological 3D model should always be designed for specific requirements according to the application, due to its complexity and variation [26]. For this reason, GEODIM describes the

Table 7.3 Geometric properties allowed for every geometric primitive

Primitive	Number of points	Length	Height	Radius	Diameter	Minor radius	Major radius
Cylinder	✓	✓	✓	✓	✓		
Torus	✓	✓				✓	✓
Sphere	✓			✓	✓		
Plane	✓	✓	✓				

topology of objects by means of two concepts: spatial representation and spatial relationships via the design of semantic rules, depending on aspects such as the type of object and relationships permitted in the real world. To accomplish all assertions and constrains, different semantic rules were created and organized into three groups: (I) Spatial Relationship rules, (II) Spatial Representation rules, and (III) Redundant Shapes rules and Union of Shapes rules.

I. **Spatial representations**. In this case, "Meet, Overlap, Equal, Inside, and Contains" are the five spatial representations for two simple 3D objects without embedded holes selected for GEODIM. These are the most common relationships in industrial scenes. Table 7.4 shows the spatial representations that may exist in industrial scenes. For instance, to GEODIM spatial relationships between two pipes can only be of type Overlap, Equal, Inside, and Contains. However, the same pipe may have Overlap and Meet spatial representations with an elbow.

The assertions and constraints that describe the spatial relationships are verified in GEODIM by semantic rules, which were created and added to the semantic model. There is a rule for every combination according to the objects participating in the relationship. An example of this type of rules is described below.

A spatial representation Meet or Overlap of a pipe is valid if—and only if—the pipe has free connections, the Relatum is an elbow, a pipe, or a tee, and if there is a spatial relationship type Meet between them of at least 10% of their points.

SRP-Pipe-Elbow:

$$(?i\ rdf:type\ ont:SpatialRelation)\ \wedge\ (?i\ ont:has\_a\_relatum\ ?rel)\ \wedge$$
$$(?i\ ont:has\_a\_referent\ ?ref)\ \wedge\ (?rel\ rdf:type\ ont:Elbow)\ \wedge$$
$$(?ref\ rdf:type\ ont:Pipe)\ \wedge\ (?ref\ ont:number\_connections\ ?num)\ \wedge$$
$$lessThan(?num;\ 3)\ \wedge\ (?i\ ont:hassome\ ?adj)\ \wedge$$
$$(?adj\ rdf:type\ ont:Total\_Adjacency)\ \wedge\ (?adj\ ont:percentage\ ?per)\ \wedge$$
$$greaterThan(?per;\ 10)\ \wedge\ (?i\ ont:has\_some\ ?adjs)\ \wedge$$
$$(?adjs\ rdf:type\ ont:Side\_Adjacency)\ \wedge\ (?adjs\ ont:isValid\ ont:TRUE)$$
$$\rightarrow (?adj\ ont:isValid\ ont:TRUE)$$

Table 7.4 Spatial representations allowed for each entity

Object	Pipe	Elbow	Tee	Valve	Plane
Pipe	O, E, I, C	O, M	O, M		O, M
Elbow	O, M	E, I, C			
Tee	O, M		E, I, C	O, M	
Valve			O, M	E	
Plane	O, M				E

O Overlap, *E* Equal, *I* Inside, *C* Contains, *M* Meet

SRP-Pipe-Elbow rule describes the conditions for the relationship between a pipe and an elbow. The rule verifies conditions about spatial relationship, adjacency, and the adjacency type of this relationship. In this case, an elbow is necessary as Relatum and a pipe as reference. If the pipe has free connections, 10% or more should be the percentage of total adjacency in this relationship and a valid one.

II. **Spatial relationships**. Spatial relationships between objects are described according to the position of the first relative object called *Referent* towards intrinsic orientation of another object called *Relatum*. In a relative reference system, the relative position of a *Referent* toward its *Relatum* is described from the point of view of a third position called the *Origin* [13]. That is, the relative spatial position of an object depends on the viewing angle from which it is observed and its sensitivity to the rotation angle of the figure. Hence, for this model the *Origin* is equal to the position of the laser with respect to the scene. In GEODIM, the spatial relationships of objects are described by "Front," "Back," "Left," "Right," "Above," and "Below," a reference system of projective-relative relationships [27–30]. From this perspective, and considering that GEODIM is focused on the analysis of industrial scenes, elements to recognize have been limited to pipes, elbows, tees, and valves. Therefore, not all spatial relationships are permitted. Table 7.5 introduces the possible relationships for every *Referent* object according to its real features with its expected *Relatum*. For instance, a tube may have only two connections, which can be elbows or tees, while a valve has only one connection, since valves are usually only connected to tees.

Different spatial relationship rules were created, which are responsible for constructing spatial relationships between two objects called A and B. The following example shows the semantic rule for the relationship LEFT, considering that the remaining spatial relationships have similar behaviors.

Table 7.5 Spatial relationships for GEODIM

Referent	Possible relations	*Relatum* expected
Pipe	2	Elbow, Tee, or Plane
Elbow	2	Pipe
Tee	3	2 Pipes and 1 valve, or 3 pipes
Valve	1	Tee
Plane	n	Pipe, or Plane

A connection type A Left B is valid if—and only if—the left points of Shape A are at least 90% close to the right points of Shape B.

SRL-Left:

$$(?i\ ont:percentage\ ?x)\ \wedge\ greaterThan(?x;\ 90)\ \wedge$$

$$(?i\ ont:refers\_to\_side\ ont:SIDE\_LEFT)\ \wedge\ (?i\ rdf:type\ ont:Side\_Adjacency)$$

$$\rightarrow (?i\ ont:isValid\ ont:TRUE)$$

SRL-Left rule describes the conditions for the relationship between a pipe and a tee. The rule verifies conditions about the adjacency of the relationship; 90% or more should be the percentage of adjacency in this relationship and an adjacency type = LEFT.

III. **Semantic validation**. Results from the recognition process of geometric primitives sometimes contain some over- and under-segmentation problems. GEODIM tries to avoid these issues through two types of semantic rules: Redundant Shapes and Union of Shapes rules. The former prevent over-segmentation by identifying the elements with the issue and joining them together as a single segment. The latter focuses on the same issue but under different conditions; it seeks the over-segmentation of objects that compose a real element. GEODIM has a Redundant Shapes rule for every type of combination of elements. The rule for two elbows is showed below as an example.

Combine segments A and B if—and only if—the Referent A and the Relatum B are of the same type, their centroids are at least 95% close, and all points are at least 85% close. That is, if A Equals B, A Inside B, or A Contains B.

RS-Elbow-Elbow:

$$(?i\ rdf:type\ ont:SpatialRelation)\ \wedge\ (?i\ ont:has\ a\ relatum\ ?rel)\ \wedge$$

$$(?i\ ont:has\ a\ referent\ ?ref)\ \wedge\ (?rel\ rdf:type\ ont:Elbow)\ \wedge$$

$$(?ref\ rdf:type\ ont:Elbow)\ \wedge\ (?i\ ont:has\ some\ ?adj)\ \wedge$$

$$(?adj\ rdf:type\ ont:TotalAdjacency)\ \wedge\ (?adj\ ont:percentage\ ?per)\ \wedge$$

$$greaterThan(?per;\ 85)\ \wedge\ (?rel\ ont:has\ some\ ?cd)\ \wedge$$

$$(?cd\ rdf:type\ ont:CentroidDistance)\ \wedge\ (?cd\ ont:percentage\ ?cdP\ erc)\ \wedge$$

$$greaterThan(?cdP\ erc;\ 95)$$

$$\rightarrow (?adj\ ont:isV\ alid\ ont:TRUE)$$

The RS-Elbow-Elbow rule describes the conditions for the relationship between two redundant elbows. The rule verifies conditions about spatial relationship, adjacency, and the centroids of the objects. Two Elbows as Relatum and as Reference are verified, 85% or more should be the percentage of total adjacency in this relationship, and the centroids between the two elbows should be 95% close.

Over-segmentation of segments belonging to the same cylinder has been observed as a result of primitives recognition algorithms. As previously stated,

GEODIM seeks to abolish these incorrectly classified segments and join the points in a single segment by means of Union Shapes rules:

Join segments if – and only if – the Referent A and the Relatum B are pipes and have a spatial representation Meet or Overlap.

US-Pipe-Pipe:

(?i rdf : type ont : SpatialRelation) ^ (?i ont : has a relatum ?rel) ^

(?i ont : has a referent ?ref) ^ (?rel rdf : type ont : Pipe) ^

(?ref rdf : type ont : Pipe) ^ (?i ont : hassome ?adj) ^

(?adj rdf : type ont : Total Adjacency) ^ (?adj ont : percentage ?per) ^

greaterThan(?per; 10) ^ (?i ont : has some ?adjs) ^

(?adjs rdf : type ont : Side Adjacency) ^ (?adjs ont : isV alid ont : TRUE)

→ (?adj ont : isV alid ont : TRUE)

The US-Pipe-Pipe rule shows the conditions for the relationship between two connected pipes. The rule verifies conditions of spatial relationship, adjacency, and adjacency side. Both the Relatum and the Referent are pipes, and 10% or more should be both the percentage of total adjacency in this relationship and a valid adjacency percentage.

c. **Validation by experts**. GEODIM is a semi-automatic model that suggests a classification of objects into the point cloud. In this sub-process an expert can modify items according to his/her knowledge. The expert can thus reclassify objects, delete them, or modify their properties. This sub-process was considered to support the quality of data classification.

7.3.3   Semantic Model

Semantic modeling helps define data in entities and their relationships. The set of entities of the proposed semantic model includes a taxonomy of classes used to support GEODIM in the representation of a digital scene as it is perceived in the real world. This approach is represented by an ontology. Figure 7.3 shows the semantic model proposed by GEODIM, which faithfully represents the behavior of the entities and their relationships in a real industrial scene, including assertions and limitations.

Table 7.6 describes the main elements of the proposed ontology.

The combination of geometric primitives allows for the generation of complex geometric figures that describe real-world objects. This ontology defines the objects (and their topology) recognized by the proposed model and it can link to external ontologies that will enrich the semantic meaning of such items.
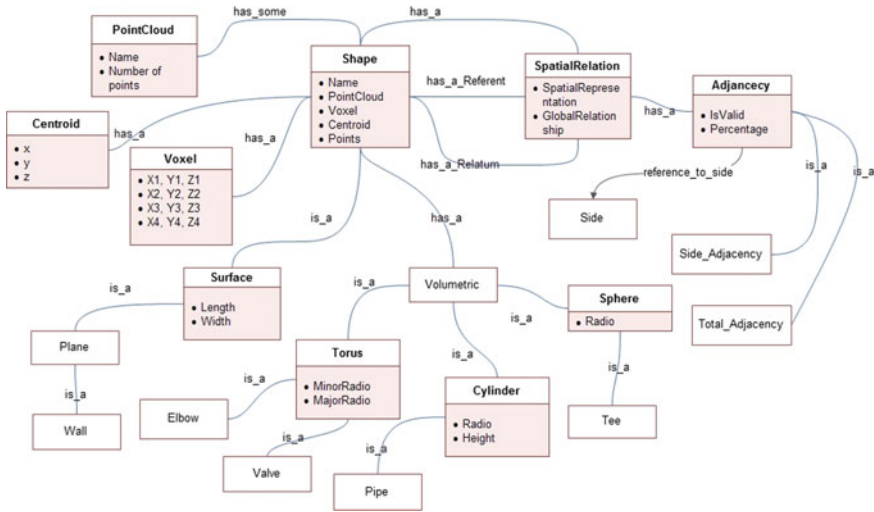
Fig. 7.3 GEODIM ontology

7.4 A Real Use Case: Objects Recognition in an Industrial Facility

This section describes the real use case that shows the functionality of GEODIM by explaining the stages of 3D reconstruction process. The problem is described below:

A great number of industrial facilities nowadays were built during the decades of 1960, 1970, or even 1980. Therefore, their plans have always been available in 2D, which makes them outdated and obsolete for today's needs due to continuous modifications, additions, or alterations that those establishments have suffered. As a result, the maintenance, reparation, and expansion of these places may become an extremely expensive, laborious, and—to a great extent—dangerous task. Since it is impossible to be completely familiar with the exact type, number, location, or dimension of objects, foremen and/or builders usually must leave the field and observe. This delays works and implies extra working hours. However, the solution to this issue is scanning inside the industrial facilities and recognizing existing objects with the help of GEODIM.

GEODIM was therefore employed in the use case for the creation of a digital 3D mockup from a scanned 3D point cloud of an industrial facility of medium-large size. Figure 7.4 represents the 360° view of a section of the industrial facility, which has 2,884,079 points. First, the shapes that GEODIM had to recognize for this studio were pipes and planes. Such information was useful to create more intelligent objects and enrich the 3D digital mockup.

The two main process of GEODIM, primitives recognition and semantic enrichment, are now described:

Table 7.6 Description of ontology

Entity	Description
PointCloud	The *PointCloud* object represents a set of points in a three-dimensional coordinate system, which describes a real-world scene
Centroid	It describes a point in the three-dimensional coordinate system that is located just at the center of an object
Voxel	It represents a grid in 3D space on the cloud of points. A Voxel object delimits, under spatial boundaries, a Shape belonging to the point cloud
Shape	It is a three-dimensional geometric object that occupies a place in space. A shape is classified as *Surface* or *Volumetric*. Every Shape element belongs to a particular *CloudPoint*
SpatialRelation	It describes spatial interactions between a *Referent* and a *Relatum*
Surface	It is a type of Shape describing an object that has length and width
Plane	It is a surface composed of points following a same direction. It is easy to find planar objects in the real world. Examples are tables, ceilings, walls, sidewalks, streets, and paper sheets
Volumetric	It describes a three-dimensional geometric Shape that has volume. A *Volumetric* element can be of type *Sphere*, *Torus*, *Cone*, or *Cylinder*
Sphere	It describes all elements having a spherical surface. Spheres can be found in both artificial environments and nature. A ball, a lemon, a doorknob, and a light bulb are examples of spheres
Torus	It is an element of the ontology describing a small circle that turns along a line drawn by another circle. Many real-life objects have a torus, such as donuts, a Roman cushion, a tire, or a light bulb
Cylinder	It describes elements formed by the simplest curves created by moving a straight line around the circumference of a circle. It is possible to find a cylinder in the real world in form of bottles, soda cans, columns, and pipes, for instance
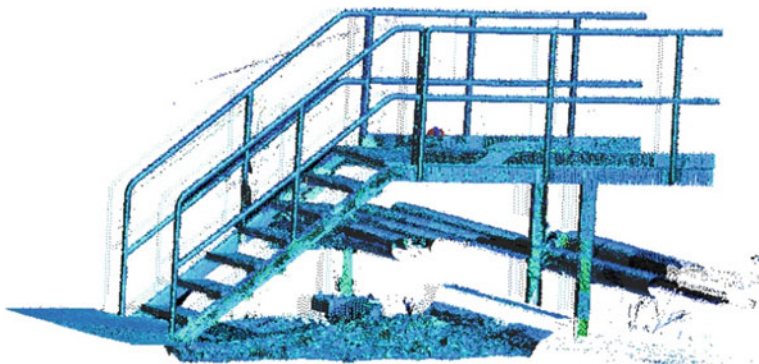
Fig. 7.4 Industrial point cloud

7.4.1 Recognition Process of Geometric Primitives

In this process the scene was segmented according to its geometric properties by the primitive recognition algorithm (detailed description of this algorithm is beyond the scope of this paper). Table 7.7 summarizes segments obtained by the algorithm.

At this stage GEODIM has a list of classified elements belonging to the real industrial scene without semantic sense. That is, GEODIM has created the logical representation of the scene. The next step is the semantic enrichment process.

7.4.2 Semantic Enrichment Process

In this process GEODIM calculated extra information into describe the industrial scene with logical, physical, and semantic information by applying the following four sub-processes:

(a) **Calculating geometric characteristics**. GEODIM calculated specific geometric properties for every element of the list previously obtained. The system calculated the properties according to restrictions showed in Table 7.3, and it generated a list of elements classified with logical information.

(b) **Calculating topology**. The topology in GEODIM is defined by calculating spatial relationship, spatial representation, and a semantic validation of elements that belong to the point cloud by using semantic rules. The spatial relationship rules were applied for this use case to clarify the obtained result. An example of a created spatial relationships is presented: the relationships of plane_006 are detailed by a graphic and a scheme. In the real world, plane_006 represented the floor of the industrial walkway surrounded by two handrails; tubes represented the basis of these handrails. Figure 7.5 shows six spatial relationships created for plane_006 type ABOVE with 6 pipes: pipe_150, pipe_153, pipe_066, pipe_056, pipe_035, and pipe_033.

Next, the spatial representations (Meet, Overlap, Equal, Inside, and Contains) of elements were calculated by applying Spatial Representations rules. A visual example of the created spatial relationship is shown in Fig. 7.6, where a real industrial walkway is shown. This walkway was composed of different objects, six

Table 7.7 Classified segments

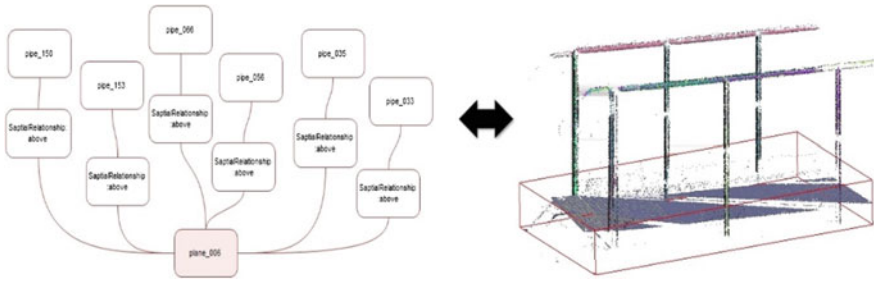| Size (number of points)	Primitive	Total segments	Segments classified correctly	Success rate (%)
2,884,079	Cylinder	112	103	86
	Planes	49	8	88
	Tori	0	0	0
	Total	161	111	68

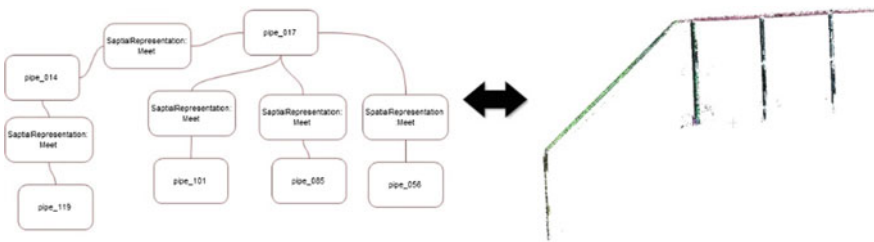**Fig. 7.5** Result of spatial relationships

Fig. 7.6 Result of spatial representations

of which are depicted. Therefore, pipe_017 has four spatial representations type MEET with pipe_101, pipe_085, pipe_056, and with pipe_014, which in turn has a spatial representation with pipe_119.

Semantic validation tries to avoid issues in the process of geometric primitives recognition in terms of over- and under-segmentation. Figure 7.7 shows a sample of over-segmentation, where pipe_017 and pipe_075 stood for the same real object and had a spatial representation type EQUAL. Once the Semantic Validation rules were applied, the elements with a spatial representation type EQUAL were joined, creating only one element. In this example, the brand new element created was called pipe_u017.

Up to this moment, GEODIM always has an industrial scene with an extra definition of logical, physical, and semantic information. However, certain errors of classification or semantics may still exist. An expert validation is thus needed.

(a) **Validation by experts**. Expert users checked the elements in order to verify the classifications, properties, and the topology. They could correct any faults within the semantic model according to their experience. Figure 7.8 shows a classification error modified by the expert in the plane classification.

After all corrections, the semantic enrichment process was repeated until information was correct according to the viewpoint of experts. The final result of GEODIM was an industrial scene described with logical, physical, and semantic
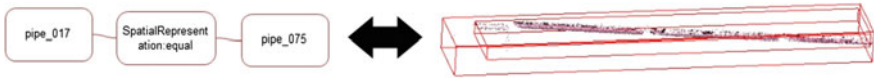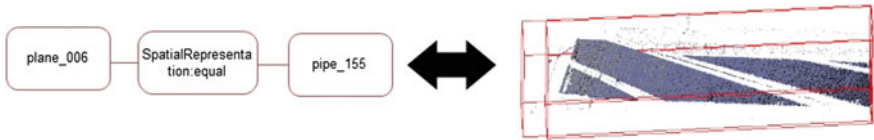
Fig. 7.7 Result from semantic validation

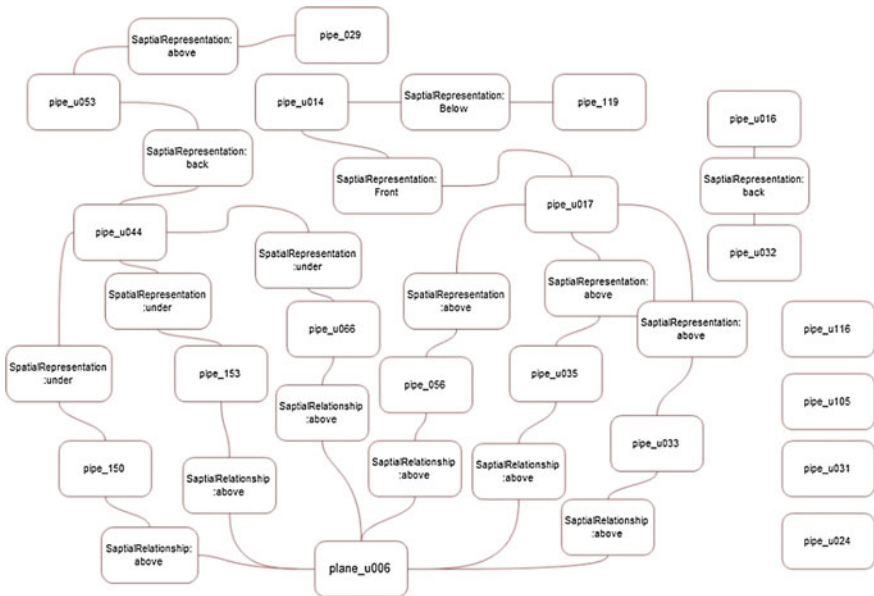**Fig. 7.8** Error recognition process

Fig. 7.9 Portion of the final result with GEODIM implementation

information. This model enabled the description of all objects into the industrial scene via semantic modeling. Figure 7.9 shows a portion of the final result after the use of GEODIM. The end result of GEODIM was the logical, physical, and semantic description of the elements that comprised the scanned point cloud. All this information allowed for the creation of the industrial inventory required by the factory.

7.5   Evaluation

Evaluating a geometric primitives recognition system is not an easy task. Literature carried out for this research managed to describe and carry out different types of evaluations, although these cannot always be useful for other systems. Algorithms evaluation actually varies depending on diverse factors, such as the amount of points in the point cloud, density, or the type of elements involved. However, from a general perspective, two different methods are possible to assess software and tools:

(a) Quantitative evaluation methods. They are based on the assumption that the software product has at least a measurable property that can change as a result of using the methods/tools to be evaluated. Quantitative evaluations can be developed in three different ways: case studies, formal experiments, and surveys. For instance, authors in [31] introduced experiments to validate the effectiveness of their proposed QDFT descriptor under geometric transformation. Similarly, the one-versus-rest multiclass classification experiments was addressed by [32], where true and false samples were labeled differently. Authors used accuracy rates to evaluate the performance of different methods. Furthermore [33], proposed a performance comparison of their work and presented seven other studies. The comparison aimed at measuring average precision. Finally, authors in [9] discussed the success of object recognition of the research, and different aspects were measured, such as the inclusion of contextual information about the objects, their geometric and appearance features, and their classification based on their type. Metrics used were precision and recall.

(b) Qualitative methods. The term Feature Analysis is used in literature to describe a qualitative evaluation, which is based on (1) identifying the requirements that users have for a particular task or activity and (2) mapping those requirements to features that a method/tool aimed at supporting that task/activity should possess. An example of this qualitative assessment is the qualitative comparison provided by [34], where two methods were analyzed. In total, 31 clusters were extracted out of the computed hierarchy for a particular model. Then, results from the two methods were compared to determine which one managed to generate the best stylized model.

In the end, a quantitative approach was selected to evaluate the quality of classification of GEODIM. Measurements employed were: Segments Classified and Segments Classified Correctly; i.e., true positive (TP), false positive (FP), false negative (FN), precision (P), and recall (R). Two evaluations were proposed. The former was used for the classification results in the process of geometric primitives recognition, whilst the latter was carried out for the final results of GEODIM in order to compare results and demonstrate how GEODIM improved the traditional classification in an industrial point cloud.

Segments Classified referred to the number of segmented elements belonging to the original point cloud, while Segments Classified Correctly or TP concerned those segments correctly classified according to their geometric properties. Similarly, FP stood for segments incorrectly classified, whilst FN comprised segments not classified based on their corresponding geometric primitive. Also, in this case P and R were defined as a set of classified segments and a set of relevant segments. That is, P referred to the fraction of classified segments relevant to the process, while R indicated the fraction of the segments relevant to the process but which were also successfully classified.

$$P = TP/(TP + FP). \tag{7.1}$$

$$R = TP/(TP + FN) \tag{7.2}$$

Table 7.8 shows the result from the process of geometric primitives recognition. The point cloud was segmented in 161 elements, most of which were cylinders and planes, in addition to some tori elements. The table shows low precision and recall metrics, since the algorithm misclassified 32 elements as tori. In the case of cylinders, precision P was 0.86, whilst planes had a precision of 0.88, although its recall measure was low due to 41 elements that were incorrectly classified. In total, this method had a precision of 0.69 and a recall of 0.68.

The implementation of the semantic enrichment process of GEODIM improved these results, such as Table 7.9 shows. It can be observed from the new table that precision and recall for cylinders and planes became 100%. However, certain

Table 7.8 Results from object recognition process

S	Primitive	TS	TP	FN	FP	P	R
2,884,079	Cylinders	112	103	9	16	0.86	0.92
	Planes	49	8	41	1	0.88	0.16
	Tori	0	0	0	32	0	0
Total			161	111	50	49	0.69	0.68

S Size of point cloud, *TS* Total of segments, *TP* True Positive, *FP* False Positive, *P* Precision, *R* Recall

Table 7.9 Results from semantic enrichment process

S	Primitive	TS	TP	FN	FP	P	R
2,884,079	Cylinders	84	84	0	0	1	1
	Planes	41	41	0	0	1	1
	Tori	0	0	0	18	0	0
Total			125	125	0	18	0.87	1

S Size of point cloud, *TS* Total of segments, *TP* True Positive, *FP* False Positive, *P* Precision, *R* Recall

elements were still misclassified, since they were sets of points that did not fit any geometric primitive analyzed. This can be considered as lost information, and, for this reason, precision in the process was 87%. The table also shows that the total number of segments decreased due to the semantic validation process, wherein the over-segmentation was solved.

7.6 Conclusions and Future Work

This research introduced GEODIM, a semantic model-based system for the recognition of industrial scenes that creates a conceptual model for semantic representation of digital mockups (i.e., a semi-automatic inventory). GEODIM enables users to enrich models of indoor scenes of factories with logical, physical, and semantic information by using a semantic model and applying two processes: geometric primitives recognition and semantic enrichment.

Geometric primitives recognition actually classifies the elements of the scanned scene according to their geometric characteristics, while the semantic enrichment process calculates valuable information for every item, including their topology, by using semantic rules. These rules describe the assertions and restrictions of behavior in the real world of industrial elements used (pipes, planes, elbows, and valves). To validate the functionality of GEODIM, the real use case of an industrial facility was thus introduced, and the research also addressed two quantitative evaluations in order to show the classification quality of GEODIM.

Results obtained showed that GEODIM does improve the classification of objects with 87% of success in this case, although elements incorrectly classified remained present due to under-segmentation issues. Future work will thus seek to include the composition of complex industrial elements based on the combination of several recognized objects in order to improve the recognition of a wide range of both simple and complex real industrial objects. Further research will also aim at including automatic element recognition to improve the quality of the classification and enrich industrial scenes. Mechanisms to avoid under-classification of elements will also be pursued.

References

1. Hofer, M., Odehnal, B., Pottmann, H., Steiner, T., Wallner, J.: 3D shape recognition and reconstruction based on line element geometry. Proc. IEEE Int. Conf. Comput. Vis. **II**, 1532–1538 (2005)
2. Bhuyan, M., Neog, D., Kar, M.: Hand pose recognition using geometric features. In: Communications (NCC), 2011, pp. 0–4 (2011)

3. El-Sayed, M., Radwan, E., Zubair, A.: Abductive neural network modeling for hand recognition using geometric features. In: Neural Information Processing, pp. 593–602 (2012)
4. Pasqualotto, G., Zanuttigh, P., Cortelazzo, G.M.: Combining color and shape descriptors for 3D model retrieval. Signal Process. Image Commun. **28**(6), 608–623 (2013)
5. Soysal, M., Alatan, A.A.: Joint utilization of local appearance and geometric invariants for 3D object recognition. Multimed. Tools Appl. **74**(8), 2611–2637 (2013)
6. Hejrati, M., Ramanan, D.: Analysis by synthesis: 3D object recognition by object reconstruction. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2449–2456 (2014)
7. Majumder, A., Behera, L., Subramanian, V.K.: Emotion recognition from geometric facial features using self-organizing map. Pattern Recogn. **47**(3), 1282–1293 (2014)
8. Junyan, L., Qingju, T., Yang, W., Yumei, L., Zhiping, Z.: Defects' geometric feature recognition based on infrared image edge detection. Infrared Phys. Technol. **67**, 387–390 (2014)
9. Ruiz-Sarmiento, J.-R., Galindo, C., Gonzalez-Jimenez, J.: Scene object recognition for mobile robots through Semantic Knowledge and Probabilistic Graphical Models. Expert Syst. Appl. **42**(22), 8805–8816 (2015)
10. Nasr, E.S.A., Khan, A.A., Alahmari, A.M., Hussein, H.M.A.: A feature recognition system using geometric reasoning. Procedia CIRP **18**, 238–243 (2014)
11. Gaither, N., Frazier, G.: Administración de producción y operaciones (2000)
12. Leifman, G., Meir, R., Tal, A.: Semantic-oriented 3d shape retrieval using relevance feedback. Vis. Comput. (2005)
13. Hois, J., Wünstel, M., Bateman, J., Röfer, T.: Dialog-based 3D-image recognition using a domain ontology. In: Spatial Cognition V Reasoning, Action, Interaction (2007)
14. Golovinskiy, A., Kim, V.G., Funkhouser, T.: Shape-based recognition of 3D point clouds in urban environments. In: 2009 IEEE 12th International Conference on Computer Vision, no. ICCV, pp. 2154–2161 (2009)
15. Rusu, R., Blodow, N.: Close-range scene segmentation and reconstruction of 3D point cloud maps for mobile manipulation in domestic environments. In: Intelligent Robots and Systems (2009)
16. Günther, M., Wiemann, T.: Model-based object recognition from 3d laser data. In: KI 2011 Advances in Artificial Intelligence (2011)
17. Wu, Y., Liu, Y., Yuan, Z., Zheng, N.: IAIR-CarPed: a psychophysically annotated dataset with fine-grained and layered semantic labels for object recognition. Pattern Recognit. Lett. **33**(2), 218–226 (2012)
18. Hmida, H., Cruz, C., Boochs, F., Nicolle, C.: Knowledge base approach for 3d objects detection in point clouds using 3d processing and specialists knowledge. arXiv Prepr. arXiv1301.4991 (2013)
19. Yang, L., Xie X.: Exploiting object semantic cues for Multi-label Material Recognition. Neurocomputing **173**, 1646–1654 (2015)
20. Sheng, W., Du, J., Cheng, Q., Li, G., Zhu, C., Liu, M., Xu, G.: Robot semantic mapping through human activity recognition: a wearable sensing and computing approach. Robot. Auton. Syst. **68**, 47–58 (2015)
21. Park, S.-J., Hong, K.-S.: Recovering an indoor 3D layout with top-down semantic segmentation from a single image. Pattern Recognit. Lett. **68**, 70–75 (2015)
22. Attene, M., Patane, G.: Hierarchical structure recovery of point-sampled surfaces. Comput. Graph. Forum **29**(6), 1905–1920 (2010)
23. Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 3rd edn. (2006)
24. Mountrakis, G., Im, J., Ogole, C.: Support vector machines in remote sensing: a review. ISPRS J. Photogramm. Remote Sens. **66**, 247–259 (2011)
25. Liu, L., Zsu, M.: Encyclopedia of Database Systems (2009)
26. Zlatanova, S., Rahman, A.A., Shi, W.: Topological models and frameworks for 3D spatial objects. Comput. Geosci. **30**(4), 419–428 (2004)

27. Moratz, R., Nebel, B., Freksa, C.: Qualitative spatial reasoning about relative position. In: Spatial Cognition III (2003)
28. Moratz, R., Tenbrink, T., Bateman, J., Fischer, K.: Spatial knowledge representation for human-robot interaction. In: Spatial Cognition III (2003)
29. Méndez, V., Rosell-Polo, J., Sanz, R.: Deciduous tree reconstruction algorithm based on cylinder fitting from mobile terrestrial laser scanned point clouds. Biosyst. Eng. **124**, 78–88 (2014)
30. Levinson, S.: Frames of reference and Molyneux's question: crosslinguistic evidence. Lang. Space (1996)
31. Li, H., Liu, Z., Huang, Y., Shi, Y.: Quaternion generic Fourier descriptor for color object recognition. Pattern Recognit. **48**(12), 3895–3903 (2015)
32. Hong, C., Yu, J., You, J., Chen, X., Tao, D.: Multi-view ensemble manifold regularization for 3D object recognition. Inf. Sci. **320**, 395–405 (2015)
33. Rubio, J.C., Eigenstetter, A., Ommer, B.: Generative regularization with latent topics for discriminative object recognition. Pattern Recognit. **48**(12), 3871–3880 (2015)
34. Attene, M., Falcidieno, B., Spagnuolo, M.: Hierarchical mesh segmentation based on fitting primitives. Vis. Comput. **22**, 181–193 (2006)