Identification and classification of objects in 3D point clouds based on a semantic concept

Jean-Jacques PONCIANO, Frank BOOCHS und Alain TREMEAU

Abstract

Our real world is increasingly subject to digitization processes producing huge unstructured 3D data sets. In order to extract objects contained in these data sets subsequent analysis steps are necessary. The most efficient but also the most expensive and time-consuming analysis is based on manual editing, which allows integrating human knowledge and intelligence. Computer based methods are less effective, as they mainly use implicit knowledge allowing to parametrize algorithms, which are part of a defined processing chain. We want to overcome these limitations through a more general and more flexible integration of any kind of useful knowledge into the processing. This paper presents an approach fully driven by semantic technologies and uses expert knowledge. This expert knowledge is modeled into an ontology and describes objects, data, and algorithms. This ontology guides an iterative reasoning process using the semantic technologies (e.g. OWL2, SPARQL, an engine reasoner) to provide a prime flexibility. This process firstly identifies relevant algorithms to parametrize and combine them. Secondly, it interprets the result provided by algorithms to enrich the knowledge base. Thirdly, a semantic reasoning uses added information to classify objects and determine, if needed, other relevant algorithms. This process is thus, dynamically and iteratively adapted to the results of executed algorithms. Its efficiency is presented through the result comparison of its results according to other possible approaches.

1 Introduction

The evolution of model acquisition techniques produces nowadays various data sets coming from different domains (e.g. cultural heritage, 3D-model of building, etc). The use of these 3D data requires to be annotated semantically on its content. To obtain such annotated datasets, the objects that compose the data have to be detected and then, annotated. The object detection process in various data sets faces to problems of a diversity of objects to be detected, whose the representation is sometimes incomplete (low density, occluded objects, etc) and composed of irregular shapes. The existing detection approaches are based on object geometry. Therefore they face problems to detect incomplete objects, Moreover, some context as cultural heritage does not allow for using machine learning approach due to the lack of annotated data to train the system. That is why we have developed a knowledge-based approach to detect objects inside 3D data. Contrary to other existing knowledge-based approach, our approach is full semantically guided by the knowledge. The knowledge base containing the description of objects, data and algorithm allow for selecting relevant algorithms according to the data and execute them. The result of an algorithm execution is semantically interpreted to enrich the knowledge base which will classify detected objects. After a classification step, the detected objects are analyzed to identify new characteristics in common to all objects of the same class. Then, it identifies if new algorithms must be executed. This iterative process provides flexibility and adaptability to detect objects according to the processed data and without being limited to a static semantic description. This process is illustrated through an application case on a 3D point cloud resulting from a LiDAR scan of a test area in Freiburg, provided by Fraunhofer IPM. Finally, this approach is compared to others.

2 Related work

With the increase of data and their diversity, some researches in computer vision have focused in the design of the automatic process of detection able to detect object inside the diversity of data or the diversity of data context. The review (Grilli et al., 2017) presents the most popular methodologies and algorithms used to segment and classify 3D point clouds. Likewise, the review (Pal and Pal, 1993) presents the most popular methodologies to segment an object in a 2D image. Two main approaches emerge from the family of data-driven methodology: machine learning methods and stochastic methods. Among the use of machine learning methods, methods based on a convolutional neural network as it is presented in (Ouyang et al., 2015) and (Wirges et al., 2018) provide an accurate object detection. Nevertheless, the use of machine learning methods is limited due to the requirement of large data sets to obtain a satisfying result. When large data sets are not available for a domain of application, a stochastic method is suitable. This method is often based on shape and feature and aims at the recognition of the context. Diverse semantic information can be provided thanks to a recognition of context as for example, a scene description (Anguelov et al., 2005) or the description of a geometry (Triebel et al., 2007). The work (Torralba et al., 2010) presents a shape-based recognition using a prestructured knowledge to identify semantic/geometric classes. In this work, semantic techniques are used to represent the content of the 3D model resulting from its process. However, a semantic technique can also be used directly in the process of recognition. The interest in ontology-based recognition process appears mainly in the domain of image processing. It is the case of the work (Hoiem et al., 2005), which presents a recognition method using a domain ontology. Combined with the use of ontologies, the techniques of reasoning intervene also, in the detection and classification of objects, as it is presented in the Ph.D. dissertation (Maillot, 2005). However, this related work shows that the use of the semantic techniques addresses only some steps of the processing. Contrary to them, our approach uses the semantic techniques at each step of the processing. Such a fully semantic based processing provides advantages as the flexibility of adaptation to guide the process of computer-based modeling (through an adaptive selection of algorithms and an iterative classification) and provide an annotated 3D model understandable by machines and humans.

3 Approach

The approach presented is based on semantic modeling and its dynamic enrichment. The dynamic evolution of the semantic model offers adaptability and flexibility from a given semantic model sufficient for the detection of objects whose description derives from their representation within the data. The dynamic enrichment of the semantic modeling aims at improving the object description by an iterative process of object detection and feature analysis of well-detected objects to deduce new relevant features of objects. Thus, semantic reasoning can determine other algorithms to detect objects.

3.1 Knowledge base

An OWL-based ontology (Ponciano et al., 2017) has been developed to support a generic process that is adaptable according to the knowledge about the scene, the objects inside, and algorithms to detect and classify scenes and objects. This ontology contains connected expert knowledge to all relevant domains (object, data, algorithms, scene,...) organized in a hierarchical structure allowing to go from most common and generic information to more specialized content. The connection inside the ontology is done through properties as illustrated in figure 1.

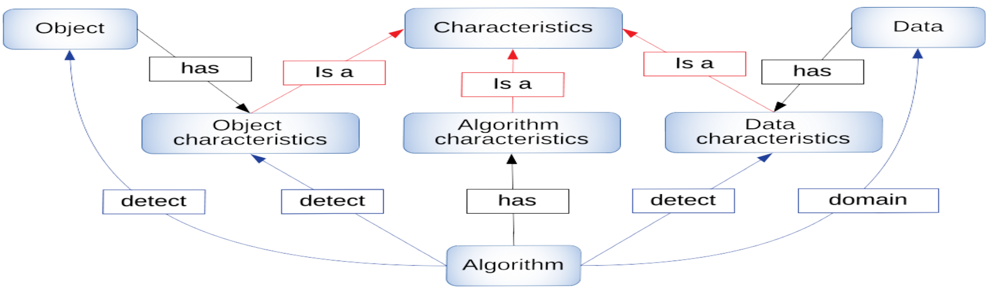

Abb.1: Overview of the knowledge base

Algorithms are thus, defined by their inputs, parameters and outputs. Characteristics describing an object can be geometric or topological according to other objects. Data characteristics depend on their acquisition technologies and qualified their quality (e.g. density, noise, etc).

3.2 Overall process

The object detection process begins by preprocessing tasks like the computation of octree to improve the efficiency of algorithms execution. A priority of detection is then computed for each object described into the knowledge base. This priority assignment is based on the topological relationship of object and defined according to the object size in decreasing order.

This priority order between objects is used to determine the data segmentation through the dynamic selection of algorithms. The selection of algorithms is done step by step by processing objects according to their priority and their geometric characteristics. This process results in an automatic segmentation of the data.

The created segments are classified by the reasoning engine using exclusively the description logic of objects to classify instances representing segments into a class of object. The classification of segments provides information about how an object is represented inside the data. Therefore, all possible features of an object are analysed and extracted from every classified segment to enrich the object description by new geometric features and topological links between other objects. The analysis is based on the occurrence of a feature for all segment classified in the same object category to deduce new relevant characteristics of the object. These new characteristics are added into the knowledge base, to enrich it and improve the detection process. Afterwards, the overall process is iterated with the new object description. Moreover, after each iteration, all points of classified segments are extracted for labelling the point cloud. Then all points labelled are removed from the point cloud in process. Thus, the labelling process allows the simplification of the object detection process by decreasing the size of the point cloud. The workflow of this iterative process is presented in figure 2.

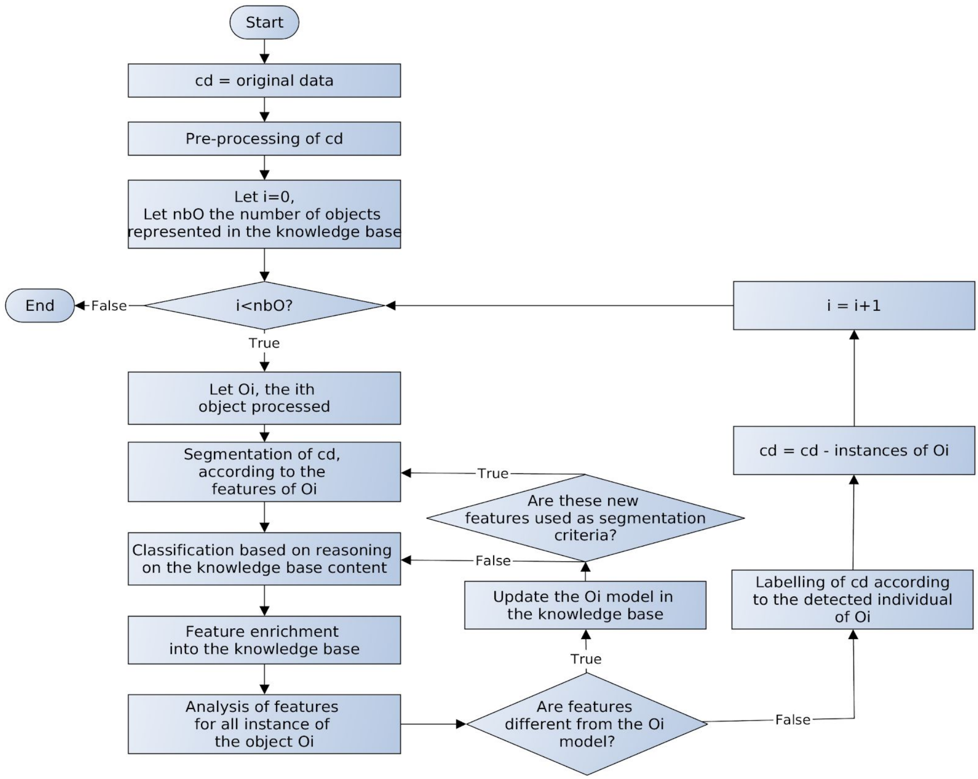

Abb. 2: Workflow of the overall process

3.3 Algorithm selection, parametrization, and execution

The algorithms selection process requires three main steps based on an analysis of the knowledge base. At first, it identifies the most relevant set of algorithms to detect each object. Secondly, it identifies algorithms that are relevant to the data. Finally, it determines algorithms required for the execution process of other algorithms as it is the case for preprocessing algorithms. All of these steps are done through reasoning on the knowledge base thanks to a set of specifications.

Moreover, algorithms parameters are automatically defined according to the characteristics of the object sought or the data used. Thanks to these definitions, the detection process is designed to determine step by step the algorithms to be executed at each step. This approach provides a detection process, which adapts itself according to the results provided by previous algorithm execution. This process to design the algorithms sequence for the detection is further explained and detailed in (Ponciano et al. 2017).

3.3.1 Application case

This approach has been applied on a point cloud (see figure 3 a) which is a LiDAR scan of a test area in Freiburg, provided by Fraunhofer IPM. This point cloud has a varying quality and it has a wide variety of representation of an object caused by irregular shapes as well as the occlusion of an element by another. These characteristics greatly alter the geometric representation of objects, making their detection a challenge.

The objects sought inside the point cloud are ground, wall, building element, tree, car and traffic sign. The size of objects and relationship to each other determines their detection priority. In this case, the ground is the biggest possible object. Moreover, cars are described as being on the ground, walls are defined as perpendicular to the ground as well as trees and traffic signs. Therefore the ground is the element whose detection is a priority. Then, walls and trees have no topological connection with other elements than the ground and the wall has a bigger maximum size than the maximum size of a tree. Therefore the next most important elements to be detected after the ground is the wall.

According to the objects detection order and the characteristics of the objects, the point cloud is firstly segmented for ground detection. As the ground is defined as the biggest horizontal plan in the scene and is composed of every element at the same altitude as this plane, the segmentation process selected begins by applying a normal estimation on the original point cloud (see figure 3a), following by a sampling of the point cloud corresponding to 1% of the minimal size of the ground. Then horizontal normal is filtered (figure 3b) and the biggest segment is classified as the ground (figure 3c). Thanks to the detection of the first ground, the altitude of every segment is computed to recursively classify segments having the same altitude of the ground as a ground (figure 3d). This analysis is performed for each new segment resulting from a segmentation designed for another object. Thus, ground detection becomes more and more accurate as the point cloud is segmented and objects are detected (figure 3e).

After the ground detection, the wall is the next objects sought. The wall detection starts at first by segment the point cloud according to the vertical plane. Thus, the point cloud is filtered with vertical normal from the point cloud without the detected ground (figure 3f) and then segmented by the orientation (figure 3g). Then, every vertical plan which are perpendicular to the ground and which satisfies the wall size constraints are classified as a wall(figure 3h). As well as for ground detection, more the point cloud is segmented and objects are detected, more the wall detection becomes accurate.

After the detection of the biggest objects, all unclassified points are gathered in another point cloud (figure 3i). Then the detection process iteratively segments the point cloud according to the geometry of other objects sought. Afterwards, segments are analysed to compute every relevant feature used to classify them. Finally, segments are classified in car, tree, building element and traffic sign by the same process as for ground and wall (figure 3j).

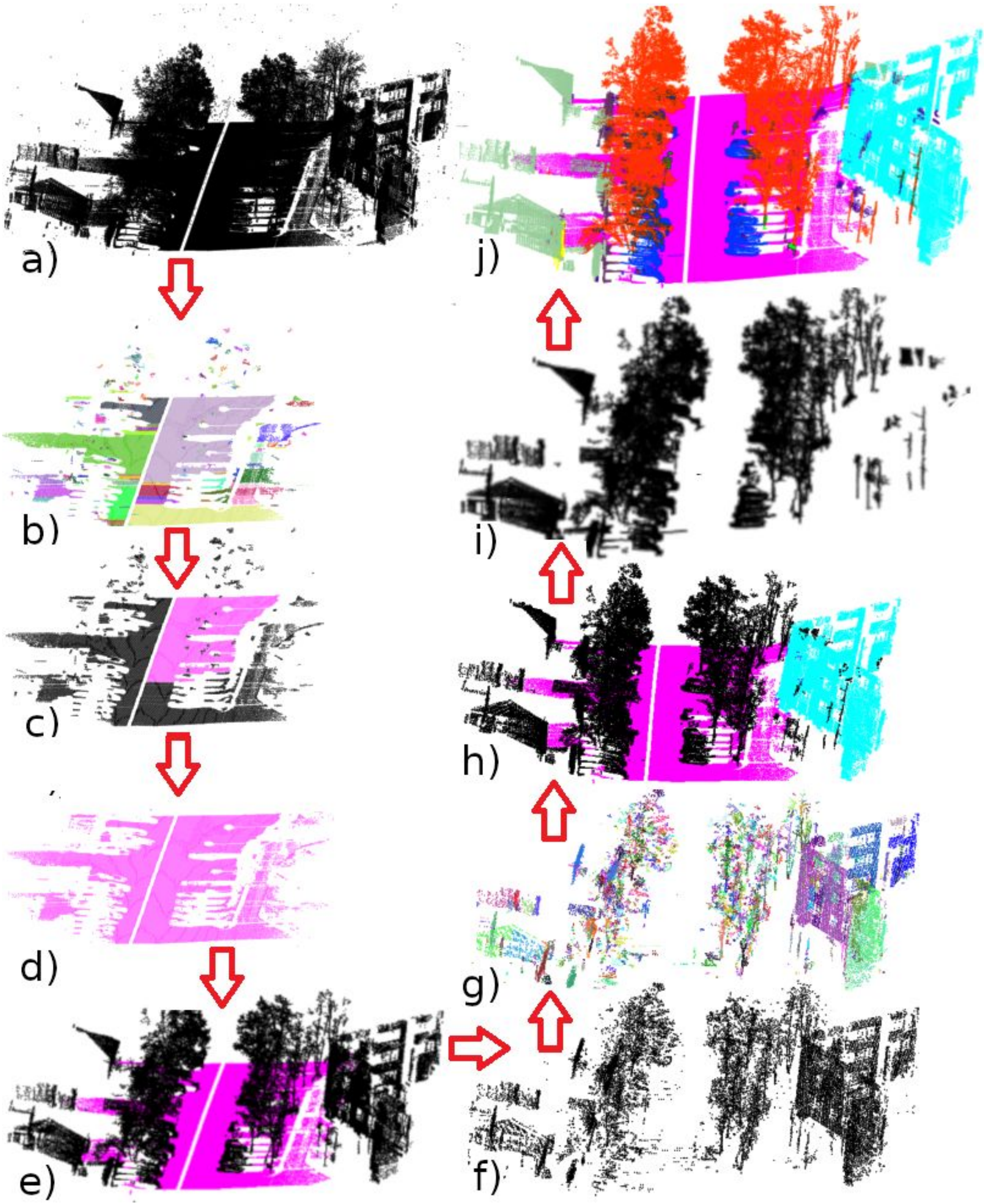

Abb. 3: The result of the object detection process step by step applied to a point cloud provided by Fraunhofer IPM

4 Discussion

The presented approach is compared to two other approaches specialized in the detection of walls and grounds. The three approaches have been applied on the same test point cloud to obtain comparable results. The result of this comparison is illustrated in Table 1.

Table 1: Result comparison between the presented approach and two other approaches on the point cloud corresponding to the use case composed of 3190793 points.

Approach / Result	Classify	Fail
Presented approach	3075320	115473
	96.38%	3.62%
Anagnostopoulos et al. (2016)	2877572	313221
	90,18%	9,82%
Xing et al. (2018)	2746148	444645
	86.06%	13.94%

In the paper (Anagnostopoulos et al., 2016), the authors proposed a method to exclusively detect walls and grounds in point cloud through the combination of linear algorithms. This method well detects the ground but fails to detect some walls due to a lot of missing part in it.

In the paper (Xing et al., 2018), the authors proposed a method of feature recognition based on the application of SWRL-rules to classify walls and ground in an urban context. This method proposes to apply a planar segmentation and uses then the semantic to classify planar segments. Nevertheless, this approach depends on the planar segmentation result and thus, fails to detect walls which are primarily composed of vegetation.

Generally, generic object detection approaches are less accurate than detection approaches dedicated to the detection of one or two objects. However, in this comparison study, our approach is more precise than two other approaches that are specialized in ground and wall detection. Our approach thus shows that it is both generic (see application case of a watermill detection into a point cloud from cultural heritage, Ponciano et al. 2019) and more robust than the other approaches.

5 Conclusion

The approach we present in this article is able to detect a wide variety of different objects regardless of their size (such as very large walls or smaller traffic signs), geometric complexity (such as cars) or different representations (such as trees). In addition, the detection approach remains robust and effective despite the divergence in the representation of objects in the data. Indeed, objects occluded by others or having a low point density due to the acquisition process (scanner too far from the object for example) are correctly detected. Nevertheless, this approach is based on the basic knowledge of the objects described, and although this knowledge is enriched by an intelligent iterative process, 3.62% of the data is not well recognized. This is due to a description of the objects that are not exclusive and not enough accurate. That is why future work consists in improving the description of objects and increasing their diversity, thus strengthening detection and increasing the use of our approach in different fields.

Literatur

Anagnostopoulos, I., Pătrăucean, V., & Brilakis, I., Vela, P. (2016). Detection of Walls, Floors, and Ceilings in Point Cloud Data. 2302-2311. 10.1061/9780784479827.229.

Anguelov, D., Taskarf, B., Chatalbashev, V., Koller, D., Gupta, D., Heitz, G., & Ng, A. (2005). Discriminative learning of markov random fields for segmentation of 3d scan data. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 2, pages 169–176. IEEE.

Grilli, E., Menna, F., & Remondino, F. (2017). A review of point clouds segmentation and classification algorithms. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42:339.

Hoiem, D., Efros, A. A., & Hebert, M. (2005). Geometric context from a single image. In Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on, volume 1, pages 654–661. IEEE.

Maillot, N. (2005). Ontology based object learning and recognition. PhD thesis, Universite Nice Sophia Antipolis.

Ouyang, W., Wang, X., Zeng, X., Qiu, S., Luo, P., Tian, Y., Li, H., Yang, S., Wang, Z., Loy, C.-C., & Tang, X. (2015). Deepid-net: Deformable deep convolutional neural networks for object detection. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Pal, N. R. & Pal, S. K. (1993). A review on image segmentation techniques. Pattern recognition, 26(9):1277– 1294.

Ponciano, J.-J., Boochs, F., & Tremeau, A. (2017). *Knowledge-based object recognition in point clouds and image data sets*. gis.Science, 3:97104.

Ponciano, J.-J., Karmacharya, A., Wefers, S., Atorf, P., & Boochs, F. (2019). Connected semantic concepts as a base for optimal recording and computer-based modelling of cultural heritage objects. In Structural Analysis of Historical Constructions, pages 297–304. Springer.

Torralba, A., Murphy, K. P., & Freeman, W. T. (2010). Using the forest to see the trees: exploiting context for visual object detection and localization. Communications of the ACM, 53(3):107–114.

Triebel, R., Schmidt, R., Mozos, O. M., & Burgard, W. (2007). Instance-based amn classification for improved object recognition in 2d and 3d laser range data. In Proceedings of the 20th international joint conference on Artificial intelligence, pages 2225– 2230. Morgan Kaufmann Publishers Inc.

Wirges, S., Fischer, T., Frias, J. B., & Stiller, C. (2018). Object detection and classification in occupancy grid maps using deep convolutional networks. arXiv preprint arXiv:1805.08689.

Xing, X.-F., Mostafavi, M.-A., & Chavoshi, S. H. (2018). A knowledge base for automatic feature recognition from point clouds in an urban scene. ISPRS International Journal of Geo-Information, 7(1):28.

Ponciano, Jean-Jacques
Hochschule Mainz, Institut für Raumbezogene Informations- und Messtechnik
jean-jacques.ponciano@hs-mainz.de

Boochs, Frank
Hochschule Mainz, Institut für Raumbezogene Informations- und Messtechnik
frank.boochs@hs-mainz.de

Trémeau, Alain
Laboratoire Hubert Curien, Université Jean Monnet, Saint-Étienne (France)
alain.tremeau@univ-st-etienne.fr