THE SMART POINT CLOUD
Structuring 3D intelligent point data

Florent Poux

UNIVERSITY OF LIÈGE
Doctoral dissertation
2019

Université de Liège

THE SMART POINT CLOUD
Structuring 3D intelligent point data

par Florent Poux

Département de Géographie
Faculté des Sciences

Thèse présentée
en vue de l'obtention du grade de Docteur en Sciences

Juin 2019

This dissertation is approved by the promotor: Prof. Dr. R. Billen

Composition of the doctoral committee:

Prof. Dr Roland Billen	University of Liège, Belgium
Prof. Dr. Pierre Grussenmeyer	INSA Strasbourg, France
Prof. Dr. Pierre Hallot	University of Liège, Belgium
Prof. Dr. Fabio Remondino	Foundation Bruno Kessler, Italy

Composition of the Jury:

Prof. Dr. Pierre Hallot	President (University of Liège, BE)
Prof. Dr. Yves Cornet	Secretary (University of Liège, BE)
Prof. Dr. Roland Billen	Expert (University of Liège, BE)
Prof. Dr. Pierre Grussenmeyer	Expert (INSA Strasbourg, FR)
Prof. Dr. Fabio Remondino	Expert (Foundation Bruno Kessler, IT)
Prof. Dr. Peter Van Oosterom	Expert (Delft University of Technology, NL)

LIÈGE université

The research described in this doctoral dissertation was carried out at the Geomatics Unit part of the University of Liège, from October 2015 to June 2019.

Typeset by the author in Minion Pro using Word.
Cover design by Marina Tünsmeyer

Printed in the Netherlands, by Ridderprint BV

ISBN 978-94-6375-422-4
An electronic version of this dissertation is available at pointcloudproject.com

ABSTRACT

Discrete spatial datasets known as point clouds often lay the groundwork for decision-making applications. E.g., we can use such data as a reference for autonomous cars and robot's navigation, as a layer for floor-plan's creation and building's construction, as a digital asset for environment modelling and incident prediction... Applications are numerous, and potentially increasing if we consider point clouds as digital reality assets. Yet, this expansion faces technical limitations mainly from the lack of semantic information within point ensembles. Connecting knowledge sources is still a very manual and time-consuming process suffering from error-prone human interpretation. This highlights a strong need for domain-related data analysis to create a coherent and structured information. The thesis clearly tries to solve automation problematics in point cloud processing to create intelligent environments, i.e. virtual copies that can be used/integrated in fully autonomous reasoning services. We tackle point cloud questions associated with knowledge extraction – particularly segmentation and classification – structuration, visualisation and interaction with cognitive decision systems. We propose to connect both point cloud properties and formalized knowledge to rapidly extract pertinent information using domain-centered graphs. The dissertation delivers the concept of a Smart Point Cloud (SPC) Infrastructure which serves as an interoperable and modular architecture for a unified processing. It permits an easy integration to existing workflows and a multi-domain specialization through device knowledge, analytic knowledge or domain knowledge. Concepts, algorithms, code and materials are given to replicate findings and extend current applications.

Keywords: Knowledge Discovery; Knowledge Extraction; Knowledge Integration; Knowledge Representation; Cognitive Decision System; Intelligent Support System; Intelligent environment; 3D Semantics; 3D Indoor; 3D Database; 3D GIS; Ontology; Smart Point Cloud; Point Cloud Database; Point Cloud Structuration; Point Cloud Topology; Point Cloud Knowledge-Base; Point Cloud Training platform; 3D Automation; 3D Pattern recognition; 3D Clustering; Information extraction; Voxel; Octree; AI Inference; Segmentation; Classification

RÉSUMÉ

Les ensembles discrets de données spatiales, appelés nuages de points, forment souvent le support principal pour des scénarios d'aide à la décision. Par exemple, nous pouvons utiliser ces données comme référence pour les voitures autonomes et la navigation des robots, comme couche pour la création de plans et la construction de bâtiments, comme actif numérique pour la modélisation de l'environnement et la prédiction d'incidents... Les applications sont nombreuses et potentiellement croissantes si l'on considère les nuages de points comme des actifs de réalité numérique. Cependant, cette expansion se heurte à des limites techniques dues principalement au manque d'information sémantique au sein des ensembles de points. La création de liens avec des sources de connaissances est encore un processus très manuel, chronophage et lié à une interprétation humaine sujette à l'erreur. Cela met en évidence la nécessité d'une analyse automatisée des données relatives au domaine étudié afin de créer une information cohérente et structurée. La thèse tente clairement de résoudre les problèmes d'automatisation dans le traitement des nuages de points pour créer des environnements intelligents, c'est-à-dire des copies virtuelles qui peuvent être utilisées/intégrées dans des services de raisonnement totalement autonomes. Nous abordons plusieurs problématiques liées aux nuages de points et associées à l'extraction des connaissances - en particulier la segmentation et la classification - la structuration, la visualisation et l'interaction avec les systèmes cognitifs de décision. Nous proposons de relier à la fois les propriétés des nuages de points et les connaissances formalisées pour extraire rapidement les informations pertinentes à l'aide de graphes centrés sur le domaine. La dissertation propose le concept d'une infrastructure SPC (Smart Point Cloud) qui sert d'architecture interopérable et modulaire pour un traitement unifié. Elle permet une intégration facile aux flux de travail existants et une spécialisation multi-domaine grâce aux connaissances liée aux capteurs, aux connaissances analytiques ou aux connaissances de domaine. Plusieurs concepts, algorithmes, codes et supports sont fournis pour reproduire les résultats et étendre les applications actuelles.

Mots-clés : Nuage de points ; Découverte des connaissances ; Extraction des connaissances ; Intégration des connaissances ; Représentation des connaissances ; Système décisionnel cognitif ; Système de soutien intelligent ; Sémantique 3D Environnement intelligent ; Intérieur 3D ; Base de données 3D ; SIG 3D ; Ontologie ; Base de données de nuages de points ; Structure du nuage de points Topologie du nuage de points ; Base de connaissances du nuage de points ; Segmentation ; Classification ; Automatisation 3D ; Reconnaissance de formes 3D ; Cluster 3D ; Extraction d'informations ; Voxel ; Octree ; Inférence AI

KURZFASSUNG

Diskrete räumliche Datensätze, so genannte Punktwolken, bilden oft die Grundlage für Entscheidungsanwendungen. Beispielsweise können wir solche Daten als Referenz für autonome Autos und Roboternavigation, als Ebene für die Erstellung von Grundrissen und Gebäudekonstruktionen, als digitales Gut für die Umgebungsmodellierung und Ereignisprognose verwenden... Die Anwendungen sind zahlreich und nehmen potenziell zu, wenn wir Punktwolken als Digital Reality Assets betrachten. Allerdings stößt diese Erweiterung vor allem durch den Mangel an semantischen Informationen innerhalb von Punkt-Ensembles auf technische Grenzen. Die Verbindung von Wissensquellen ist immer noch ein sehr manueller und zeitaufwendiger Prozess, der unter fehleranfälliger menschlicher Interpretation leidet. Dies verdeutlicht den starken Bedarf an domänenbezogenen Datenanalysen, um eine kohärente und strukturierte Information zu schaffen. Die Arbeit versucht eindeutig, Automatisierungsprobleme in der Punktwolkenverarbeitung zu lösen, um intelligente Umgebungen zu schaffen, d.h. virtuelle Kopien, die in vollständig autonome Argumentationsdienste verwendet/integriert werden können. Wir befassen uns mit Punktwolkenfragen im Zusammenhang mit der Wissensextraktion - insbesondere Segmentierung und Klassifizierung - Strukturierung, Visualisierung und Interaktion mit kognitiven Entscheidungssystemen. Wir schlagen vor, sowohl Punktwolkeneigenschaften als auch formalisiertes Wissen zu verbinden, um schnell relevante Informationen mithilfe von domänenzentrierten Grafiken zu extrahieren. Die Dissertation liefert das Konzept einer Smart Point Cloud (SPC) Infrastruktur, die als interoperable und modulare Architektur für eine einheitliche Verarbeitung dient. Es ermöglicht eine einfache Integration in bestehende Workflows und eine multidimensionale Spezialisierung durch Gerätewissen, analytisches Wissen oder Domänenwissen. Konzepte, Algorithmen, Code und Materialien werden zur Verfügung gestellt, um Erkenntnisse zu replizieren und aktuelle Anwendungen zu erweitern.

Schlüsselwörter : Punktwolke; Wissensentdeckung; Wissensgewinnung; Wissensintegration; Wissensrepräsentation; Kognitives Entscheidungssystem; Intelligentes Unterstützungssystem; Intelligente Umgebung; 3D Semantik; 3D Innenbereich; 3D Datenbank; 3D GIS; Ontologie; Smart Point Cloud; Punktwolke Datenbank; Punktwolke Strukturierung; Punktwolke Topologie; Punktwolke Knowledge-Base; Punktwolke Trainingsplattform; 3D Automatisierung; 3D Mustererkennung; 3D Clustering; Informationsextraktion; Voxel; Octree; AI Inferenz; Segmentation; Classification

CONTENT

Abstract 5

Résumé 6

Kurzfassung 7

Content 9

Acknowledgments 11

Chapter 1 Introduction 15

Chapter 2 Background 27

Chapter 3 A Smart Point Cloud Data Structure 43

Chapter 4 Knowledge Extraction 79

Chapter 5 Application to archaeology 121

Chapter 6 Knowledge integration & representation 165

Chapter 7 Conclusion and research perspectives 213

List of tables 226

List of figures 228

Bibliography 234

Curriculum vitae i

List of publications ii

Appendix A v

Appendix B vi

Appendix C vii

Everything has a point.

ACKNOWLEDGMENTS

Undertaking this PhD has been a truly life-changing experience and it would not have been possible to accomplish without the support and guidance that I received from many people.

I would like to first say a truly massive thank you to my supervisor Prof. Roland Billen. His unfailing trust and guidance were a large motor of motivation to surpass myself in what I can achieve. Moreover, Roland gave me the possibility to travel abroad, conduct research visits and publish in Open Access Journals while leaving me a large autonomy. For this and for being a very kind, understanding and open-minded mentor, I am very grateful.

My gratitude secondly goes to Prof. Fabio Remondino and Prof. Pierre Grussenmeyer for their leadership, expertise and their valuable feedback. Most importantly, I would like to thank them as well as Prof. Peter Van Oosterom for being eminent researchers, dedicated to high-quality works which were an invaluable inspiration for this thesis.

I would also like to thank the Geomatics Unit academics for their warm integration. Specifically, I thank the Prof. Yves Cornet for exchanging ideas and for agreeing to be in my Jury; I thank the Prof. Pierre Hallot for his contribution to the doctoral committee, for always being full of resources and for his help navigating the institution's administration.

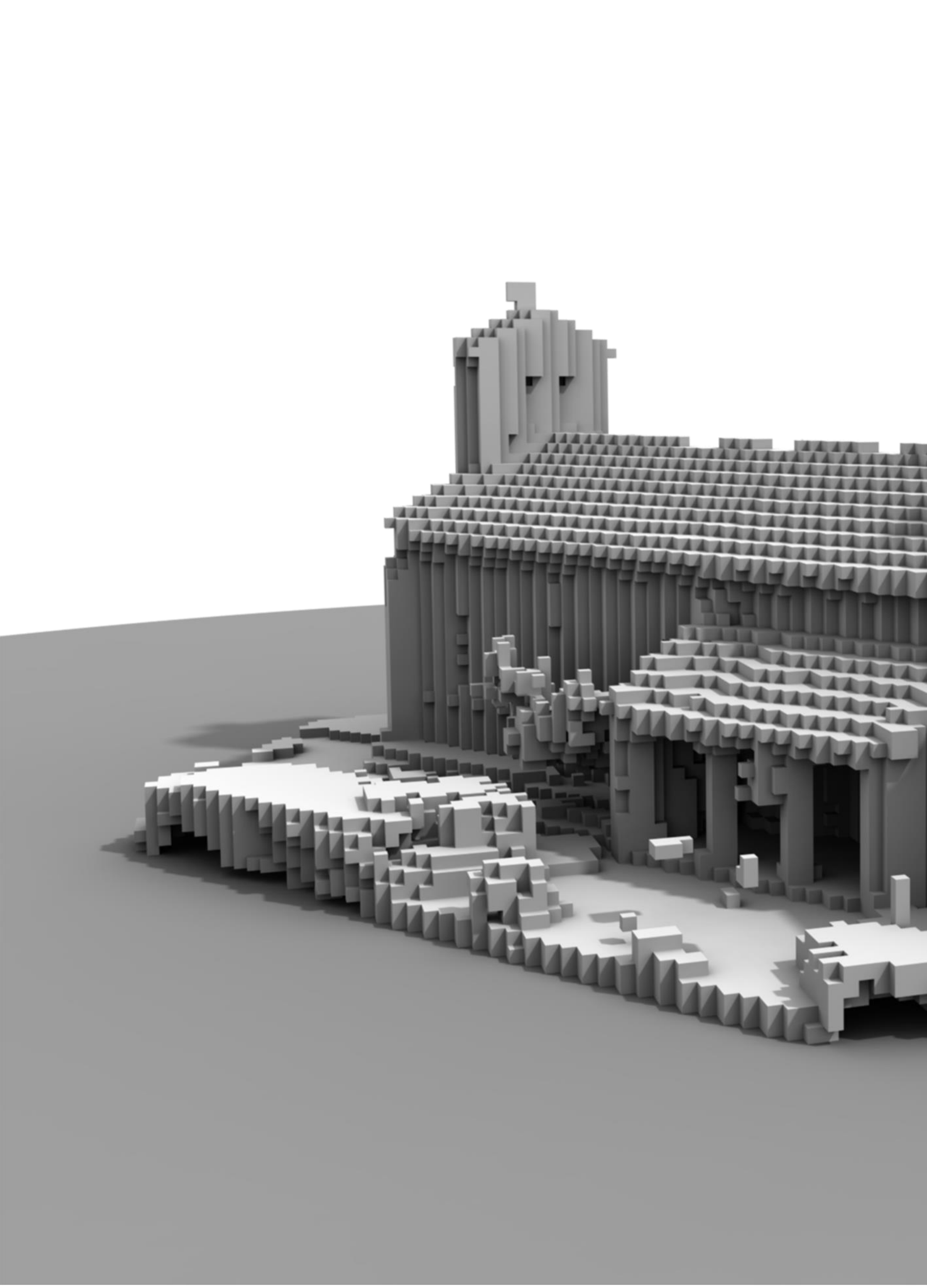
A very special gratitude goes out to all down at the Geomatics Unit as well as the Geography Department with a special mention to the other PhD students Cécile Deprez, Romain Neuville and Gilles-Antoine Nys. It was fantastic to have the opportunity to accomplish most of my research in your facilities, and to exchange so many ideas about work/life topics. Such an important spiritual contribution!

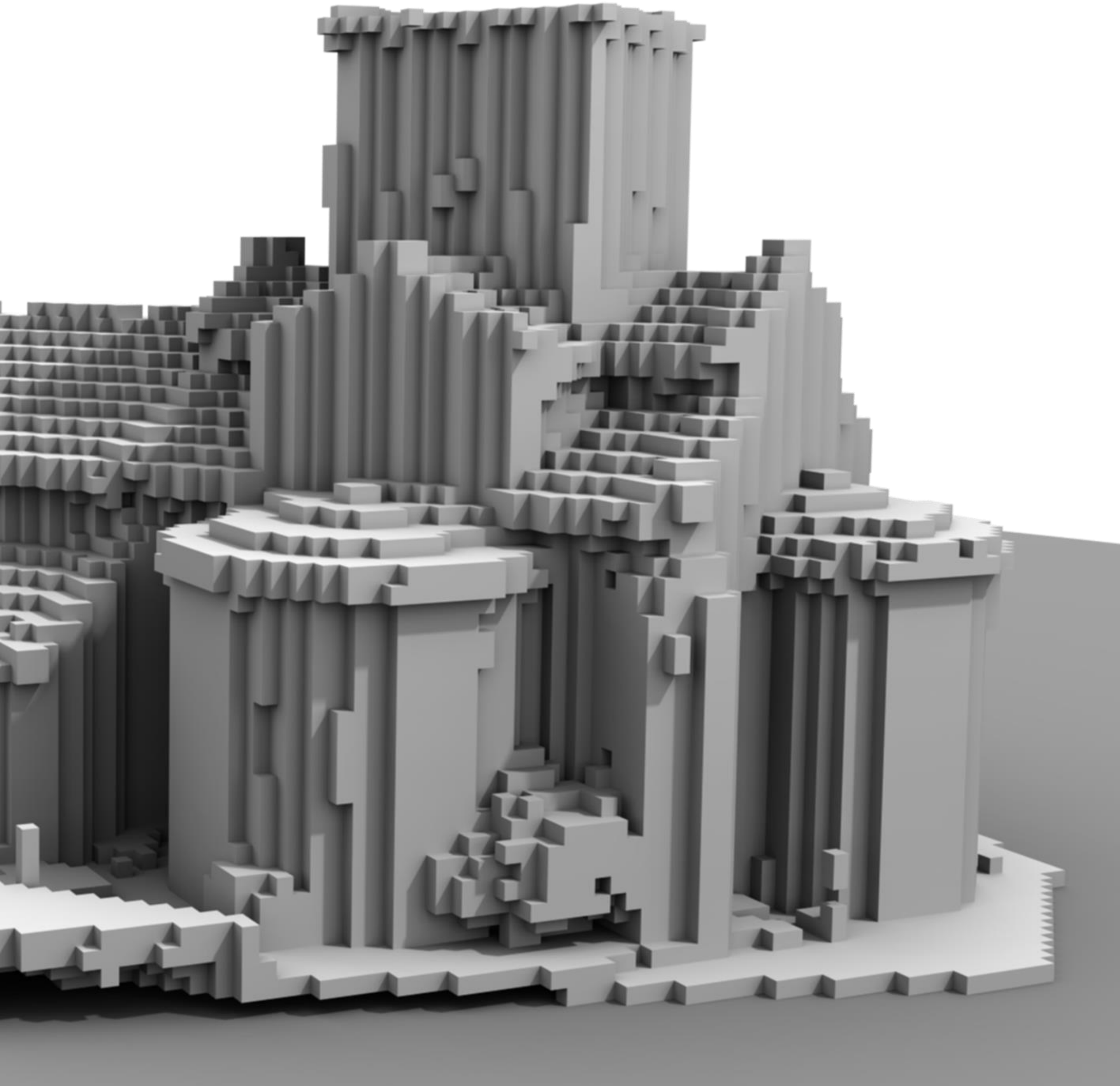
To my exceptional girls Marina & Lara, my life-balance and the main reason I get up in the morning. They are the guiding lights that will boost a momentary disorientation when everything seems dark. An eternal gratefulness for being there.

To my endlessly interested, admirable, encouraging and always enthusiastic grand-parents. They are and were (In the loving memory of Pierre Poux) always keen to know what I was doing and how I was proceeding, although it is likely that the content was a bit elusive!

I am forever grateful to my siblings and parents – Thibaut, Sonia, Alice, Hélène (alias Maman) and Michel (alias Papa) – who have provided me through moral and emotional support all my life.

To all my earthly friends, thanks for all your encouragement!





Point Cloud voxelisation of the church situated in Germigny des prés, France.

CHAPTER 1

Introduction

Chapter 1 - Introduction 14

1.1 Context 17
1.2 Problem statement 20
1.3 Research questions 21
1.4 Outline of the thesis and scope 21

1.1 CONTEXT

"when we open our eyes on a familiar scene, we form an immediate impression of recognizable objects, organized coherently in a spatial framework" [1]. In 1980, Treisman defines in simple terms the complex mechanism behind our human sight-perception. For non-impaired human-being, it is often the primary source of information which our cognitive decision system can use to act on. This is extendable using our brain which quickly adapts to new surroundings and only uses the most important material captured though our eyes. In fact, the brain receives just three "images" every second, which are sorted and combined with prior knowledge to create the reality that we experience. This mechanism is exceptionally fast and efficient allowing to brake when we see a red light, or simply to read this thesis and understand the spatial organization of words. Even more impressive, our vision can be adapted for an "orientation attention" – energy saving mode where the brain does not develop a full understanding of the surroundings – or a "discover attention" – which runs slower as the brain collects data from our memory to obtain a full understanding of the scene. With today's computational power and high level of dematerialization, virtually replicating such a process is not only very attractive but seems feasible. While the operation is genuinely hard to mimic, studying how we interact with our environment permits to better grasp the boundaries and usable mechanisms. It first translates into the use of sensors that can capture key inputs usable by a computer. We then aim at a procedure based on gathered data and accessible information repositories to produce a "semantic representation": a depiction of a scene integrating concepts and their meaning. In such a scenario, a spatial sensor plays the role of our eyes to obtain a digital spatial asset further refined into a semantic representation using available knowledge.

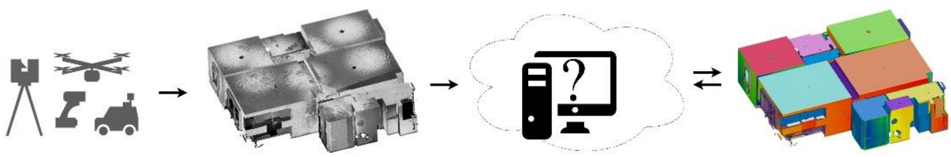


Figure 1. The sensor plays the role of our eyes, the spatial framework becomes a semantic representation, and the scene is tagged familiar using available knowledge

But this availability is often a first complication. Our online cognitive perception uses our memory and is structured to access needed evidence in a very short time. Mirroring this stage using a computer (Figure 1) is extremely complex and aiming for a solution as generalist as possible is an important challenge. The second bottleneck when trying to virtualize a cognitive decision system is the creation of a semantic representation (E.g. Figure 2). Gathering and attaching domain knowledge to underlying spatial data is linked to colossal integration and mining complications regarding data types, sources or representations.

Figure 2. 3D point cloud representation vs 3D semantic representation

The last main challenge revolves around the specificity of the data collected by the sensor(s). Single raster images or video streams are great when depth cues are not necessary, but emulating our 3D visual cognition demands a richer data basis. Reality Capture devices permit to obtain such an exhaustive 3D spatial information primarily as a point cloud: a {X, Y, Z} (+ attributes) spatial ensemble which digitally represents the recorded environment w.r.t the sensor strengths and limitations. The landscape of these instruments and acquisition methodologies is mature enough to allow digital replicas of the real world ranging from the object scale to the country scale (Figure 3).

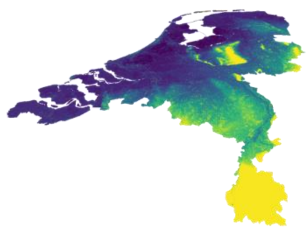
Object-scale point cloud

Terrestrial laser scanner building-scale point cloud

Mobile-based area-scale point cloud

Street level MMS point cloud

Aerial city-scale photogrammetric point cloud



LiDAR-based country-scale point cloud

Figure 3. Point cloud datasets using different techniques at different scales and using different methodologies

Moreover, the acquisition of these so-called point clouds has become easier, faster and is even accessible from very low-cost solutions. All these hardware evolutions were unfortunately not followed by their software counterpart, which are heavily impacted by the 5 V's of Big Data problematics (Figure 4).

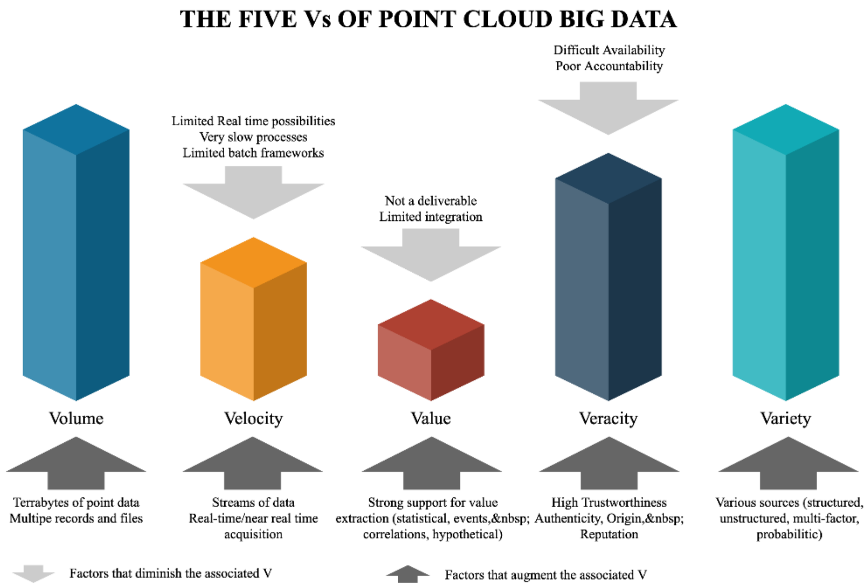


Figure 4. The Five Vs of Big Data in the context of point clouds.

Connecting numerous sensors/approaches creates heterogeneous point cloud datasets (Variety) and participate in the constitution of massive data repositories (Volume). In turn, it reduces the processing efficiency (Velocity) and creates new needs to turn huge amounts of point data into trustworthy (Veracity) and actionable information (Value). Specifically, the procedures to convert point clouds in application-specific deliverables are very costly in time/manual intervention. It is getting ever more complicated for the human expertise to handle adequately the large and complex volumes of information, often contradictorily disseminated among different actors/supports of one project. Thus, it is key for a sustainable system that big point cloud data translates into more efficient processes opening a new generation of services that help decision-making and information extraction. We need to find ways for massive automation and structuration to avoid task-specific manual processing and non-sustainable collaboration.

Interoperable approaches which permits several actors to leverage one common information system (E.g. Facility Management 4.0) based on a digital twin is a great exploration motor. In this continuum, the reflexion to go from a human-centered process to an autonomous workflow orient our research to develop automation and AI to speed-up inference processes, crucial to the development of point clouds in 3D capture workflows.

1.2 PROBLEM STATEMENT

Point cloud acquisition and processing workflows are usually application-dependent following a classic progression from data gathering to deliverable creation. While the collection step may be specific to the sensor at hands, point-cloud-as-a-deliverable upsurges, becoming one de-facto choice for many industries. This task-oriented scenario mainly considers these as a spatial reference – which is used by experts to create other deliverables – thus being a project's closest link to reality. It brings accurate real-world information which could allow decision-making based on digital-reality instead of interpreted or not up-to-date information. However, there are several considerations to address for a suitable integration. Point clouds are often very large depending on how much data is collected – usually in the realms of Gigabytes, if not Terabytes – and are destined to be archived as a reusable support to create new type of data and products. This can lead to a dead-end with exponential storage needs, incompatibility between outputs, loss of information and complicated collaboration. These practices also show limited to no attempt to generalize a framework which could in turn play as a common ground for further interoperability and generalization. This lack is counterproductive and could lead in term to a chaotic data repartition among actors and worsen the dependency to several outsourced service each aiming an application independently. This emphasize a strong need to study interoperable scenarios in which one point cloud could be used by many users from different domains, each having a different need. This will in turn introduce new constraints at the acquisition level to define the needed exhaustivity of the 3D representation for use with reasoning engines. Of course, this serialize additional challenges for interconnecting processes and insuring a compatibility with the different sources, volumes and other data-driven parameters.

Finally, robotics research has made a leap forward providing autonomous 3D recording systems, where we obtain a 3D point cloud of environments with no human intervention. Of course, following this idea to develop autonomous surveying means demand that the data can be used for decision-making. The collected point cloud without context does not permit to take a valid decision, and the knowledge of experts is needed to extract the necessary information and to creates a viable data support for decision-making. Automating this process for fully autonomous cognitive decision systems is very tempting but poses many challenges mainly link to Knowledge Extraction, Knowledge Integration and Knowledge Representation from point cloud. Therefore, point cloud structuration must be specifically designed to allow the computer to use it as a base for information extraction using reasoning and agent-based systems.

1.3 RESEARCH QUESTIONS

In this dissertation, I explore various topics related to the concept of Smart Point Clouds. The main research question that I seek to answer in this thesis is:

> How to extract and integrate knowledge within 3D point clouds for autonomous decision-making systems?

The research question is subdivided into multiple interrogations, targeting a theoretical, implementation-wise and experimental side of the global problematic.

Theoretical part enquires the following aspects:

1. How to structure efficiently point clouds with domain knowledge for interoperable workflows?
2. How to leverage a data structure to automate object detection over massive and heterogeneous point clouds?
3. How to connect reasoning services for autonomous decision-making scenarios?

Implementation and experimental parts cover the following questions:

4. How can open-source database systems integrate point cloud data and semantics? How should one provide domain connectivity?
5. How can a system handle the heterogeneity found within point clouds datasets and semantics for object detection? Can unsupervised frameworks relate to domain concepts?
6. How modular is the Smart Point Cloud Infrastructure? How efficient are the proposed point cloud processing modules (segmentation, classification, semantic injection, semantic modelling)?

1.4 OUTLINE OF THE THESIS AND SCOPE

This thesis is mainly based on papers that I have published during the course of my doctoral research and listed on pp. 234.

It provides all the necessary information for the development of an infrastructure to (1) handle point cloud data, (2) manage heterogeneity, (3) process and group points that retain a relationship (4) regarding a specific domain ontology and (5) that allow to query and reason for (6) providing a decision-making tool (7) including smart modelling. The thesis is structured in 7 chapters (including Chapter 1: Introduction and Chapter 7: Conclusion), each giving extended details on the proposed Smart Point Cloud Infrastructure showed in Figure 5.

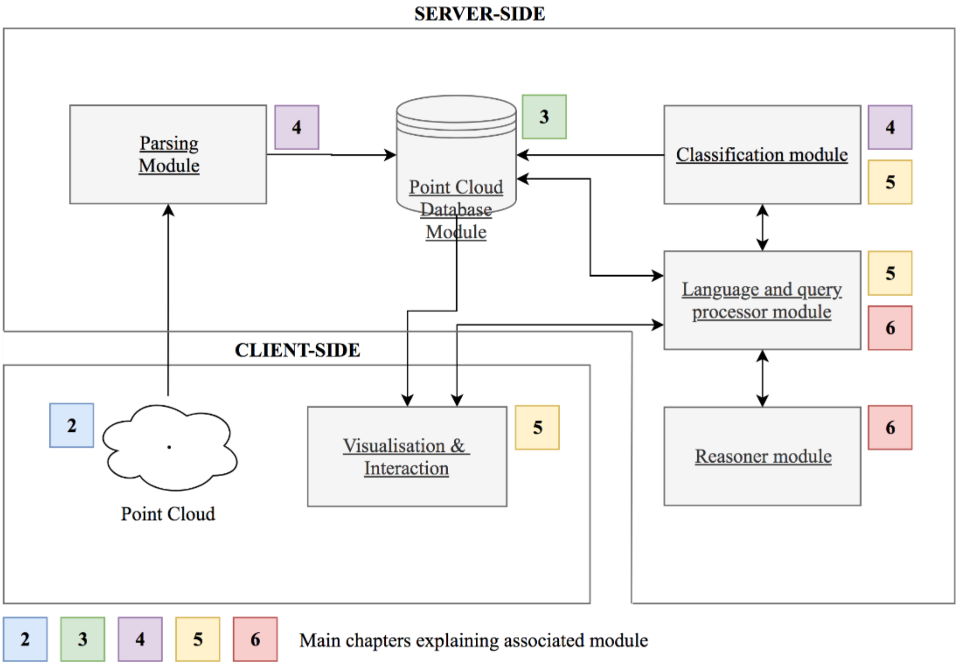


Figure 5. The Smart Point Cloud Infrastructure and its modular architecture associated to the thesis chapters.

Chapter 2 gives an overview of the most in-use point cloud workflows from acquisition to delivery. I first provide a comprehensive list of applications associated with used sensors and platforms. This introduces the main challenges by giving data-driven specificities linked to reality-capture devices, followed by problematics to structure & represent such datasets. Then, I address automation fundamentals and related limitations. The sections are voluntary succinct to give the reader a quick and structured outline with references to pertinent works if one wants more details. While this chapter doesn't participate into new research breakthrough, it is essential to the good comprehension of the thesis and to develop a full understanding of the motivations behind the formulated hypothesis.

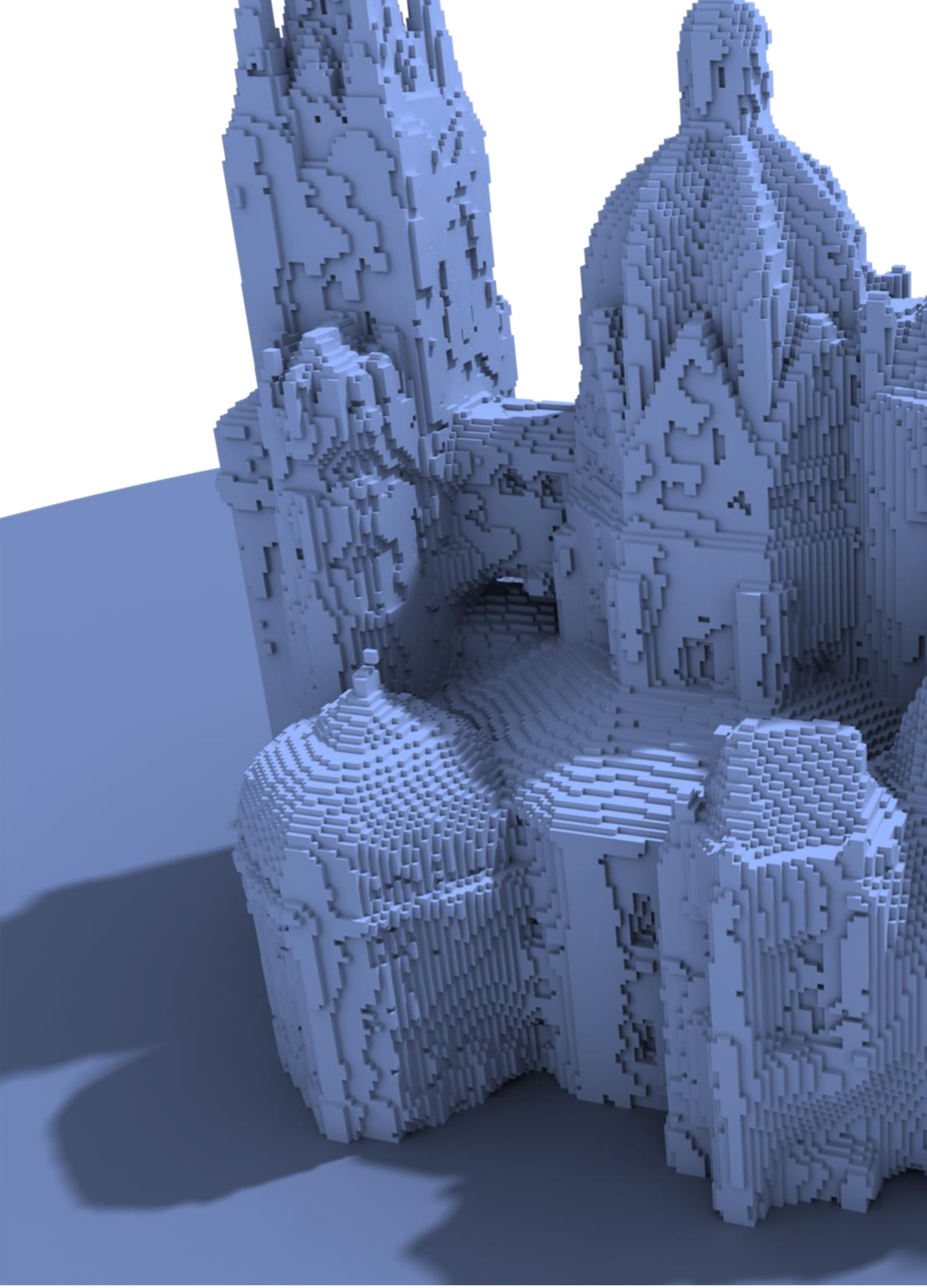
Chapter 3 details the proposed Smart Point Cloud (SPC) Infrastructure for a modular and centralized point cloud framework. It provides a conceptual data model to structure 3D point data, semantics and topology proficiently. It aims at creating a point-based digital twin usable by Cognitive Decision Systems. A multi-modal infrastructure integrating this data model is presented that includes Knowledge Extraction, Knowledge Integration and Knowledge Representation for automatic agent-based decision-making over enriched point cloud data. It constitutes the backbone of the thesis, and its modular nature permits efficient extensibility to several applications. Each module is then described in following chapters, giving the details of their specificity and their role toward interoperable digital reality. This chapter is based on the book chapter *"A Smart Point Cloud Infrastructure For Intelligent Environments"* [2] to be published in 2019 in ISPRS Book Series.

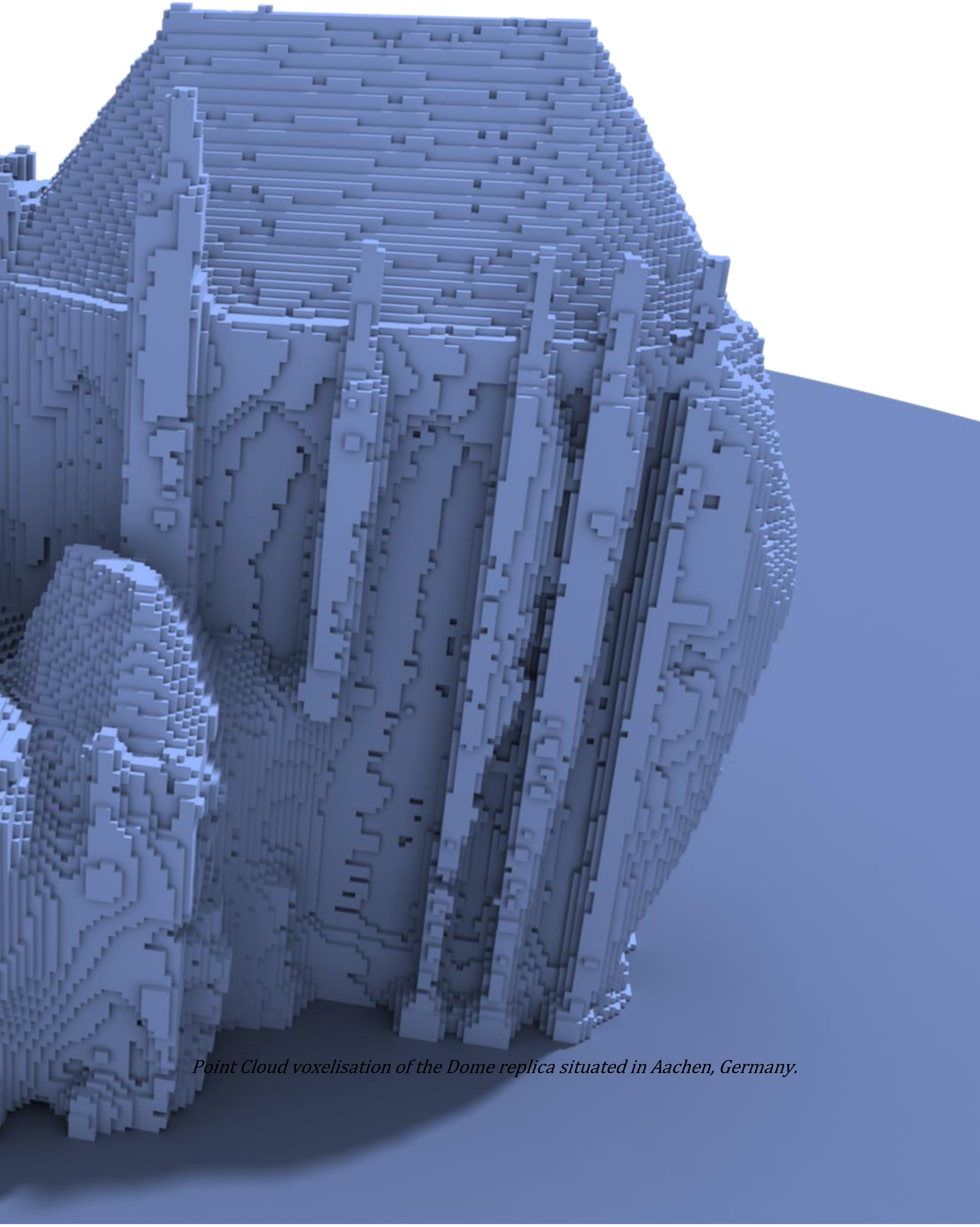
Chapter 4 develops the parsing module to best integrate point cloud data within the SPC Infrastructure while managing efficiently source heterogeneity. The chapter gives extended details on a voxel-based feature engineering method followed by an unsupervised segmentation approach that is used to better characterize point clouds and integrate them within automated workflows. The parser is also extendable through domain related classifications illustrated in the chapter. Algorithms, methods and metrics are given to benchmark the approach against best-performing deep learning mechanisms. This chapter is based on the article *"Voxel-based 3D point cloud semantic segmentation"* [3] to be published in 2019 in the Open Access ISPRS International Journal of Geo-Information.

Chapter 5 illustrates the flexibility of the SPC parser by plugging a formalized knowledge classifier for multi-sensory data. The chapter presents a domain-based mechanism exemplifying Knowledge Extraction, Knowledge Integration and Knowledge Representation over point cloud data in the context of archaeological research. As such, I give an exhaustive review of point cloud integration within archaeological applications and related 3D GIS. An acquisition, pre-processing, and ontology-based classification method on hybrid point clouds is proposed leveraging the SPC Infrastructure. It also proposes a web-based prototype for the visualisation of complex queries. It provides an example of ontology-based classification as a module while demonstrating the possible interactions with a 3D semantic representation. It is based on the paper *"3D point cloud in archaeology"* [4] published in 2017 in the Open Access journal Geosciences.

Chapter 6 shows the possibility to attach domain-specific agent-based modules to provide automated inferences. It presents an integrated 3D semantic reconstruction framework that leverages segmented point cloud data and domain ontologies. The approach follows a part-to-whole conception which models a point cloud in parametric elements usable per instance and aggregated to obtain a global 3D model. SPC heuristics, context and relationships are used to deepen object characterization. Then, it proposes a multi-representation modelling mechanism augmented by automatic recognition and fitting by massive 3D data mining. Its conception permits a direct integration within the landscape of 3D data by providing an interoperable way to map point clouds with other 3D representations. This chapter is based on the article *"3D point cloud semantic modelling"* [5] published in 2018 in the Open Access journal Remote Sensing.

Chapter 7 concludes the thesis with the key takeaways and answers the research questions by summarizing the main contributions of the research. Finally, I propose a roadmap for future work by giving research questions that were raised by doing this research.





Point Cloud voxelisation of the Dome replica situated in Aachen, Germany.

CHAPTER 2

Background

Chapter 2 - Background 27

2.1 Applications & Industries 29
2.2 3D point cloud data specificities 32
2.3 Representation & structuration 34
2.4 Automation 37

The dissertation is inspired by previous works gradually presented along chapters 2, 3, 4, 5 and 6 following Figure 6. These references are introduced in the second section of each chapter and organized to fit its content. Figure 6 acts as a point of reference to identify where the most pertinent state-of-the-art resources can be found and to help the reader navigate the essay.

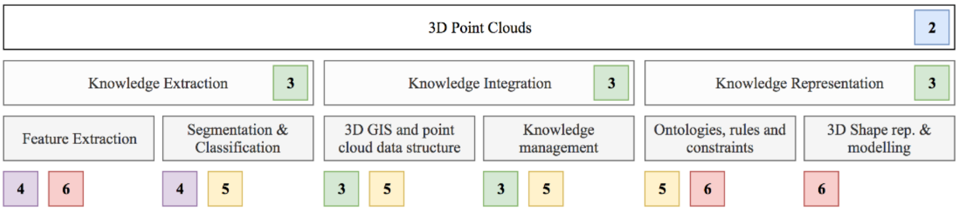


Figure 6. State-of-the art resources given in the thesis. It is gradually organized to first present point clouds specificities and challenges in Chapter 2, then decomposed in Chapter 3 regarding Knowledge Extraction, Integration and Representation further specialized toward automation and structural tasks.

Every addressed topic is related to 3D point clouds which we present in the current Chapter 2. We believe that the more we dive into digital processes, the more important it is that we keep a link to real-world applications. Thus, to understand the potential of point clouds, it is central to overview main applications (2.1) and the challenges that limit a broader use. We categorized these challenges by first giving data-driven specificities linked to reality-capture devices (2.2). Then we list the main problematics we face to structure & represent such datasets (2.3) and finally we address automation fundamentals. These subsections are voluntary succinct to give the reader a quick and structured overview with references to pertinent works if one wants more details.

2.1 APPLICATIONS & INDUSTRIES

The way digital disruption is affecting industries changes the practices and creates new applications every day. Reality capture and 3D point clouds are part of these disruptive technologies, which permits the creation of spatial datasets at various scales. This rich information then serves industries such as surveying and engineering; buildings and architecture; public safety and environmental issues; heavy construction, power and plants; transportation and navigation; mining and natural resources; environmental and man-made structures monitoring. Their utility is found at many levels and seeing the number of in-use application gives a good idea of the potential they hold. We list in Table 1 the main applications from these industries categorized by sensor-related scale.

Scale/Sensor	Main applications
Object-scale TrLS[1], TPHO[2]	Molecular density, target identification, free-form component inspection, reverse-engineering, manufacturing documentation, quality control, gaming asset creation
Building-scale TLS[3], HHLS[4], TPHO	Structural deformation, BIM, 3D mapping, architectural analysis, gaming asset/environment creation, façade inspection, built environment, space optimization, construction progress monitoring, deformation control, coordination, reconstruction, restoration, conservation, asset management, offsite production, navigation maps, robotics, ship build documentation, facility management, planning technical modifications
Area-scale TLS, BMLS[5], APHO[6]	Crime-scene investigations, site mapping and discovery, virtual tours, 3D designs, forensics, agriculture, bullet-path reconstruction, fire investigation, security and pre-planning, volume calculation
Street-scale TLS, BMLS, MLS[7], APHO	Roadway design, infrastructure (bridge, railway …) inspection, parking utilization, corridor mapping, traffic congestion, road signs and utility management, forensics and accident investigations, passive safety of cars, autonomous vehicles and self-driving cars.
City-scale APHO, MLS, ALS[8]	Gaming environment creation, urban survey, light simulations, airport infrastructures, utility planning, vulnerabity studies, earthquake damage assessment, cellular network planning, solar energy planning, tourism and park management, urban planning, 3D cadaster, building classification
Region-scale ALS, APHO	DEM, DTM, DSM, Micro-topography, Forestry, environmental monitoring (land, glacier, coastline, dune, windfarm, shoreline, change detection …), agriculture, flood/pollutant modelling, ecological and land classification, mapping, meteorology, geology, astronomy, topographical mapping, tsunami prediction, mining, oil and gaz exploration

Table 1. Applications making use of reality-based point cloud data organized by looking at the scale of the studied environment.

From this high number of applications today in use, one is particularly interesting for its numerous challenges: self-driving cars. These are equipped with various LiDAR sensors [6] which work as a "powerful" eye for the autonomous vehicle: an eye that allow one to see in all directions all the time. Imagine if, instead of guessing, we could always know the precise distance of objects in relation to us. LiDAR enables a self-driving car to view the surroundings with these special "powers". These are a big research motor allowing sensor mapping for guidance systems. Gerla et al. [7] state that "Vehicles will be a distributed transport fabric capable to make its own decisions about driving customers to their destinations". This statement hides big challenges: It demands massive automation for object

[1] Arm/Triangulation Laser Scanner
[2] Terrestrial photogrammetry
[3] Terrestrial Laser Scanner
[4] Hand-held Laser Scanner
[5] Backpack-Mounted Laser Scanner
[6] Aerial Photogrammetry
[7] Mobile Laser Scanner
[8] Aerial Laser Scanner (Aerial LiDAR)

detection while managing highly dimensional data from heterogenic sources. This implies looking into big data practices and data mining principles to correctly handle the volume, the velocity, the veracity and the variety of information.

The same problematics are found in scene understanding applications, a well-established research area in robotics [8]. Tasks like navigation, grasping or scene manipulation is essential to its applications, and depth sensors [9] are highly used to create environment in which robots can evolve and interact. The rapidly growing interest for UAV-based solutions [10] has made passive sensors such as thermal, infrared and RGB camera a common tool for creating point cloud via photogrammetry [11] and computer vision implementations [12]. On the industry side, new software based on Structure from motion [13] and multi-view stereo with bundle adjustment and SIFT, SURF, ORB, M-SURF, BinBoost descriptors [14,15] allowed a wide range of professionals and non-expert to recreate 3D content from 2D poses [16–25]. Use cases range from object-scale reconstruction to city-scale reconstruction, making this technique a promising way to get 3D point clouds. However, While reconstruction precision for middle to large scale are getting increasingly better [26], remote sensing via active sensors is favoured in several infrastructure-related industries.

Also mentioned in the Table 1, active triangulation, structured light and computer tomography for reverse engineering and modelling is highly used at the object-scale due to its high precision, and adaptation to small isolated objects [27]. Typically, deviation analysis and prototyping will be based on the point cloud acquired, and parametrization tasks will convert the point cloud to a mesh interpolation [28], or solid parametric geometry. To some extents, ALS point cloud processes relates closely to the same pipeline for mesh generation, specifically DTM, DSM and DEM. In both techniques, a critical step includes layering data, therefore classifying to obtain correct representations. I.e., the generation of a DTM needs to take into account only points belonging to what is considered "ground" [29], which present major automation challenges. Applications are numerous especially to map regions (corridor, roads, railway, landscapes), for product extraction, forest and coastal management, flood mapping, volume calculation, 3D city models creation for urban planning, glacier monitoring and other region-scale measurement's extraction listed in Table 1. This can be further refined using the full-waveform LiDAR data [30] providing better structural and physical information within the backscattered signal of the illuminated surface, allowing new data analysis and interpretation. Some Terrestrial Laser Scanners (TLS) also propose full waveform processing, and an example of application to determine canopy height is given in [31].

TLS have driven an engagement within manufacturers including Trimble, Topcon, Faro, Maptek, Optech, Riegl, Zoller + Fröhlich, Smart Max Geosystems, Neptec Technologies, TI Asahi, Clauss, Leica, focusing on Phase Based and Time of Flight technology [32] enabling different use cases depending mainly on the desired precision, the range and resolution of the point cloud. Topography, indoor mapping, AEC, monitoring, reconstruction, archaeology, cultural heritage conservation benefits of the high precision and versatility laser scanners offers. The high speed and rate generation of 3D points has become a convenient way to get instantly data, constituting datasets of up to Terabytes, so redundant and rich that control operation can take place in a remote location. However, it is still rare that point clouds are directly shared or used as an end-product. Rather, they are interpreted and analysis reports, simulations, maps, BIM and 3D models will be considered as deliverables. Rising from the static concept, Mobile Laser Scanning (MLS) [33] has scaled up the data rate generation of TLS by allowing dynamic capture allowing rapid street point cloud generation and public domain mapping. New concepts and technology including Solid State LiDAR and simultaneous localization and mapping (SLAM) has pushed dynamic acquisition for quickly mapping the surroundings, extending cases for indoor mapping using Hand-held Laser Scanners HHLS [34], Mobile Mapping Systems (MMS) [35], or more recently Backpack-mounted Laser Scanners (BMLS) [36].

While 3D point clouds are well-established as a source of spatial data, we often note that they are not the preferred default data type in the industry, mainly due to data specificity and structuration problematics. On top, both these characteristic and the variety of domain complicate automation workflows which industries aim to increase reliability and efficiency in established workflows.

2.2 3D POINT CLOUD DATA SPECIFICITIES

Despite the ease of capturing point clouds, processing them becomes a challenging task in part due to the problems listed Table 2. Indeed, point clouds obtained from the sensors described in sub-section 2.1 suffers mainly several artefacts (Table 2). These are the main data-driven obstacles to a wider dissemination, often related to their data-structure or capture-related environment specificities. The structure-related problems usually emerge due to a lack of connectivity within point ensembles, which can make the surface information ambiguous [37]. Furthermore, environment-related problems are usually present in real world acquisition setup which happens due to different reasons such as limitations of the sensors or some physical factors in the scene.

Artefact	Main source
Misadjusted density	Point clouds exhibit locally variable densities based on surface orientation, nature and distance from the capturing device. Occlusions from surface irregularities and adjacent objects also produce regions with missing data. Variations in density can be attenuated by sub-sampling techniques but not fully eliminated.
Clutter	A scene can contain small objects represented by very few points, moving objects and multiple objects in proximity which for an application are considered "noise". These are often making feature detection, structuration and automatic recognition difficult.
Occlusion	A scene can contain objects of significant size that occlude objects behind them (E.g cars, buildings…). This produces incomplete and disjointed descriptions of the background surfaces thus large missing areas. This phenomenon can be limited through an adapted acquisition technique, E.g. reduced using BMLS, MLS, MMS or HHLS systems compared to static devices.
Random errors (noise)	This is mainly due to absorption linked to the operation frequency, scattering, and taking into account the properties of the observed object + the wavelength of the incident energy. These influence the sensor's choice considering the application, firstly between active and passive sensors. As passive sensors rely solely on the light emittance of the measured object, they are more influenced by atmospheric conditions and errors.
Systematic errors	The sensor will measure a property in the scene, and to be highly representative, must be sensitive to only the value measured, without influencing the backscattered signal. However, errors such as zero input, sensitivity error will create additional noise, but these can be calibrated.
Surface properties	The physical texture of common surfaces can range from smooth (steel, marble) to very irregular (grass, crushed stone). Because a given scene can contain a wide range of surface roughness's, no priors about noise levels can be reliably used.
Misalignement (i.e. georeferencing errors)	Assembling point data in one reference system is a task that is primarily dependent on the acquisition platform (airborne, vehicle, tripod, satellite). Indeed, the data obtained by different LiDAR platforms are heterogeneous in three respects: (1) different perspectives: data collected by space-based laser scanning (SLS) and airborne laser scanning (ALS) systems are from a top view, while data collected by MLS or TLS systems are from a side view; (2) different spatial resolution: the resolution of ALS data is generally at the meter scale, while MLS and TLS data are at the centimeter scale, with TLS being more precise; (3) different content of focus: ALS data cover general features, while MLS data cover both trajectory sides. Furthermore, registration can introduce local noise due to imperfect correspondence between point clouds.

Table 2. Artefact commonly found in reality-based 3D point clouds. These are obstacles to a broader adoption of point cloud data as the default spatial data type.

As referred in [38], these artefacts produce missing and erroneous (noise, outliers, misalignment) data, which can arise from an improper set-up on the scene. It also happens when the surface doesn't permit a correct survey thus return an incomplete dataset. E.g. transparent and very reflective surfaces will produce unpredictable data. Objects like windows, mirrors, and shiny metal frames or ducts can generate either no-points associated to them, or worse, points on virtual reflected surfaces that place the object at a false location in space. Filtering techniques and knowledge based-interpretation are possible solutions to these problems to get the most complete description of the scene. Complementary, getting a high number of representative signal descriptors and parameters for each point permits a more reliable description important for automation through segmentation, classification and domain adaptation.

2.3 REPRESENTATION & STRUCTURATION

The 3D datasets in our computerized ecosystem – of which an increasing number come directly from reality capture devices presented in 2.1 – are found in different forms that vary in both the structure and the properties. Interestingly, they can be somehow mapped with success to point clouds thanks to its canonical nature. We provide a quick review of the main 3D data representations modes bindings to point clouds and we invite the reader to study [39–42] for more details.

First, point clouds can be mapped to shape descriptors [43,44]. These can be seen as a signature of the 3D shape to provide a compact representation of 3D objects by capturing some key properties to ease processing and computations. The nature and the meaning of this signature depend on the characteristic of the shape descriptor used and its definition. For example, global descriptors provide a concise yet informative description for the whole 3D shape while local descriptors provide a more localized representation for smaller patches in the shape. The work of Kazmi et al. [43], Zhang et al. [44] and more recently Rostami et al. [45] provide comprehensive surveys about such 3D shape descriptors.

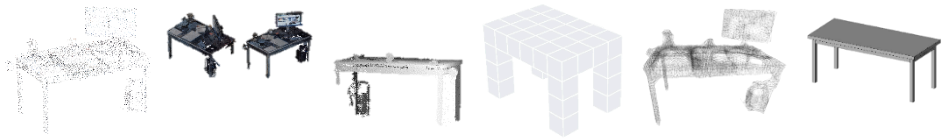
Secondly, projecting 3D data into another 2D space is another representation for raw 3D data where the projected data encapsulates some of the key properties of the original 3D shape [46]. Multiple projections have been proposed in the literature where each of them converts the 3D object into a 2D grid with specific features. Projecting 3D data into the spherical and cylindrical domains (e.g. [47]) has been a common practice for representing the 3D data in such format. Such projections help the projected data to be invariant to rotations around the principal axis of the projection and ease the processing of 3D data due to the Euclidean grid structure of the resulting projections. However, such representations are not optimal for complicated 3D computer vision tasks such as dense correspondence due to the information loss in projection [48].

Third, representing 3D data as RGB-D images has become popular in the recent years thanks to the popularity of RGB-D sensors. RGB-D data provides a 2,5D information about the captured 3D object by attaching the depth map along with 2D colour information (RGB). Besides being inexpensive, RGB-D data are simple yet effective representations for 3D objects to be used for different tasks such as identity recognition [49], pose regression [50] and correspondence [49]. The number of available RGB-D datasets is huge compared to other 3D datasets such as point clouds or 3D meshes [51] but they do not permit a direct 3D immersion.

Fourth, 3D data can be represented as a regular grid in the three-dimensional space. Voxels are used to model 3D data by describing how the 3D object is distributed through the three-dimensions of the scene. Viewpoint information about the 3D shape can be encoded as well by classifying the occupied voxels into visible, occluded or self-occluded. Despite the simplicity of the voxel-based representation it suffers from some constraining limitations [52]. Voxel-based representation is not always efficient because it represents both occupied and non-occupied parts of the scene, which can create an unnecessary demand for computer storage. More efficient 3D volumetric representations are often linked to indexing techniques addressed later.

Fifth, we can access 3D information from a multi-view image, which is a 2D-based 3D representation where one access the information by matching several 2D images for the same object from different point of views. Representing 3D data in this manner can lead to learning multiple feature sets to reduce the effect of noise, incompleteness, occlusion and illumination problems on the captured data. However, the question of how many views are enough to model the 3D shape is still open, and linked to the acquisition methodology for photogrammetric reconstructions: a 3D object with an insufficiently small number of views might not capture the properties of the whole 3D shape (especially for 3D scenes) and might cause an over-fitting problem. Both volumetric and multi-view data are more suitable for analysing rigid data where the deformations are minimal.

Sixth, 3D meshes are one of the most popular representations for 3D shapes. A 3D mesh structure consists of a set of polygons called faces described in terms of a set of vertices that describe how the mesh coordinates exist in the 3D space. These vertices are associated with a connectivity list which describes how these vertices are connected to each other. The local geometry of the meshes can be realized as a subset of a Euclidean space following the grid-structured data. However, on a global aspect, meshes are non-Euclidean data where the familiar properties of the Euclidean space are not well defined such as shift-invariance, operations of the vector space and the global parametrization system. 3D meshes can also be presented as graph-structured data where the nodes of the graph correspond to the vertices of the mesh and the edges represent the connectivity between these vertices.



(a) (b) (c) (d) (e) (f)

Figure 7. 3D data representation: (a) point cloud; (b) multi-view image; (c) depth map; (d) volumetric; (e) polygonal mesh; (f) parametric model.

Thus, the versatility given by point cloud mappings to 3D representations modes is very interesting and permits to leverage the strength of several approaches. However, this demands that we study ways to integrate point clouds directly in computerized systems.

This is a major challenge as the large discrete datasets that point clouds constitute cannot directly fit in the main memory. Handling efficiently these massive unstructured datasets (heterogeneous and from different sources) demands high scalability, speed (when data must be mined in a near or real-time style) and computational adaptation to answer specific needs.

This again relates to Big Data problematics, or how to efficiently process big semi-structured / unstructured datasets. relational Database Management Systems (DBMS) and NoSQL DBMS for such application provide interesting research tracks which are further explored in this thesis. The fundamental component in DBMS is the data model that determines the logical structure of a database determining in which manner the data can be stored, organized, and manipulated. We refer to the extensive works of Otepka et al. [53], Van Oosterom et al. [54] and Cural et al. [55] which provides fundamentals over existing large point cloud data structures including attribute and geometrical information organization. They rightfully state that the secondary storage access limits data-intensive tasks that could be solve through streaming algorithms to keep small parts in-memory. However, this implies pre-sorting and structuring a priori the data to handle the large volume and high resolution, thus linked to indexations problematics.

Indexation for 3D point clouds via spatial indices that subdivide the space through different approaches are a solution to reduce the overhead via chunk memory loading. The exhaustive paper presented by [56] state that the spatial subdivision of k-d tree are not suited for updates (e.g. add, remove) operations over point clouds because the tree structure becomes unbalanced. However, k-d trees perform well regarding Nearest-Neighbor searches by efficiently eliminating large portions of the search space ($\approx 0(\log n)$). Octree structures, a 3D analogy of quad-tree [57], as opposed to kd-tree perform well for update operations thanks to their uniform spatial subdivision, which makes them particularly interesting considering point cloud varying resolution, distribution and density. Linked to voxelized representations, an octree [58] is a tree-like data structure that models occupancy by dividing a 3D scene [56] into multiple hierarchical cubes. They are powerful in representing the fine details of 3D scenes compared to voxels with less computations because of their ability to share the same value for large regions of space. However, as stated by [59] , octrees are "*not able to dynamically adjust the tree structure according to the actual object layout. As a result, the tree depth is high where there are many objects, and this also results in unstable query performance*". Other interesting work such as 3D R-Tree [60], modified nested octrees or sparse voxel octrees [61] are also solutions for spatial data indexing techniques.

Point clouds are thus good candidates for Level of Detail (LoD) implementations for a better management (and rendering) of the data. But to be usable in extended workflows, ways to map semantics and domain knowledge to point clouds are to be addressed. A good introduction to knowledge injection within 3D scenes is accessible in [62]. The authors structure knowledge information in formal domain ontology [63], which could then be accessed through different queries. Formal ontologies thus are a great research track for leveraging domain specificities, but their integration with data artefacts and data structure is very manual. By extension, the extraction of knowledge from gathered data suffers from automation problematics that are very important obstacles to the usability and efficiency of point clouds within workflows.

2.4 AUTOMATION

The concept of 3D point cloud automation is found at different stages, illustrated in Figure 8 presenting a classical workflow.

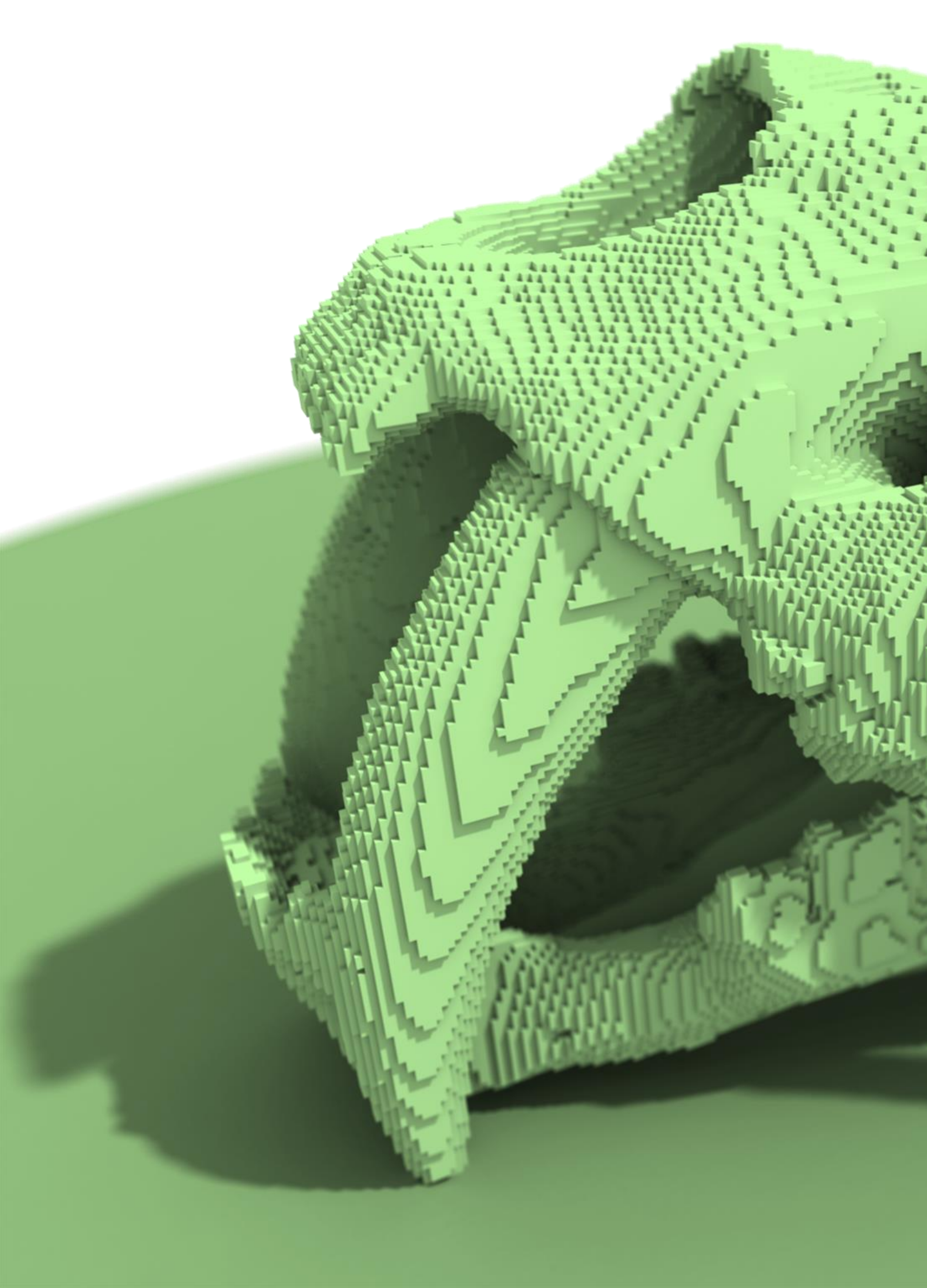
Figure 8. Classical point cloud workflow from acquisition to delivery of a product for a specific application.

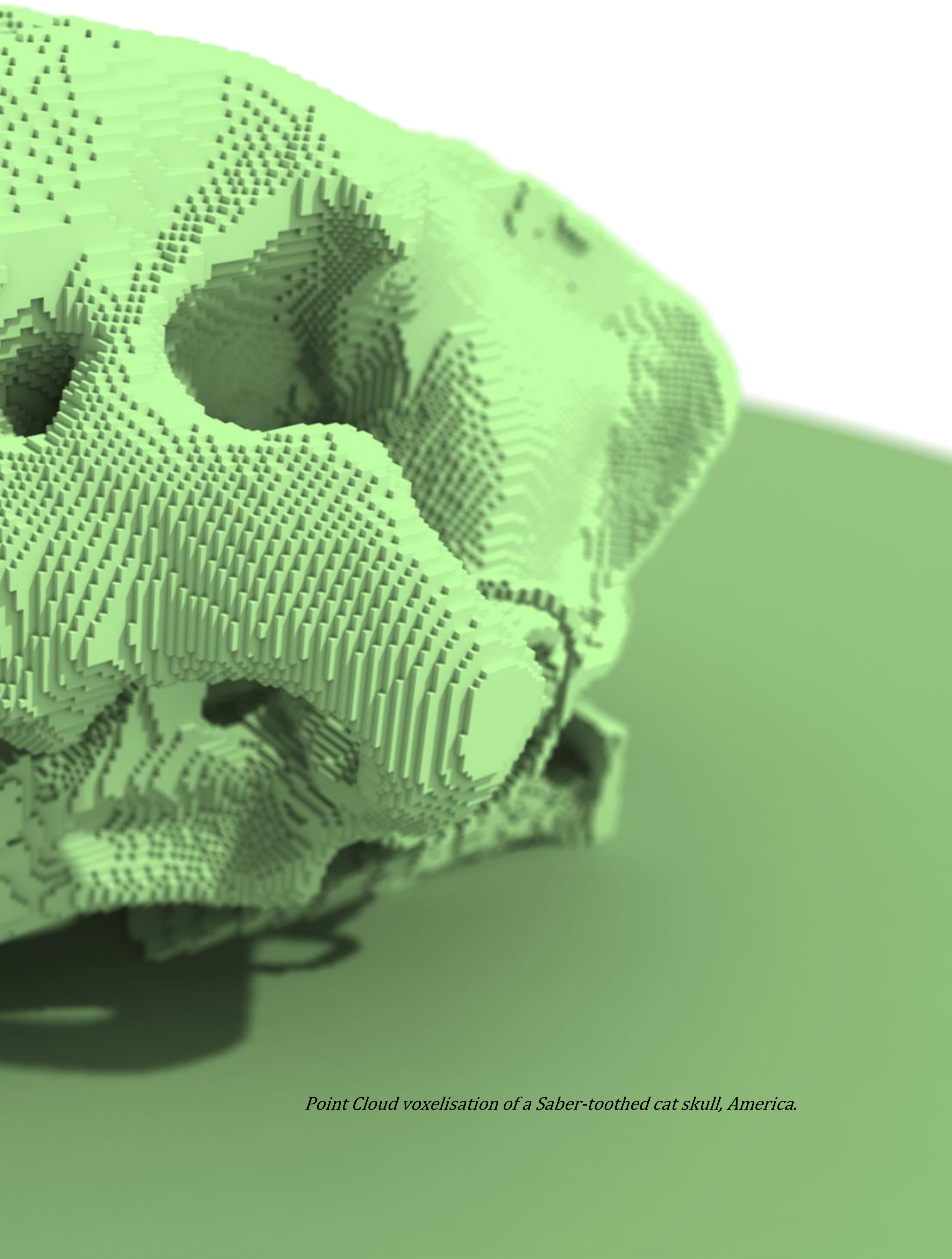
While automation for the acquisition of point clouds, pre-processing and registration opens multiple research tracks, we focus in this dissertation on segmentation, classification, structuration and application-related automation. However, the reader can find further information regarding the acquisition phase automation which mainly deals with capturing platforms [64], such as UAVs [10], multi-sensory robots [65] and efficient navigation systems [66]. For the pre-processing stage which mainly deals with filtering techniques to avoid sensor-related problematics and for calibrating devices, we refer to the work of Kaasalainen et al. [67]. As for the registration step, we refer to feature-based and featureless automatic registration methodologies reviewed in [68,69], SLAM reviewed in [70] and dense-matching in [15,71].

Automation in detecting objects by grouping points that share a similarity and decisive criterion is the basis for segmentation, thus classification. This step is crucial since accuracy of the subsequent processes depends on the validity of the segmentation results [72] and requires to balance flexibility with the ease of use. As point cloud sets get larger, segmentation methods need to be scalable in terms of time and memory complexity. Special attention needs to be paid to offline processes – that can be run over night without user supervision – and online responsive processes that leverage user interaction and geared toward real-time applications. State-of-the-art segmentation and classification applied to point cloud references are reviewed in [73], and deep learning techniques for semantic segmentation in [74]. It is important to note that the major challenges concern the domain specialization, which will majorly orient a classification approach. Classifiers that learn from previous or available knowledge differ from their approach, thus their results. They are usually categorized in 3 ensembles as in being supervised learning (from a set of features to a labelled data) unsupervised learning (structure detection by pattern recognition) and reinforcement learning (functional inference through a set of state and actions). While supervised method often proves more efficient than its unsupervised analogue, they can suffer from over-fitting which limits their extensibility and interoperability with wider applications. This is a main concern for massive automation, and it is tackled in the dissertation. Several approach demands a feature selection and engineering among the most in-use 3D descriptors reviewed in [75], which permit to extract a fine-tuned description of the underlying data. However, it is important to note that in point cloud classification frameworks and feature estimation, the suitability of features should privilege quality over quantity [76]. This shows a need to prioritize and find robust and relevant features to address the heterogeneity in a point cloud structure. Several approaches make use of Decision-trees [77], Random Forests (RF) [78] extended to Streaming Random Forests [79], Support Vector Machines (SVM) [80], Conditional Random Fields (CRF) [81], Neural Networks [82,83] and multiple variants [39,45,91–93,83–90].

As stated by [94] *"any fully-fledged system should apply as much domain knowledge as possible, in order to make shape retrieval effective"*. With the rise of online solutions, there is a great potential in using formalized knowledge for classification to analogically associate shapes and groups of points with similar features. This association through analogy *"is carried out by rational thinking and focuses on structural/functional similarities between two things and hence their differences. Thus, analogy helps us understand the unknown through the known and bridge gap between an image and a logical model"* [95]. This introduces the concept of data association for data mining, and relationships between seemingly unrelated data in an information repository. The use of domain knowledge over point cloud data by separating domain knowledge from operational knowledge refers to ontologies, although knowledge-based applications do not always refer to ontology reasoning. These concepts orient research in structural and application-related automation for a more interoperable and generalizable framework. Some attempt were made such as in [96] to turn a kitchen point cloud into a meaningful representation for robot interaction; [97] developing an interesting web object recognition workflow; [66] tackling robot-vision understanding and navigation challenges by proposing an abstracted spatial hierarchy graph and a semantic hierarchy that model domain concepts. Overall, attaching semantic concepts to point clouds for further reasoning was attempted in several disciplines but never generalizable. On top, some propose a way to infer knowledge in segmentation and classification method, papers rarely cover the topic of data structuration. To keep a record and use ontologies over analysis process, the point cloud needs to be structured retaining spatial and relation information deducted or useful for classification and segmentation. For data visualisation, it is also very important to work over a structure as flexible as possible to handle billions of records and queries over different attributes for validation through visual perception.

It appears that 3D data capture workflows would benefit from semantically rich point cloud in order to automate reasoning for an application, end-point of Figure 8. Using connectivity, material properties, date stamp or even description of chunk-wise group of points can be useful in almost all scenarios, even for mesh derivation. Managing highly dimensional data and heterogenic sources therefore goes through the definition of efficient automated procedures that describe the nature and properties of a point cloud sample in order to classify and establish relations between point segments in a new data structure.





Point Cloud voxelisation of a Saber-toothed cat skull, America.

CHAPTER 3

A Smart Point Cloud Data Structure

Chapter 3 - A Smart Point Cloud Data Structure 42

Chapter's preface 45
3.1 Introduction 48
3.2 Semantics & 3D Data 50
3.2.1 3D spatial information systems 50
3.2.2 Knowledge integration solutions for 3D semantically-rich data 51
3.2.3 Point cloud solutions for knowledge integration 53
3.3 The Smart Point Cloud (SPC) Data model 55
3.3.1 Knowledge categorization 55
3.3.2 Conceptual SPC (Smart Point Cloud) Model 56
3.3.3 Level-0: Generalized SPC meta-model 58
3.3.4 Level-1: Connection-layer meta-model 60
3.3.5 Level-2: Domain adaptation 62
3.4 Framework articulation & automation for intelligent environments 63
3.4.1 Infrastructure modularity and extensibility 64
3.4.2 SPC Requirements benchmarking 69
3.5 Limitations, Perspectives and possibilities 74
3.6 Conclusion 75

CHAPTER'S PREFACE

In this chapter 3, I will give the necessary details for the realisation of a Smart Point Cloud Infrastructure. The main element is the conceptual data model which permits the structuration of the Point Cloud Database Module as highlighted in Figure 9.

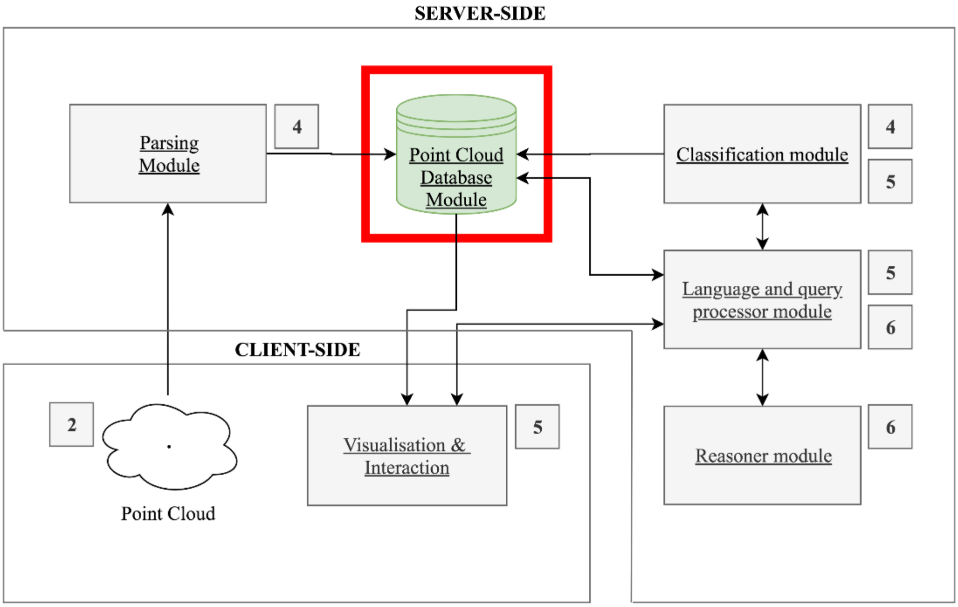


Figure 9. Chapter 3: A Smart Point Cloud Data Structure

Details about the integration and compatibility of each module are also given. This chapter lays the groundwork to comprehend how every subsequent module described in chapters 4, 5 and 6 interact, and permits to get a precise understanding of the proposed framework.

Based on Book Chapter [2]

A Smart Point Cloud Infrastructure for intelligent environments

Abstract: 3D Point cloud data describes our physical world spatially. Knowledge discovery processes including semantic segmentation and classification are a great way to complement this information by leveraging analytic or domain knowledge to extract semantics. Combining efficiently these information's is an opening on intelligent environments and deep automation. This chapter provides a conceptual data model to structure 3D point data, semantics and topology proficiently. It aims at creating an interactive clone of the real world usable by Cognitive Decision Systems. A multi-modal infrastructure integrating this data model is presented that includes Knowledge Extraction, Knowledge Integration and Knowledge Representation for automatic agent-based decision-making over enriched point cloud data. A knowledge-base processing with ontologies is provided for extended interoperability.

Keywords: 3D Point Cloud, Intelligent Support Systems, 3D Database, Segmentation, Classification, Structure, Ontology, Semantics, Cognitive Decision.

To be published in 2019
Type: Open Access Article
Publisher: Taylor and Francis
Serie: ISPRS Series
Book name: Laser scanning: an emerging technology in structural engineering

3.1 INTRODUCTION

Knowledge extraction (KE), also known as Knowledge Discovery Process [98,99] is the process to mine[9] information and create new knowledge from structured / unstructured data. Specifically, KE-oriented processes such as semantic segmentation and classification permit to extract relevant information regarding an application domain. However, KE is only one step from a global pipeline that aims at creating an interactive space for autonomous decision-making: to represent the world in a form that a computer can use to reason (Figure 10).

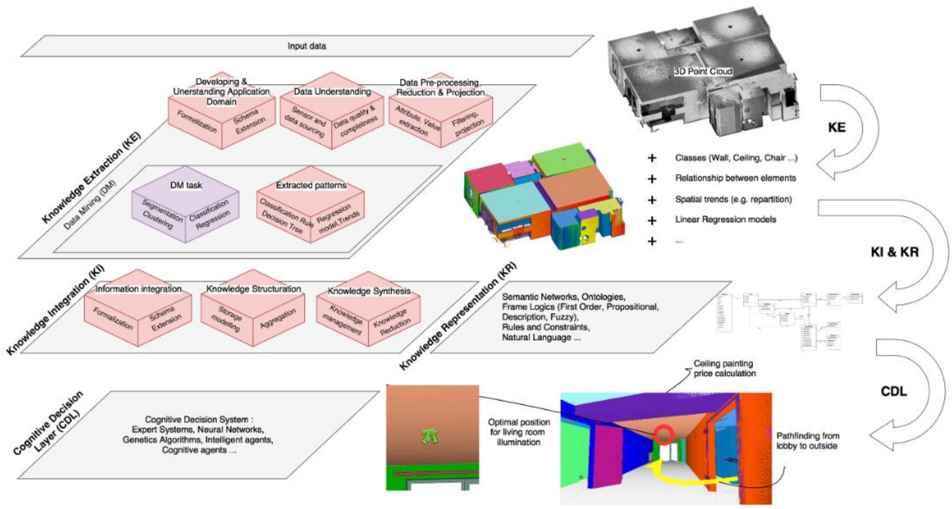


Figure 10. Modular Framework for the creation of an intelligent virtual environment, illustrated over an Indoor Point Cloud. (1) KE; (2) Knowledge Integration; (3) Knowledge Representation; (4) Reasoning from Cognitive Decisions Layer (CDL)

Looking at this modular framework from raw data to Intelligent Environment [100], the processes of integrating, structuring, reasoning and interacting with the data are very challenging especially when dealing with massive datasets from heterogeneous sources. In this chapter, we explore a solution driven by a need in automation (Figure 10) to progressively achieve fully autonomous cognitive decision-making based on 3D digital data.

State of the art in KE-automation applied to point clouds permits nowadays to efficiently extract different information's by means of feature-based algorithms [76] , by using Knowledge-based inference [4] or more actively through machine

[9] The purpose of data mining is to extract knowledge from large amounts of data using automatic or semi-automatic methods.

learning[10] with promising results using neural networks [90,91,101–105] and decision trees [106]. Extracted knowledge often comprises new patterns, rules, associations or classifications and is ultimately useful for one application. But adapting this extra information to be usable within various domains is a great interoperable challenge. A transversal field coverage demands bijective communications through a great generalization and normalization effort.

The processing module defined as knowledge integration (KI) in Figure 10 specifically addresses this task of synthetizing multiple knowledge sources, representations and perspectives often layering multiple domains. This can be decomposed into (1) information integration (i.e. merging information that is based on different schemas and representation models), and (2) synthetizing the understanding of one domain into a common index that keep track of the variance within perspectives. KI specifically addresses integration and structuration of the data (often within database systems), which may resolve conflicts with hitherto assumed knowledge.

Thus, the extended data structure must be translated into an explicit Knowledge representation (KR) to permit a computer to achieve intelligent behaviour, and access knowledge reasoning through a set of logical and inference rules. This is very motivating if you want to deepen the operations made by the computer rather than interpreting on the fly (brain work). Once metadata is attached to 3D data, then you can more easily grasp, or even make calculations impossible before (e.g. in Figure 10 for 3D indoor pathfinding, indoor lighting simulations, reasoning for optimal positioning, extraction of surfaces per room for digital quotations and inventories).

While such a universal solution to decision-making situation is very attractive, KE, KI and KR highly depends on the initial application domain definition, data understanding and its integration within a complete infrastructure. This strains that the underlying data structure must synthetize knowledge through pertinent KR while permitting inference reasoning based on an efficient Cognitive Decision Layer (CDL).

In a first part, we study the attempts, standards and existing reflections to define a common scheme for exchanging relevant 3D information while situating the current integration state of point clouds within these normalizations. Secondly, we present a data model for point cloud structuration that retains knowledge. We then propose a point cloud infrastructure integrating this data model to allow KE, KI and KR following the modular framework presented in Figure 10. Finally, we benchmark such a global solution against several datasets to test its response to established requirements and identify limitations for new research directions.

---

[10] supervised / semi-supervised / unsupervised

3.2 SEMANTICS & 3D DATA

In the context of 3D data, the wide array of applications implies a vast diversity on how the data is used/conceived. This research environment adds important complexities to the integration within any generalized workflow, of which 3D point cloud data are quasi-inexistent. This explicitly demands that data-driven applications enable targeted information extraction specific to each use case. While this is rather convenient for one use case, making a general rule that applies to all models is a daunting task. In this section, we review existing attempts and standards favouring interoperability through well-established 3D spatial information systems (3.2.1), KI reflexions over 3D data (3.2.2) and point clouds (3.2.3).

3.2.1 3D spatial information systems

Datasets that explicitly include spatial information are typically distinguished regarding the data models and structures used to create, manage, process, and visualize the data. In a 3D context, we analogically to [107,108] differentiate three main categories:

- 3D GIS: GIS systems usually model the world itself, retaining information about networks, conductivity, connectivity, topology, and associativity. This enables geospatial analysis, often carried on large collections of 3D instances stored in data warehouses with coordinates expressed in a frame of reference.
- 3D CAD (Computer Aided Design): CAD/CAM techniques model objects from the real world through parametric and triangular modelling tools. The topology is often partial or planar (although vendors extend functions to include semantics and higher descriptive topology [109]) and the data usually plays on a visual scale. The distinction between visualisation and storage is not as clear as in 3D GIS systems, and one file generally describes one complex 3D object. CAD files carry visualization information that is not relevant to the data itself. A simplistic difference consists in thinking of 3D GIS systems as 3D spatial database whereas 3D CAD models are rather related to 3D drawings. The coordinate system is therefore linked to a defined point of interest (often the centroid) in the scene.
- BIM: it constitutes working methods and a 3D parametric digital model that contains "intelligent" and structured data initially for planning and management purposes. It is often studied for its integration with 3D GIS systems with an extensive review in [110], but their parallel evolution (conditioned by temporal and hermetic domain research) and fundamentally different application scopes are slowing down their common assimilation. BIM models share many properties with 3D CAD models,

including their expression of coordinates in a local system, but benefit of a higher semantic integration.

The emergence of new data sources and evolution in data models constantly put in question the suitability of these categorizations. Established and emerging data types and their integration / characterization can become difficult for meeting the characteristics of one of these categories. For example, a more primal spatial data from a more direct data source such as 3D point clouds could benefit of their own category. Indeed, they have a very small direct integration in these groups, but rather serve as a support for the creation of CAD/CAM models, BIM models or 3D GIS systems. In some advanced cases, the information included in 3D point clouds can help extract metadata for the future data model.

However, it is important to note that while barriers between each category was well defined five years ago, the improvement and added functionalities to each category as well as interoperability and integration research and standardization plays a major part into blurring the respective frontiers.

3.2.2 Knowledge integration solutions for 3D semantically-rich data

"*Semantic interoperability is the technical analogue to human communication and cooperation*" [111]. This sentence pertinently summarizes the drive in GIS research to formalize semantics in order to facilitate the communication of data among different communities. Different levels of interoperability exist, and we are looking at the technical parts without looking at societal issues raised by enterprise-oriented information sharing [112]. However, the conceptualization of interoperability in our computerized environment remains a challenge at different levels:

- The nature of concepts that defines interoperability should not arise from simplistic assumptions as notions evolve with time;
- The ever-growing use of 3D data makes it very hard to define a common "language" to be spoken by all professionals;
- The knowledge involved is sparse enough to constraint natural language extension in a computerized formalism;
- Standardization efforts need an international cooperation to represent as thoroughly as possible the reality and benefit of effective coordination;

- Retaining semiotic relationships between concept, symbol and entity, as in the semantic triangle[11].

In a narrower context, 3D data as 3D models are largely used for a high number of applications, which vary in scope, scale, elaboration and representativity. Therefore, semantic schemes as generic as possible provide a potential solution for interoperability.

Ontologies are a good way to explicitly define knowledge in order to address semantic heterogeneity problematics arising from this large variety. However, independent work and research limits their extension to a broader audience especially looking at 3D content. But the rise in usage demands that specific solutions allow 3D data to be exchanged and used as thoroughly as possible. Independent development and uncoordinated actions in the research field of ontologies applied to GIS are addressed by entities such as the World Wide Web Consortium (W3C), the International Organization for Standardization (ISO), the Open Geospatial Consortium (OGC), the International Alliance for Interoperability (IAI) and the rise of open-source developments and repositories. Clarifying standardization processes over 3D data is especially important, with issues arising at both a technical level and a consideration level (how is 3D data considered by the community?)

In general, a standard defines a data model at two levels: properties and geometry. A well-known example is the standard GML3 issued by the OGC which is used by the CityGML data model describing the geometrical, topological, and semantic aspects of 3D city models [113]. The specification and the decomposition in Level of Details (LoD) as well as the current 2.0 version allowing to define semantic concepts has made the integration of city models easier and applicable to a wider range of use cases [114]. Indeed, this gives the possibilities for decision makers to impose a specific "abstraction figure" (LoD1, LoD2 …) that characterizes the granularity level of the wanted geometry and semantic concepts. This interoperability "tool" is a leap forward in the democratization of the standardized data model CityGML. However, its integration with other standards or ontologies is still being discussed and studied, where a discrete number of LoD with 'unconnected' (potentially uneven) levels could be a concern [115]. This illustrates the need to find interoperable systems between already established standards to benefit of higher semantics and topology integration that enhances our comprehension and usability of 3D data.

The Semantic web is a great tool standardized through Semantic Web 3.0 that can create links between already established standards, which encourages the

[11] Attaching meaning to language "objects" is a conceptual phenomenon. As such, this object concept refers to a symbol and conceptualize a "real world" entity that shapes the symbol, and is attached to a social agreement in an information community. [343]

use of web-based data formats and exchange protocols, with the Resource Description Framework (RDF) as the basic format. Indeed, this has the potential to greatly reduce the gap/frontier between each previously defined category in 3.2.1, and better integrate knowledge within 3D spatial data. This is especially efficient if we better integrate 3D point clouds, on which we today derive so many systems and data models. Indeed, in a first time it could serve as transition data but given time it could provide all the necessary information if correctly integrated. Point cloud data large volume and high resolution make it suitable for LoD management and rendering.

Finding hidden pattern and information for knowledge discovery requires complex multi-modal systems as presented in 3.2.3. Data interaction needs flexibility and scalability for different tasks: processing, data management and visualisation. To solve these challenges, both the data model and the storage model must be investigated.

3.2.3 Point cloud solutions for knowledge integration

Few solutions exist for managing semantics and geometry directly on point cloud, demanding new data model to capture key logical aspects of the data structure. On top, the large datasets that point clouds constitute cannot directly fit in the main memory, demanding adapted systems that can exploit efficiently the information regarding a storage model.

We can identify two ways that point cloud can integrate semantics, geometry and topology: through file-based solutions, or through Database Management Systems (DBMS). The storage model used will condition how efficiently the variability and redundancy of the amount of observations is handled. In available DBMS the data store is either as a Block model (i.e. points are grouped in blocks, usually neighbourhoods, which are stored in a database table, one row per block) or a Flat Table Model (i.e. points are directly stored in a database table, one row per point, resulting in tables with many rows). All file-based solutions use a type of blocks model, where points are stored in files in a certain format and processed by solution-specific software. While they are common point cloud storing systems managed through hierarchical-like database models, sharing, compatibility, query efficiency and data retrieval are the main limitations in these solutions. For example, the standard LAS[12] format allows only one byte for user data, making it a very limited choice for managing highly variable knowledge with different perspectives for categorization. Therefore, a fixed schema will quickly become obsolete when trying to cope with thousands of applications.

[12] The LAS file format is the most common file format for the interchange of 3-dimensional point cloud data.

[116] introduced the concept of point cloud database for scientific applications, based on relational tables. Classical Relational DBMS for such application exists, but binary trees limited scalability that struggle with huge datasets size and non-adapted vectorisation and indexation schemes often specific for one usage are hard to exploit on many different servers. Building on this, they point requirements of the structure for analysis of point clouds mainly filtering capabilities, key look-up and Nearest Neighbour's search, cluster analysis, outlier identification, histogram and density estimation, random sampling, interactive visualisation, data loading, insert and updates. [54] extend the concept by defining point cloud data as the third type of spatial representation (the first one being vector data – row like Single Feature Specification – and the second raster data – multipoint object). Their extensive work focus on benchmarking several available commercial Point Cloud Data Management Systems (PCDMS) (block model and flat model of PostgresQL-PostGIS, block model and flat table model of Oracle, flat model of MonetDB, file-based LAStools) to define which one is the most fitted for point cloud management. While some improvements need to be implemented to fix issues in available solutions, each provides a benefit compared to the others, but none can answer efficiently combined queries, data I/O and real-time visualisation. The interoperability stays essential to combine point cloud data with vector data and raster data. They also show in a brilliant way the need of linking user needs, user type with user experience to define a standard in point cloud design and implementation. The NoSQL database robustness to massive data with weak relationship can scale up to many computers but functionalities are today very limited.

Other research work by [53,55,62] present some solution to the integration of domain knowledge through a priori or a posteriori KR in ontologies, but the efficiency and extensibility to production processes depend on the underlying structure for efficient processing, analysis and visualisation. As stated by [53], naïve strategies especially considering query complexity of neighbour search $O(n^2)$ are unrealistic for industrial applications. Indexing techniques provide a solution to storing, compressing and managing the data [54,56,116], but efficiency and extensibility to dynamic semantic update and ontological reasoning stays limited. Queries over octree derived indexing techniques can provide an efficient solution for out-of-core rendering and parallel processing, but data structuration cannot efficiently include context adaptation and inference reasoning.

While these constitute pertinent examples of knowledge and semantic enhancing capabilities, no clear and defined structure is developed. Identifying links and relations within objects of interest becomes essential to truly understand how each spatial entity relates to its surroundings and connecting GIS, CAD and BIM concepts (as seen in 3.2.1) to 3D point clouds through contextual segmentation and object storage. The work of [4,108,117] is a first step in this direction: it proposes a global framework that classify, organise, structure and validate objects detected through a flexible and highly contextual structure that can adapt to three identified knowledge sources being domain, device and analytic knowledge. This lays the groundwork for the development of a new data model – the smart point cloud – that can address previously identified issues while retaining a high level of interoperability with existing standards (3.3).

3.3 THE SMART POINT CLOUD (SPC) DATA MODEL

The Smart Point Cloud concept gives the conceptual tools and definitions for a point cloud knowledge-based structure contextually subdivided according to KE results (namely segmentation and/or classification). It presents a broad framework for the semantic enrichment and structuration of point clouds for intelligent agents and decision-making systems. Our approach infers initial relationships / topology and separate spatial / attribute information to provide efficient data mining capabilities. The domain specialization relies on ontologies to allow high interoperability and specialization through derived Semantic Enrichment Layers [118]. The structuration of the in-base knowledge relies on a categorization first introduced in [117], extended in 3.3.1, and a conceptual model proposed in 3.3.2 and described in 3.3.3, 3.3.4, 3.3.5.

3.3.1 Knowledge categorization

KR and reasoning are the area of Artificial Intelligence (AI) where we study how knowledge can be represented symbolically and manipulated automatically by reasoning engines[13]. Mainly, the use of logic in this context will study entailment relations languages, truth conditions, and rules of inference to enable reasoning. This demands that part of the knowledge be explicitly represented (Knowledge-Base) and constitute the foundation of what the system believes. In order to structure this Knowledge Base (KB) for 3D Point Clouds, we propose a simple yet efficient categorization of knowledge for deriving implicit conclusions from our explicitly represented KB.

The first step toward knowledge integration considers knowledge categorization. In order to cope with the heterogeneity within information's and perspectives, Knowledge for point cloud processing was categorized in 3 branches: device knowledge, analytic knowledge and domain knowledge. The latter constitute what is the closest to domain applications, thus is attached to ontologies of specialization.

[13] Reasoning engines are charged of determining what sorts of computational mechanisms might allow its accessible knowledge to be made available to an agent. What allows humans to behave intelligently is that they can apply their knowledge and adapt/transform it to a new environment to achieve their goals.

3.3.2 Conceptual SPC (Smart Point Cloud) Model

The overview of current practices showed a need to improve automation, data management and interaction. The semantization process relies on geometrical descriptors as well as a domain analogy integrated in a new structuration of the point cloud data through correct indexing techniques. At a higher conceptual level, the creation of an intelligent virtual environment from point clouds is inspired by our cognitive system: recognizing an object means accessing symbolic units stored in a semantic memory and which are abstract from our previous experiences while being independent from any context. Disposing of either digital copies of the real world, invention / conception of "things" to be integrated in the world or a combination of both, we refer to geometries from the "physical space" and "fictional space" (immaterial, concept-based) as in [119]. In their paper, they propose an ontology of space in order to facilitate an explicit definition of CityGML. Extending the formalism, it constitutes a basis for semantic injection into point clouds. However, the study of ethno-physiography as well as human cognition of geospatial information is mandatory for defining information system ontologies. Indeed, the closer (and the richer) the model is to the domain concept, the better (and more extensible) the ontology will be. But the questions of how detailed an ontology should be depend on the levels of interoperability that is envisioned.

The purpose of the SPC characterization is to represent the real world spatially described by point clouds in a computerized form: a user-centered frame representation serving an intelligent environment. The definition of a generic model that applies to a general purpose is very complex, as opening on all domains that benefit from 3D semantically rich models and point clouds range from neuro-psychiatry to economics or geo-information. Our approach was thought to allow a maximum flexibility by defining a conceptualization on which different domain formalization can be attached (Figure 11).

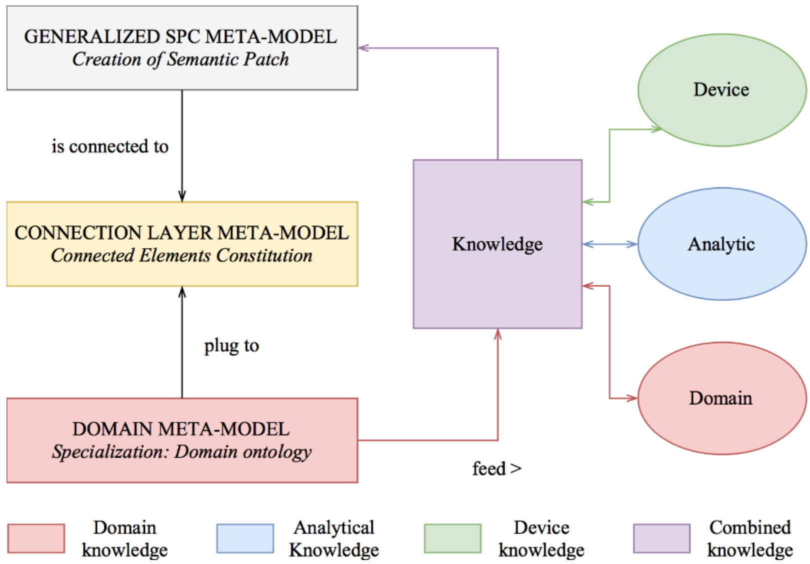


Figure 11. Meta-model articulation for the creation of a SPC

We divided the characterization (KR and data modelling) in different hierarchical levels of abstraction to (1) avoid overlap to existing models, and (2) enhance the flexibility and opening to all possible formalized structure. The core instruction is that the lower levels are closer to a domain representation than higher levels (level-0 being the highest level) but they impose their constraints. The overall structure can be seen as a pyramidal assembly, allowing the resolution of thematic problems at lower levels with reference to constraints formally imposed by the higher levels.

KI is essential to the creation of the SPC structure, as it constitutes the necessary source for the meaning and adaptation of different entities within this pyramidal model. By default, we integrate a core algorithmic module that allows to extract a raw relationship graph based on a voxel element mining routine inspired by [120,121]. This was established as it does not require any external semantic information other than pure spatial information which encourage flexibility and adaptation. However, more domain-verse classification modules such as [81] provide potentially enhanced workflows. As seen beforehand, while Relational DBMS are a great fit logically speaking, they do not perform well considering the very high number of rows. Clustering via indexing-schemes is mandatory for interactive visualisation as well as efficient data loading, inserts and updates. Building this spatial structure over an object-based binary host/guest structure enables powerful analysis and visualization exploitation. In parallel, the ontology-based KR allows inference reasoning/semantics retention and is directly linked to the spatial structure. Thus, it defines relationships and topology both at the point and objects geometrical levels. The top conceptual level, called level-0 gathers data, information, and knowledge about the core SPC components.

3.3.3 *Level-0: Generalized SPC meta-model*

For clarity, we specifically target point clouds, but the model can be extended to all kind of massive gathered data from our physical world, and in an extended version provide an opening for 3D meshes or parametric model integration. The different meta-models are formalised in UML and provide a conceptual definition for implementations. We therefore modelled as a goal to provide a clear vision and comprehension of the underlying system, but the database creation slightly differs to privilege performances, therefore adaptations are made at the relation scheme modelling level.

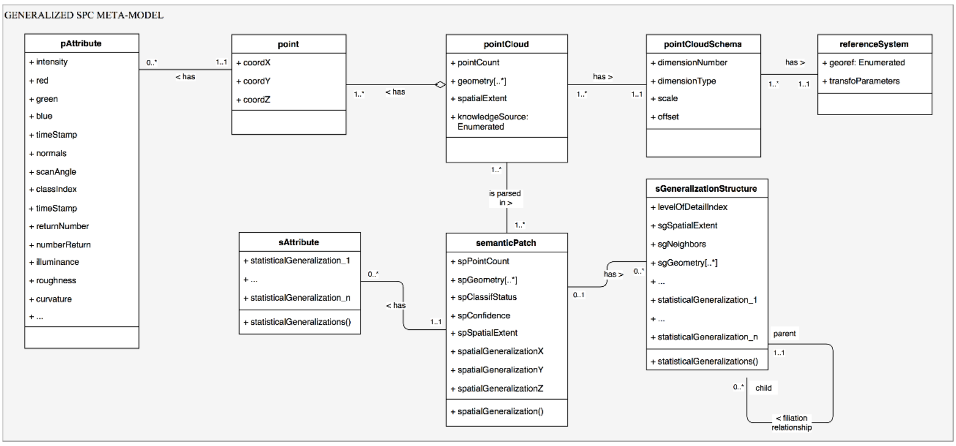


Figure 12. Level-0 Generalized SPC meta-model (UML). A point cloud constituted of points is block-wise organized through semantic patches. These can be pure spatial conglomerate or retain a coherent semantic relationship between constituting points. Generalizations via different schemes are possible using the generalisation structure to provide additional analysis flexibility.

The generalized SPC meta-model (Figure 12) formalizes the core components needed for constituting semantic point patches. It starts with the most primitive geometry: a point. It has a position defined by three coordinates in Euclidean space (R^3): X, Y and Z. Each point has a limited number of attributes, for an example in Figure 12, derived from 3 different sources: device knowledge (scan angle, intensity …), analytic knowledge (normal, curvature, roughness …) or domain knowledge (definition, representativeness …). While the UML model shows a one-to-many relationship, to avoid too many SQL joints and for performances sake, the attributes can directly be integrated within the point table (the same applies to semanticPatch). However, it is important to note that one point can have many sets of attributes (consequently, as does a semanticPatch).

A collection of points sharing the same type of dimensions (spatial and semantic) constitute a point cloud. This is a data-driven aggregation, as depending on the definition of the dataset, the point cloud object parameters will differ. However, one dataset often represents a coherent point aggregation which serves a domain purpose. This point cloud entity also benefits from a knowledge source pointer to identify which knowledge source it relates to (if multiple domain-specific ontologies are connected to the model). To cope with heterogeneity in point cloud sources, a schema is defined and attached to all point clouds that share a similar dimension number, dimension type, scale and offset, similarly to [122]. Each point cloud is then parsed in semantic patches, regarding available knowledge and an adapted subdivision technique. Arbitrary, such a technique could be point-number related, geometry related, or position related. While the existing PostgreSQL plugin Pgpointcloud defining patches in an XML scheme provides spatial patches, we propose to greatly enhance such an approach by constructing semantic patches, which retain both spatial and semantic properties. It constitutes small spatial subsets of points that share a relationship based on available (and injected) knowledge. By default, our proposed voxel-based subdivision method groups point using geometrical & topological properties that implicitly relate to abstract conceptualization of our mind (such as geometric shapes to group points belonging to a plane, others floating above it …). As such, they are indirectly semantically enriched. “semanticPatch” retains many attributes, with an emphasis on two specifics: a classification status (which can be 0: unclassified, 1: one class only or 2: many classes), and a confidence level for the classification. These are computed through a segmentation and classification routine as described in 3.4.1, which is independently developed from the proposed point cloud data model. In order to speed up computations, allow enhanced spatial & semantic searches and provide new generalization possibilities to better address our representation of the data, a LoD generalization structure definition is directly linked to the semantic patches (Figure 13).

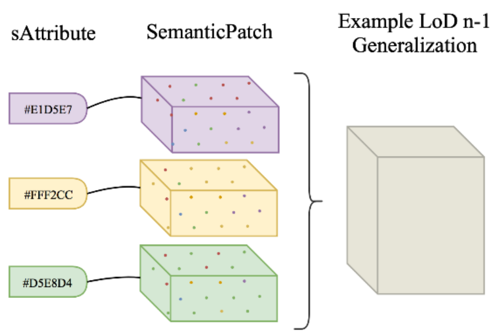


Figure 13. Example of a basic LOD n-1 Generalization of 3 SemanticPatches from a point cloud with colour attributes only.

It defines the indexation scheme used, the different levels (if any), its node spatial extent and neighbours, associated geometries (if any) and other generalized attributes derived from statistical computations (average, Gaussian mixture …). [117] suggests a 3DOR-Tree as defined in [123] for improved performances, but hashing and implicit storage [124] can also greatly improve the internal coherence.

3.3.4 Level-1: Connection-layer meta-model

The connection-layer meta-model (i.e. the strict framework that drives the use of a formalism and resolves any ambiguities about the use of its concepts) plays the role of a plug system: an interface between the core SPC level-0 generalized meta-model, and a domain ontology that formalizes the domain-specialization of a generic ontology. It is constituted of two sub-levels, L1-1 and L1-2 (Figure 14).

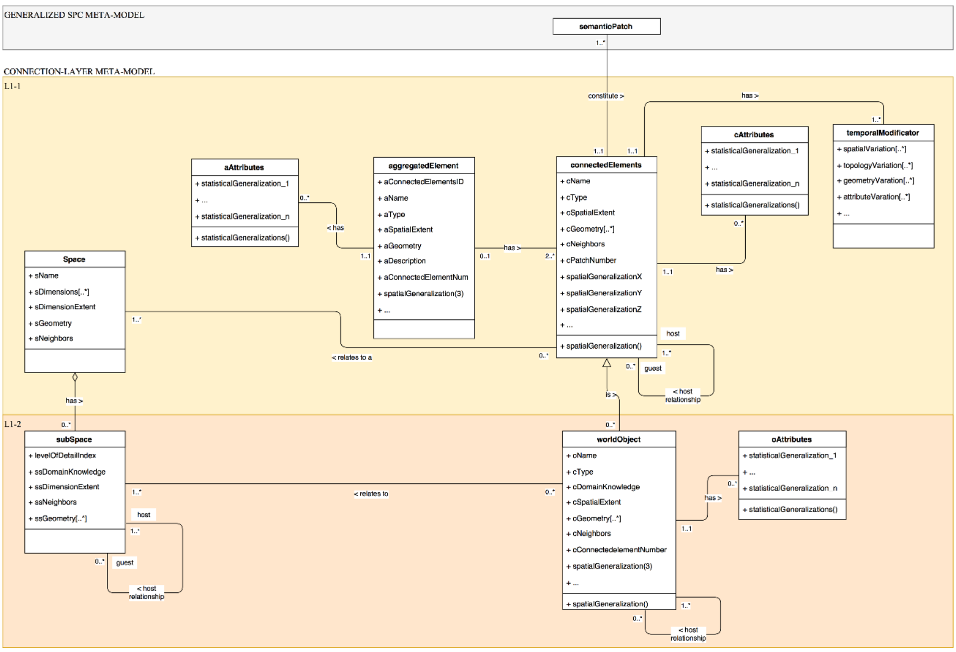


Figure 14. Level-1 Connection-layer meta-model. It is directly linked to the Level-0: semantic patches constitute ConnectedElements. AggregatedElements and topological notions gives flexibility to the deepness of an element characterization. ConnectedElements can relate to one or multiple spaces defined by their dimensions. These are subsequently divided in subspaces regarding a concept from a domain knowledge characterization, similarly to the world objects (being a specialization of ConnectedElements).

The core element in this meta-model is created by an aggregation of one or many patches that define a connected element (ConnectedElements, CEL). These are the entities that closely relate to classified objects, retaining both spatial and metadata coherence. Connected elements transparently describe a portion of the space that is by default indirectly influenced by analytical knowledge and device knowledge, from the underlying patch organization. Connected elements have a spatial extent computed from the aggregation of patches, as well as one or several geometries that can be obtained by topological calculations from the patches. Aside from geometrical attributes including a spatial generalization (which can be for example the barycentre of the spatial extent, but also more representative statistical generalization) they retain raw semantics from the underlying patch aggregation rule dependant of a CEL. The integration of domain knowledge gives the opportunity at this level to deepen the representativeness of a connected element. Nevertheless, one connected element regarding a variety of applications can have different spatio-semantic interests. Therefore, aggregated elements constitute an aggregation of connected elements which provide additional granularity and flexibility (a table, with 4 feet and one horizontal working area, which is either 5 connected elements or one aggregate element). Similarly, each connected element retains relationships with its surrounding environment: we detect and store host and guest relationship information (the table is the guest of the floor, and the floor is the host of the table). These strong concepts have an influence on how deep the selectivity can go. Retaining relations and organizing hierarchically through topological relations refers to mereology, applied on point clouds object generalization regarding DE-9IM [125]. The existing topological relations between 3D spatial objects with internal space are Disjoint, Meet, Overlap, Equal, Contain, ContainedBy, Cover, CoveredBy [126]. Therefore, a double structural definition retaining generalization and point primitives (Level-0) allows new analysis combining multi-LoD definitions. This pyramidal graph relationship formalization permits to easily access a spatially connected graph for reasoning engines that interpret topological relations. These conditions can be used to infer a physical description and combine many possible analysis, for example the possibility to recreate occluded zones, reason about position in time and space, conduct structural investigation... Connected elements also have additional properties and specific attributes inherited from the patches that it relates to. While every point can retain a date stamp, a connected element can be influenced by temporal variations, but duplicating a physical description of the connected element at every discrete temporal interval would not be enough. Therefore, ConnectedElements can have a temporal modifier that will describe the different modifications from the in-base initial state. They can also relate to multiple spaces defined as a set of dimensions in R, R^2… R^n (for example X/Y/Z in R^3).

"Space" and "ConnectedElements" are related to a lower abstraction level L1-2 within the connection-layer. A space can have many subspaces defined in respect to the space dimensions. In a spatial context, it is interesting to note that they are mostly fiat subspaces in regard to [127]. Indeed, bona fide boundaries represent physical separators whereas fiat boundaries will describe a fictional border, and most of subspaces for human cognition have a fictional border (e.g. a room with an open door). The topological inward relation allows to constitute different subspace

Levels of Detail (we can consider a building, or the first floor of that building, or the room 2/43 of that first floor …). “SubSpace” therefore retains a domain knowledge source pointer that can be dedicated to one or many specific domains (it can be a subspace regarding the ontology of buildings, to the archaeology temporal findings in Australia …). The concept of world objects results from the definition of [119], which is a mind conceptualization of an object that also follows the categorization of [127]. “WorldObject” is a specialisation of “ConnectedElements” retaining a domain related semantic pointer similarly to “SubSpace” (a knowledge source mirroring the domain conceptualization). Geometries attached to these entities are useful for topological calculations and the direct link to “SubSpace” allows many possible queries for information extraction (testing the inclusion of a world object in a subspace, testing the intersection of two objects geometries with a fiat boundary from a subspace …). “SubSpace” and “WorldObject” constitute the entry points on which domain ontologies can be associated to adapt to a specific application.

3.3.5 Level-2: Domain adaptation

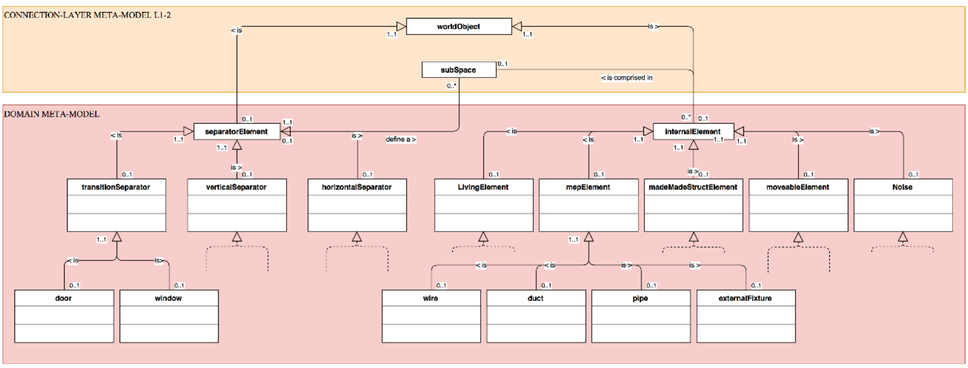


Figure 15. Level-2 meta-model example. SeparatorElement and InternalElement are connected to the Level-1 meta-model directly through “WorldObject” and “SubSpace”. It is a succession of specialization describing an indoor environment.

As stated by [94] *“any fully-fledged system should apply as much domain knowledge as possible, in order to make shape retrieval effective”*. With the rise of online solutions, we have seen a great potential in using knowledge database for classification to analogically associate shapes and groups of points with similar features. This association through analogy *“is carried out by rational thinking and focuses on structural/functional similarities between two things and hence their differences. Thus, analogy helps us understand the unknown through the known and bridge gap between an image and a logical model”* [95]. This introduces the concept of data association for data mining, and relationships between seemingly unrelated data in a relational database or other information repositories. Enabling the use and analysis of domain knowledge through explicit domain assumptions while separating domain knowledge from operational knowledge refers to domain ontologies. This shares interoperability notions with our proposed SPC structure; while one domain meta-model formalization is suited for some applications, another can be more adapted for others and create different results that will be used differently. These will dictate how the final point cloud data model should be used (for which application).

Therefore, the level-2 meta-model is directly linked to different knowledge sources, which are specified in the level-1(-2) meta-model interfaces: "SubSpace" and "WorldObject". Their conceptual abstraction in between pure spatial data (point clouds) and specific domain-verse data constitute a generic door for the potential connection to many level-2 domain specialization. This allows a great flexibility and a context adaptation to a very wide range of application, limited only by the underlying domain ontology. In fine, the domain meta-model attached to the connection-layer meta-model, and indirectly to the generalization meta-model constitute the SPC model.

In a simple example (Figure 15), we illustrate over a basic indoor ontology the connection of a level-2 meta-model to the connection-layer meta-model. It contains 2 class elements (SeparatorElement and InternalElement) specialized in 8 classes (transitionSeparator, verticalSeparator, horizontalSeparator, livingElement, mepElement, madeMadeStructElement, moveableElement, noise) which can also be specialized in a refinement process to get as close possible from the abstract idea that the human mind has of a concrete or abstract object of thought. This crude example is inspired by already established BIM standards and is used on a simple test case to enable rapid perception of natural language requests. As such, one selected "WorldObject" can be specialized and identified as an internalElement, a mepElement (Mechanical, Engineering, Plumbing), specifically a duct of the "subspace" room 4 in the higher LoD level "subspace" building 7 and attached by an "externalFixture" next to the exit "door". The possibility to play on all possible scales is therefore an opening on a flexible system that can be adapted to many real-world applications within an infrastructure.

3.4 FRAMEWORK ARTICULATION & AUTOMATION FOR INTELLIGENT ENVIRONMENTS

The SPC data model permits to structure the information (3D geometry, semantics and topology) in order to leverage knowledge for accessing decision support tools and reasoning capabilities. Indeed, at the frontier between a point cloud GIS system and a spatial infrastructure for agent-based decision support systems, its flexibility allows its extension through new developments mainly in artificial intelligence and machine learning. As such, a modular SPC-based conception permits a high extensibility, with a few constraints described in 3.4.1.

3.4.1 *Infrastructure modularity and extensibility*

The SPC data model integration within a computerized environment was though to first allow an end-to-end usage even with limited automation. Thus, each module as seen in Figure 16 has been developed and implemented in a standard version that allows immediate usage, but which can be upgraded when every defined constraint is met. The choice was oriented to provide an open, functional and evolutionary infrastructure that can easily be replicated / extended. It mainly addresses level-0 and level-1 SPC conceptual model, and for some datasets the level-2 is presented.

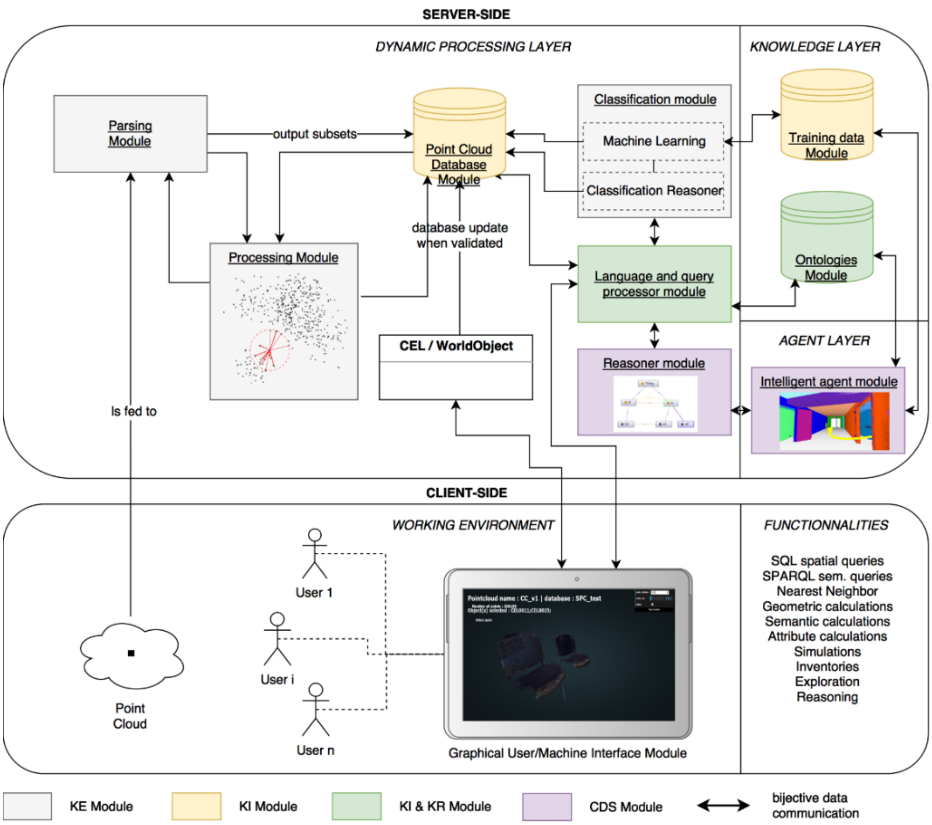


Figure 16. SPC modular framework. The point cloud is fed to the parsing module (KE) that is directly adjusted regarding the processing module (KE). The different point subsets are extracted and injected in the Point Cloud Database module (KI). This module is central and influenced by the Classification module (KE) and the language and query processing module (KI & KR) which are themselves linked to the Knowledge and agent layers.

The Point Cloud is defined by its characteristics that must follow the prescriptions detailed in 3.4.1.1. The Parsing (3.4.1.2) and processing (3.4.1.3) modules are designed for KE and for organising efficiently in SemanticPatches the data which is integrated with other spatio-semantic information in the Point Cloud Database (3.4.1.4). This will constitute the main data repository for the Knowledge processing Engine (3.4.1.5) including Classification, language and query processing, reasoning which is directly linked to the Knowledge management layer described in 3.4.1.7 (Training Data and Ontologies). Finally, when through the GUI (3.4.1.9) an interaction necessitate an agent intervention, the Query engine / reasoner and Agent Layer (3.4.1.8) permits AI-based decision making.

3.4.1.1 Point Cloud characteristics

To be processed and usable by the SPC Infrastructure, a point cloud P must a least be constituted of n points with each three spatial attributes: X, Y, Z. This criterion is the minimal condition to be compatible with the parsing module. If the point cloud has one or multiples attributes a_j, these will be kept, and a specific schema will formally describe the format as an XML document stored in the pointcloud_formats table from the SPC data model level-0.

3.4.1.2 Point Cloud parsing

In order to integrate both classified point clouds and unclassified point clouds, two methods are included in the SPC infrastructure. If knowledge information is available through classification metadata (e.g. .las attributes, ASCII with classification pointers) or file-on-disk organization (.txt file), each independent instance of each class is then divided into a point-based number of separate semantic patches pointing to each instance according to the hardware configuration of the server and the version of the used database. The Point Cloud parsing standard module for raw point cloud data is based on a semantic segmentation framework that groups points in a voxel-based space at a given octree level calculated automatically regarding device knowledge[14]. Each voxel is then studied by analytic featuring and similarity analysis to define a ConnectedElement. This process is conducted regarding an initial connected component from graph representation process after automatically detecting the main “ground” element and different perpendicular/parallel elements that are candidates to wall and ceiling. The voxels containing “edges” or multiple possible points that should belong to separate objects are further subdivided by studying the topology and features with their surrounding elements (Figure 17). Thus, the indexation is defined both spatially and semantically to define SemanticPatches either “pure” voxel or leaf nodes from “hybrid” voxels. While this is very efficient, any decomposition can be though, for example regarding a 3D-OR Tree [60], a Kd-tree [128] or a Sparse Voxels Octree … On the implementation side, the parser module was developed in Python, including the

[14] these were KB determined regarding available product-sheets specifications for each tested point cloud presented in Figure 16.

following libraries: numpy, laspy, tensorflow, scikit-learn, math, 66etwork. The Storage Model is therefore of type Block.

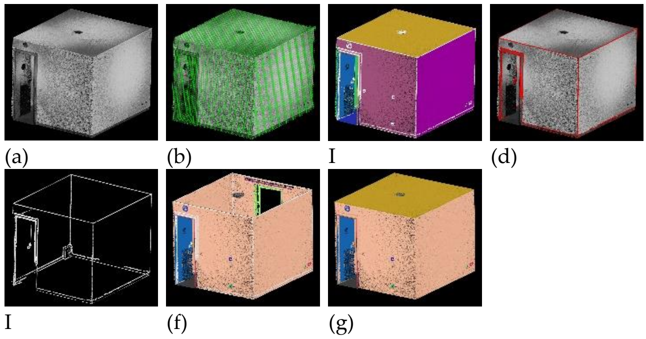


Figure 17. Point cloud parsing methodology: (a) Raw point cloud from TLS; (b) voxelisation at different Octree LoD levels; (c) Segmentation: (d) Voxel-based topology featuring; I Extraction of highly representative points; (f) voxel classification; (g) Connected Element constitution and patch decomposition.

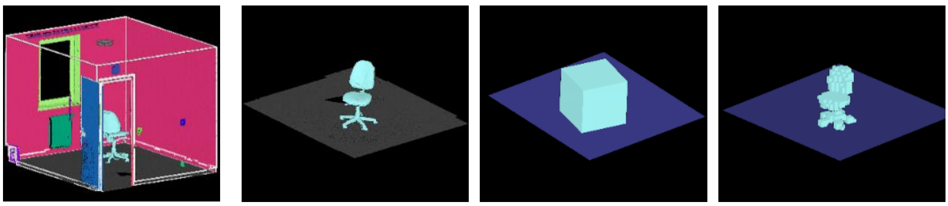
3.4.1.3 Point Cloud processing

The added functionalities play on multi-LoD (point-based or object-based) KE routines. By default, statistical generalizations are computed, as well as a topological skeleton based on voxel adjacency, object information (if available), space decomposition (if available) and in any pertinent space (XYZ, RGB …). This module then rearranges SemanticPatches to corresponding ConnectedElements, and WorldObjects if applicable. Eigen vectors and eigen values are also extracted through Principal Component Analysis, as well as parametric shapes, planarity estimators, shape regularity, Concave/Convex 3D estimator. This is especially important regarding the host/guest topological (8-relationships study) inference used for further reasoning and computations.

This module can directly be improved by joining classification procedures linked to the Knowledge processing engine for added feature computation or ConnectedElements specialisation. The constraints being the update of WorldObjects table, the topology and internal relationships. The standard processing module was developed in Python and C++ and is directly connected to the Point Cloud Database through the psycopg/JSON python library. Some calculations and extraction are directly in-Base (SQL).

3.4.1.4 Point Cloud Database

The Decision Support System is constructed over a Point Cloud DBMS which provides an interface for integrating, updating, and accessing 3D point clouds. In addition, a LoD data structure and indexing scheme is prepared for fast data access by hierarchically subdividing the spatial area or using existing indexes. It provides an access to every component as described in the Conceptual model, permits topology featuring (Figure 18) and allows various modules to be plugged thus playing the role of a centralized KI module. In addition, point/objects attributes resulting from capture, analysis, simulation, or processing stages can be stored. Efficient processing requires a certain data quality that can be ensured by applying Knowledge-Based filtering and registration methods to the input data. The module was implemented in the open-source PostgreSQL 9.6 DBMS, with the extensions PostGIS and Pgpointcloud activated.



(a) (b) I (d)

Figure 18. Example of Point Cloud in-base topology determinations for relationship extraction (a) CEL block storage; (b) host/guest CEL/ground; (c) Generalization and topology inference: (d) Higher LoD topology inference.

3.4.1.5 The Knowledge processing engine

For the training data and ontologies present in the Knowledge management layer to be usable, a processing module make use of this information to allow possible reasoning on point cloud data. As such, it is composed of a reasoner engine that can be used in the case of classification tasks for creating new knowledge (KE). It is also able to extract new information based on data stored in-base. The First Order Logic (FOL) is used for expressing logical conditions. The new knowledge is then stored as attributes and made available through a schema definition's modification. New KI is possible. A Convolutional Neural Network based on PointNet Vanilla and trained using different datasets is developed in python and an indirect connection allows automatic classification based on Machine Learning, using the trained model. Extending this provide a great opportunity for added representativity and automatic processing. The implementation was made using Protégé, the FOL reasoner Pellet and python for the link between RDF/JSON and SQL statements.

3.4.1.6 The Language and query processing engine

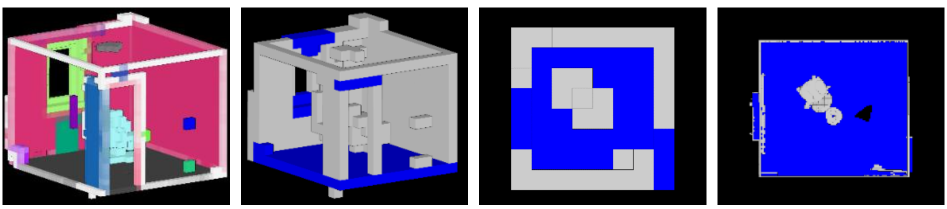
This module allows to navigate between different languages, with the aim to provide a direct interface for natural language processing. The version in the SPC covers the ability to realise SQL and NoSQL queries. The language and query processor are developed in Python, and uses the following libraries: psycopg, json, SPARQLWrapper.

3.4.1.7 The Knowledge management layer

The training data is structured to keep for each classified instance a link to every possible related domain formally expressed in an ontology of Specialization such as provided in [5] . The ontology also permits to formalize rules and constraints for KR. The reasoner used for extracting new information is Pellet, and the developments are made in Protégé, linked to the database through SPARQL queries and connected through an EndPoint, or a server hosting. The main used language is OWL/RDF. KI and specifically Knowledge structuration was constructed to allow efficient KR and reasoning while insuring interoperability with already established models.

3.4.1.8 The Intelligent Agent Layer

The Intelligent Agent Layer indirectly linked to a reasoner through a language processor permits to leverage AI through Experts Systems (E.g. semantic modelling SPC extension in [5]) / Neural Networks, Genetics Algorithms or any agent-based technology for Decision Support Systems. Our tests were conducting with an AI pathfinding agent that uses the Subspace graph connections to establish possible areas to visit. Based on an initial node and a goal node, the agent determines the nodes succession of the least costly path to the goal. We used a heuristic that works for A* returning the distance between the node and the goal. Then, the 3D Geometry representing each subspace is tested against "ST3Dwithin" SQL statement to establish the occupancy grid of any CEL (Figure 19). The implementation was made in python, SQL and JSON.



(a) (b) I (d)

Figure 19. Example of voxel space generalization from 3D Point Cloud for A* pathfinding (a) Raw point cloud from TLS; (b) voxelisation at different Octree LoD levels; (c) Segmentation: (d) Voxel-based topology featuring.

3.4.1.9 The GUI

The GUI is conceived so as to answer the 10 usability heuristics for user interface design described in [129]. Concurrency is also very important, and a platform should be able to scale up to multiple simultaneous connections. As such, a client-side application and RESTful development constitute a good solution for answering efficient interaction and high interoperability. The World Wide Web is a democratized way to share and exchange information. It constitutes a long-term mean to collaborate and is independent of the location which is very important considering the need to be able on site to work with digital copies. The web application is implemented in a WebGL framework and is accessible on any HTML5-compatible browser. The server has different roles and one interaction is to allow the user to view different point clouds. Each point cloud is linked to a corresponding database stored on the server. When the user selects a part of a point cloud, the coordinates are sent to the server which will execute different queries to determine which object is selected by comparing the intersection of these coordinates with the related geometry (e.g. bounding box) of each object in the database. Once the corresponding object is determined, the extension "Pgpointcloud" allows to retrieve in JSON format the data related to each point constituting the patches of the object. When viewing an object, the client makes AJAX calls to the server to retrieve all this data. Once retrieved, the client can then process them to display each point in the point cloud representing the object. The server is developed in Python using the Django library, psycopg2 and json libraries which makes it possible to set up this kind of infrastructure quickly and easily.

3.4.2 SPC Requirements benchmarking

In order to test the suitability of the SPC infrastructure to efficient Decision Support Systems, the following requirements were evaluated:

- (R1) Allow point cloud data loading, insert and updates; Key look-up search, Nearest Neighbour's search and Cluster analysis; Outlier identification, Histogram and density estimation, Random sampling; Filtering capabilities based on spatial attribute and semantics; [116];
- (R2) Visualisation and interaction with in-base data;
- (R3) Support KE processes for semantic segmentation / classification;
- (R4) Possibility to attach semantics to point clouds and point cloud subsets;
- (R5) Support for the interpretation and the aggregation of contextual information;
- (R6) Allow spatial, semantic and topology queries;
- (R7) Allow the user to control, manipulate, search, analyse, query and navigate within the data;
- (R8) Support agent-based inference and reasoning;
- (R9) High interoperability with established data models and extensibility through other modules;

The implementation's choice toward a client-server infrastructure (Figure 20) was thought to allow point clouds and databases storage on a server. In addition, the application can manage multiple point clouds and databases. All data is therefore centralized with the application, developed and tested on Linux (Ubuntu) and Windows 10, using the described frameworks, languages and libraries in the previous section. The application hierarchy is easy to set up, and administrator sessions for editing, adding or deleting application data are automatically implemented. For example, each object in a point cloud is automatically stored in the application database when a user selects it.

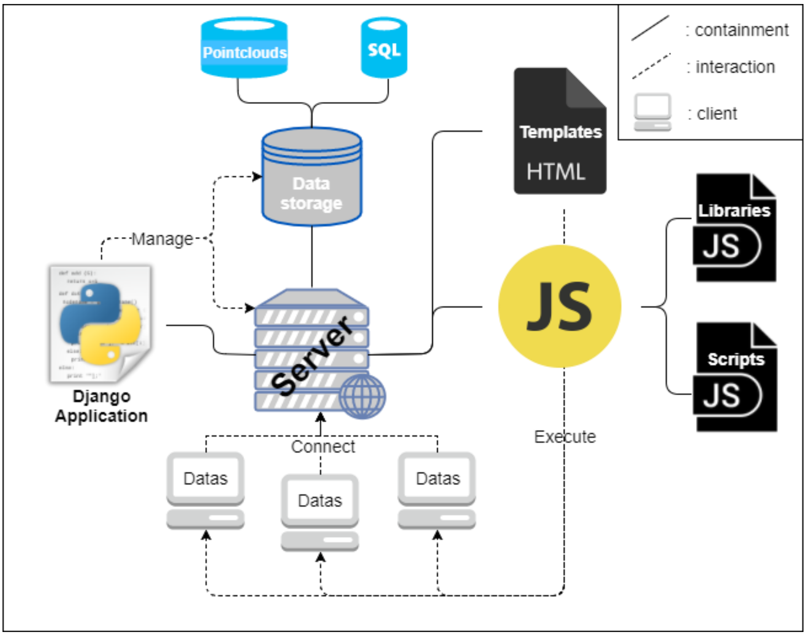


Figure 20. SPC Client/Server Infrastructure implementation

To test the suitability of the infrastructure to point clouds with varying characteristics, the following datasets (Table 3) from different sensors/methodologies were integrated:

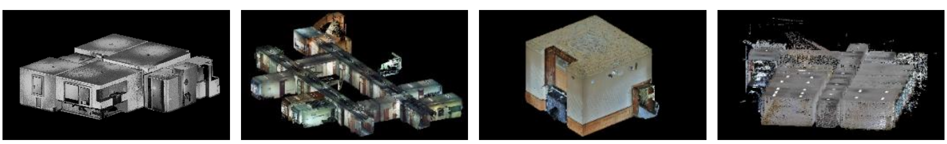
Table 3. Datasets for benchmarking the SPC Infrastructure

PCID	PCType	Sensor / Instrument	Point Number (Million)	Attributes Number	Classification Status
1	TLS	Leica P30	200	7	Unclassified
2	TLS	Trimble TX5 (eq. Faro Focus 3D)	150	8	Unclassified

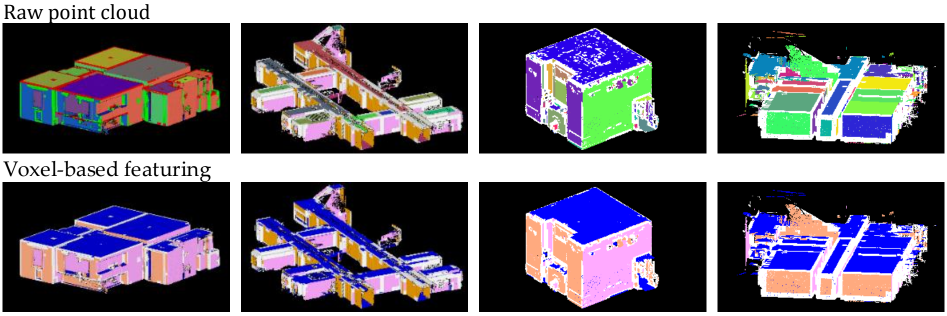
3	PS CMOS	Canon 5D Mark III + 24-104 mm	200	6	Unclassified
4	PS CMOS	Canon 70D + Fisheye Sigma 15 mm	450	6	Unclassified
5	IS	Camera Matterport Pro 3D	50	3	Classified
6	2DS+S	Zeb Revo	25	3	Unclassified
7	LiDAR	Riegl Litemapper 6800i	70	15	Classified
8	2DS	Velodyne	1	3	Classified
9	2DS+S	NavVis	8	7	Unclassified

Note: TLS= Terrestrial Laser Scanner, PS CMOS= Passive Sensor CMOS, IS= Infrared Scanner; 2DS+S= 2D Scanner + SLAM

Firstly, we notice that the SPC can directly integrate, all tested point cloud with different characteristics while addressing (R1). Mainly, the characteristics of the point cloud data influence the initial Connected Element detection for non-classified datasets. We also noticed that the quality/representativity of the point cloud data can impact the results. The most influential factors are the irregular point distribution, the point accuracy and the return signal dependence on the physical characteristics of the surface. We observed that the datasets with high noise or with complex structures (indoor and outdoor data) can become problematic for the Parsing module. However, Device Knowledge-based filters and characterization correct this phenomenon and when the sensor related device knowledge is formalized in the Knowledge Layer, it has a limited impact.



Raw point cloud



Voxel-based featuring

Grouping before topology

(a) PCID 1 (b) PCID 2 I PCID 3 (d) PCID 7

Figure 21. Visualisation of four point clouds from the benchmarking datasets and visual impact of the voxel-based featuring and grouping before topology

Requirements (R2), (R3), (R4) are also validated, with a manual interaction step when the classification module doesn't permit automatic classification. Through the GUI, a python script on the server is executed. It first determines the selected ConnectedElement object according to the x, y and z coordinates of the ray-casting. To do this, an SQL query is carried out on the database linked to the point cloud and allows to return the name of the object containing the point (e.g. Figure 21). (R7) is answered but the implementation could be extended by providing higher interactivity.

Requirement (R5) for contextual aggregation is limited to established EndPoints such as DBPedia. The integration demands that the concepts definitions are the same as the one in the different accessible ontologies. However, as the WWW standards evolve, and new semantic web resources are available, it can be extended.

Using PostGIS, Python, SQL, SPARQL statements, information extraction is possible. This gives the SPC infrastructure the ability to fully answer requirement (R6), independently from the initial tested datasets. Problem arises when the generalization geometry is not representative of the described objects (e.g. when noise is too prominent with an incorrect 3D Concave Hull / BB extracted).

Requirement (R8) depends on the depth and completeness of the semantic definition of objects. Indeed, the classification granularity may limit the available operations to a fraction of what is possible due to a lack of specialisation. For example, A* pathfinding is possible in all datasets, but for the PCID5, the use of SubSpace and objects definition information gives an additional edge to the deepness of the AI-pathfinding. The reasoning module was tested over unclassified datasets, as a classification reasoner, and allowed to automatically detect shapes regarding available knowledge over geometrical and radiometric properties of the point cloud [4]. Extending the module by providing domain ontologies and inference rules would provide very interesting for adapted domain-versed classification.

Therefore the (R9) condition is important and was tested. The addition of knowledge “pointers” to the in-base data permits to use this flexibility concerning every point cloud. The possibility to connect to EndPoints also permits a higher interoperability by gathering established knowledge from recognized knowledge repository (e.g. DBPedia). As such, information regarding classified elements can directly be linked to extend instance information. This can also be used by agent support system for inference and better guidance toward decision-making and simulation.
By combining the Point Cloud Database module with the Knowledge Layer, the SPC infrastructure allows the use of operators either spatial, semantic, topology-based or any combination of them. This was tested over the different datasets and the reasoner engine “Pellet” was used to infer new knowledge. Purely spatial operators possess spatial semantics that can be classified as 3D directional operators such as above, under, NorthOf, SouthOf, EastOf, WestOf; metric operators such as distance analysis; topological operators such as touch, contain, equal, inside; Boolean operators such as union, intersection. The ability to play with the generalization possibility provide a very high flexibility regarding possible data analysis and cognitive decision making.

Table 4. Example of Basic SQL statements over the SPC infrastructure

Goal	SQL Statement
I want to select the ‘semanticpatches’ which intersects a defined 2D polygon	SELECT pa FROM semanticpatch WHERE ST_INTERSECT(pa::geometry) = TRUE
I want to select all ‘semanticpatches’ that have been classified	SELECT pa FROM semanticpatch WHERE spclassifstatus = 1
I want to select the connected element CEL0065	SELECT pa FROM semanticpatch WHERE connectedelement_id = 0065
I want to extract the name of the connected element(s) that include the (x,y,z) position (used through ray-casting)	SELECT name FROM connected_elements WHERE((ST_Z(x, y, z) > ST_Zmin(geom::box3d)) AND (ST_Z(x, y, z) < ST_Zmax(geom::box3d)) AND (ST_Y(x, y, z) > ST_Ymin(geom::box3d)) AND (ST_Y(x, y, z) < ST_Ymax(geom::box3d))AND (ST_X(x, y, z) > ST_Xmin(geom::box3d)) AND (ST_X(x, y, z) < ST_Xmax(geom::box3d)));

The integration of semantics, e.g. “The connected element CC0065 is a desk chair named ‘comfychair’ made in 2018-10-08 for working in front of a PC” as SQL statements (INSERT INTO moveableelement_connectedelement_id, type, title, date_prod, kind) VALUES (‘65’, ‘chair’, ‘comfychair’, ‘2018-10-08’, ‘working’) from KE-routines and the Knowledge Layer permits to achieve very interesting analysis. E.g., the Natural language “I want to locate the highest table within the room A of the Building B and calculate the free space between its surface and the ceiling for determining the possible extension” leveraging the linked domain concepts from a level-2 meta-model mirror our real-world information gathering, thus heavily extend possibilities.

3.5 LIMITATIONS, PERSPECTIVES AND POSSIBILITIES

While the SPC infrastructure can play the role of intermediary between real-time acquisition and inference for decision making without the need to denature point cloud data, the infrastructure presents some limitations that suggest new research directions.

Any modelling choice is arbitrary and depends on the conscious or unconscious aspirations of the designer. Although our work responds to a concern for generalization at a spatio-semantic level, it nevertheless remains that it is not totally independent of a certain context. It is for this reason that we wanted to clearly illustrate a privileged domain of application: indoor environments (for BIM, emergency response, inventory management, UAV collision detection …). This choice permits to explore different scales and configurations for deeply and entirely testing our developments. It is also ideal for the definition of new virtual spaces, and the GIS demand associated to such environment is ever increasing. Therefore, as the formalization of domain constantly evolve, modelling and direct integrations of level-2 domain meta-models remain important.

While 3-dimensional spaces are strongly inferred in the SPC model, 4-dimension spaces integrating time or by extension n-dimensional spaces are possible characterizations for greater interoperability. Tests were conducted with static point data only, but varying positions in space and time present additional problematics that could be address through new modules or an extended SPC data model. While the developed framework constitutes the groundwork of a modular infrastructure that provide direct integration of hard-coded or inferred domain knowledge, future work including the extensibility of the proposed model to other data types, as well as a better integration of learning routines and ontologies as knowledge sources is very interesting.

The SPC infrastructure permits to link efficiently AI through agent decision support systems and reasoning. While developments were carried in a 3D voxel space (passable or non-passable at each point) with gravity constraints, using CEL Bounding-Boxes or an advanced Navmesh could enhance results. This illustrate that finding an efficient AI-based agent can become complicated for one application, so generalizing software agents can become complicated. To have an intelligent agent that performs reactively and/or pro-actively, interactive tasks need to be tailored to a user's needs without humans or other agents telling it what to do. To accomplish these tasks, it should possess the following general characteristics in regard to [130]: (1) Independence, (2) learning, (3) Cooperation, (4) Reasoning, (5) Intelligence.

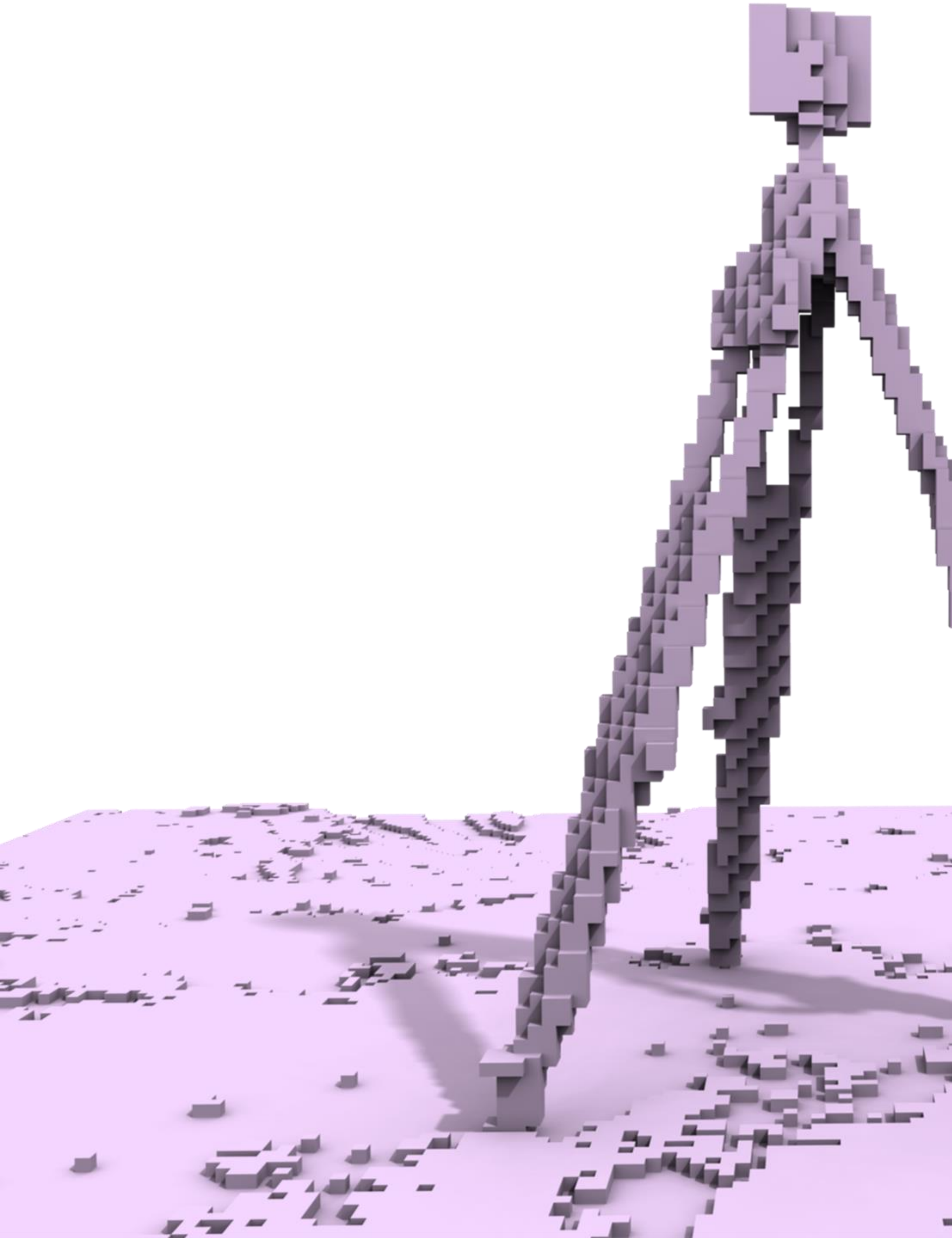
This relates to the exploratory research field of for Artificial General Intelligence[15] [131] which explicitly justify a need of virtual environments that incentivize the emergence of a cognitive toolkit, such as the SPC infrastructure. As such, SPC-based multi-agent environments provide an opening thanks to its variety (the optimal strategy must be derived optimally) and natural curriculum (the difficulty of the environment is determined by the skill of other agents). This direction while being extremely exciting to avoid purpose-specific algorithm is still a research question that need to be further explored, in which Deep Learning may provide a suitable answer.

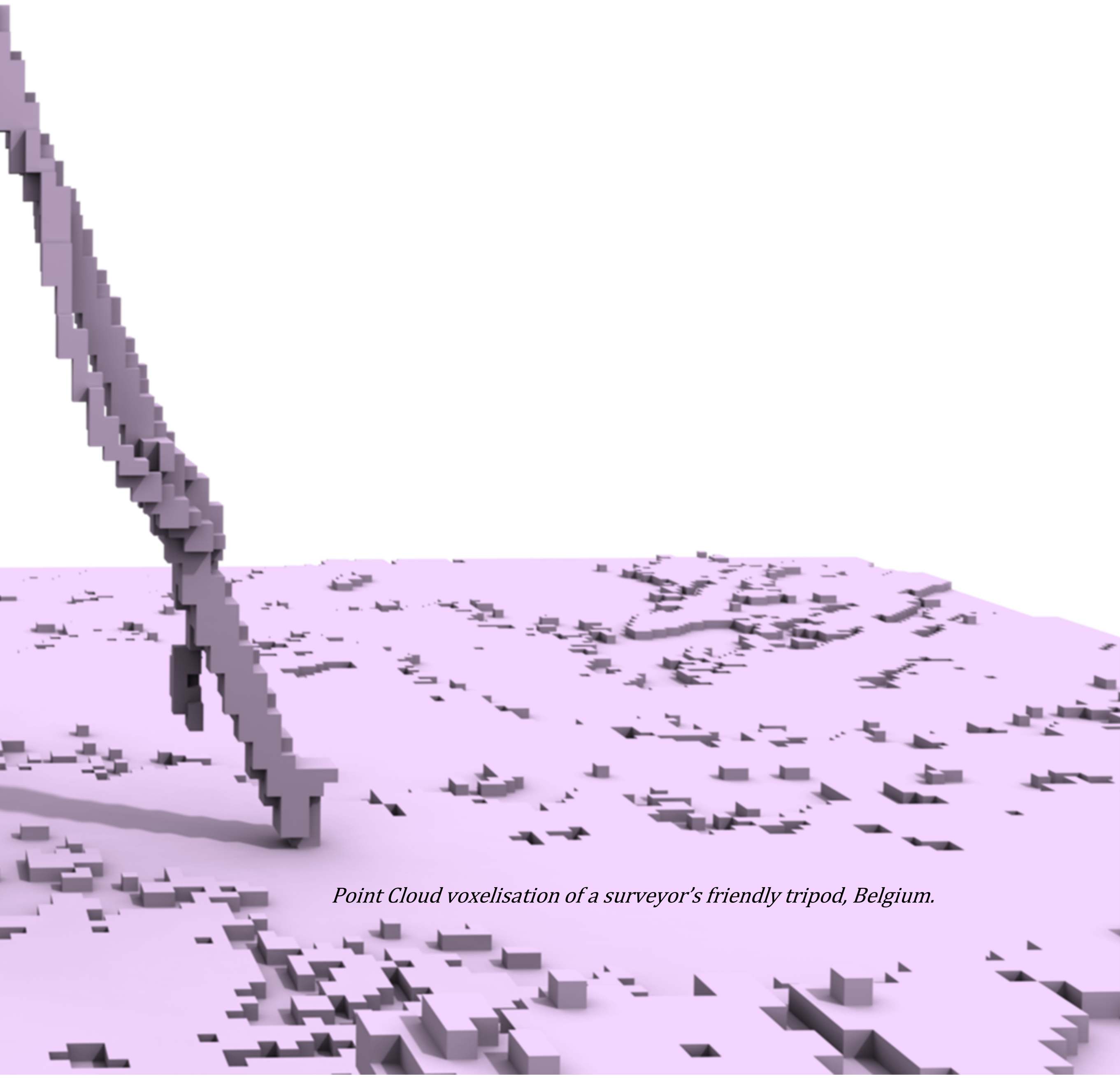
3.6 CONCLUSION

The Smart Point Cloud data model permits to structure the information (3D geometry and semantics) to leverage knowledge for accessing decision-making support tools and reasoning capabilities. At the frontier between a point cloud GIS system and a spatial infrastructure for agent-based decision support systems, its flexibility allows to evolve with new developments mainly in artificial intelligence and machine learning. The proposed modular infrastructure includes Knowledge Discovery processes with Knowledge Integration and Knowledge Representation as ontologies, proving efficient context-specific adaptation. Nine point cloud datasets were used for testing the infrastructure, successfully answering identified needs and providing new research directions as modular extensions.

---

[15] Artificial General Intelligence is the intelligence of a machine that could successfully perform any intellectual task that a human can do, including: reasoning; judgment calls under uncertainty; KR; planning; learning; natural language communication. Other important capabilities include the ability to sense (e.g. see) and the ability to act (e.g. move and manipulate objects) in the world where intelligent behaviour is to be observed.





Point Cloud voxelisation of a surveyor's friendly tripod, Belgium.

CHAPTER 4

Knowledge Extraction

Chapter 4 - Knowledge Extraction 79

Chapter's preface 81
4.1 Introduction 84
4.2 Related works 86
4.2.1 Point cloud feature extraction 86
4.2.2 Semantic segmentation applied to point clouds 87
4.3 Materials and Methods 90
4.3.1 Voxelisation grid constitution 91
4.3.2 Feature Extraction 92
4.3.3 Connected Element constitution and voxel refinement 99
4.3.4 Graph-based semantic segmentation 101
4.4 Dataset 103
4.5 Results 105
4.5.1 Metrics 105
4.5.2 Quantitative and qualitative assessments 106
4.5.3 Implementation and performances details 113
4.6 Discussion 114
4.6.1 Strengths 115
4.6.2 Limitations and research directions 116
4.7 Conclusions 117

CHAPTER'S PREFACE

In the previous chapter 3, we established the fundamentals of the Smart Point Cloud Infrastructure (SPCI) and its modular architecture. The present chapter 4 will give extended details on the Parsing module and the Classification module as highlighted in Figure 22.

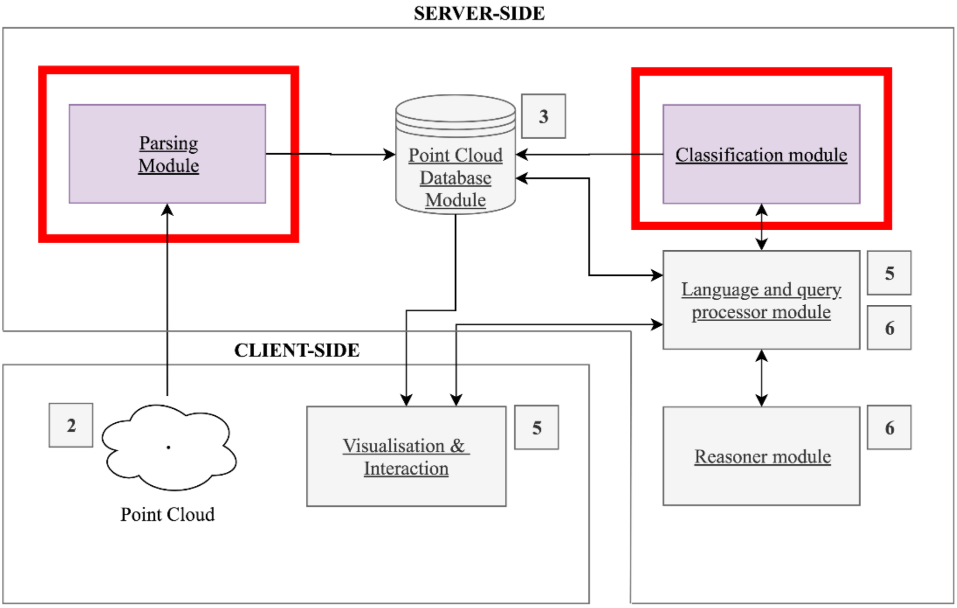


Figure 22. Chapter 4: Knowledge Extraction

In this chapter, we provide a general clustering approach that groups point in relatively low-specialization segments which can then be further refined depending on the application at hand. We then provide a naïve knowledge-based classification approach for indoor point cloud data and asset management applications. These both permit to leverage the SPCI architecture. This chapter holds targeted references as seen in chapter 2, namely feature extraction and segmentation / classification approaches. The following chapters, 5 and 6 are then building on the initial spatial structuration provided by the Parser module.

Based on Article [5]

Voxel-based 3D point cloud semantic segmentation: unsupervised geometric and relationship featuring vs deep learning methods

Abstract: Automation in point cloud data processing is central for knowledge discovery within decision-making systems. The definition of relevant features is often key for segmentation and classification, the main challenges in automated workflows. In this paper, we propose a voxel-based feature engineering that better characterize point clusters and provide a strong support to supervised or unsupervised classification. We provide different feature generalization levels to permit interoperable frameworks. First, we recommend a shape-based feature set (SF1) which only leverage raw X, Y, Z attribute of any point cloud. Then, we derive relationship and topology between voxel entities to obtain a 3D structural connectivity feature set (SF2). Finally, we provide a knowledge-based decision tree to permit infrastructure-related classification. We study SF1/SF2 synergy on a new semantic segmentation framework for the constitution of a higher semantic representation of point clouds in relevant clusters. We finally benchmark the approach against novel and best-performing deep-learning methods using the full S3DIS dataset. We highlight good performances, easy-integration and high F_1-score (>85%) for planar-dominant classes comparable to state-of-the-art deep learning.

Keywords: 3D point cloud; voxel; feature extraction; semantic segmentation; classification; 3D semantics; deep learning

To be published in 2019
Type: Open Access Article
Publisher: MDPI
Journal: ISPRS International Journal of Geo-Information
Special Issue: Point Cloud Feature Extraction

4.1 INTRODUCTION

Extracting knowledge from raw point cloud data is actively driving academic and industrial research. There is a great need for automated processes that can speed up and make existing frameworks faster and more reliable. It often integrates a classification step to extract relevant information regarding one application domain. However, one classification approach cannot efficiently satisfy all domains as the semantic concepts attached to objects and location vary depending on uses (e.g. considering a chair as an object, or its legs). Therefore, insuring that such information is transferable to benefit other applications could provide a great opening on point cloud data usage. Yet, this is a non-trivial task which necessitate highly interoperable reasoning and a flexible way to handle data, relationships and semantics. Our method considers the Gestalt's theory [132], which states that the whole is greater than the sum of its parts, and that relationships between parts can yield new properties/features. We want to leverage the human visual system predisposition to group sets of elements.

In this paper, the first goal is to provide a point cloud parsing unit to extract semantic clusters through a voxel-based partitioning of the dataset. It permits flexible usage in different domains such as Architecture, Engineering & Construction (AEC), Building Information Modelling (BIM), Facility Management (FM), indoor navigation and robotics. The module acts as a standalone within a Smart Point Cloud Infrastructure [2] – a set-up where point data is the core of decision-making processes – and handles point clouds with heterogeneous characteristics. Indeed, for the sake of interoperable data management, the possibility to incorporate Knowledge-Extraction routines in existing frameworks has become essential for an efficient international research cooperation. As such, we investigate an objective solution for versatile 3D point cloud semantic representation transparent enough to be usable on different point clouds and within different application domains. We propose to structure a point cloud in Connected Elements further refined in Semantic patches using efficient and low-level voxel-related features. This is primarily motivated by the limitations of point-based approach where the amount of data, the redundancy and the absence of relationships within points are great performance issues.

In order to assess the possibilities given by the 3D clustering scheme, a semantic segmentation solution is developed to leverage feature sets retaining both shape and relationship information. This permits to benchmark the performances and results against the best-performing state of the art deep-learning methods. Indeed, with the rise in computing power, promising machine learning techniques detailed in [39,45,91–93,83–90] are a great opening to more reliable and robust 3D objects classification. However, ground-truth extraction and dataset labelling to create training data are the main drawbacks in supervised learning. Manually annotating and insuring the quality of such datasets is a heavily dauting task. Hence, ways to alleviate these mechanisms through automated tools are essential to new findings and for training new models.

The experiments were conducted on the full S3DIS [133] indoor dataset as (E.g. Figure 23), but it is generalizable to outdoor environments with man-made objects/characteristics.

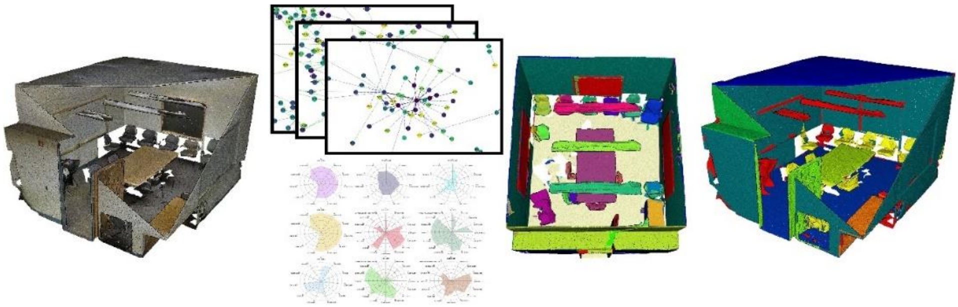


Figure 23. Voxel-based 3D semantic segmentation. From left to right: Raw point cloud, feature engineering, Connected Elements extraction, Classified point cloud

Briefly, this paper makes the following three main contributions:

- A new interoperable point cloud data clustering approach that account variability of domains for higher-end applications;
- A novel point cloud voxel-based featuring developed to accurately and robustly characterize a point cloud with local shape descriptors and topology pointers. It is robust to noise, resolution variation, clutter, occlusion and point irregularity;
- A semantic segmentation framework to efficiently decompose large point clouds in related Connected Elements (unsupervised) specialized through a graph-based approach: it is fully benchmarked against state-of-the-art deep learning methods. We specifically looked at parallelization-compatible workflows.

The reminder of this paper is structured as follows. Section 4.2 briefly reviews recent related works dealing with point cloud feature extraction, segmentation and classification. Section 4.3 gives the details of the proposed voxel-based featuring and semantic segmentation. In Section 4.4, we present the S3DIS dataset used for the different experiments and benchmarks. In Section 4.5 we study the impact of features over the results and analyse the performance of the approach against high-yielding supervised learning. In Section 4.6 we discuss our findings and highlight limitations as research directions.

4.2 RELATED WORKS

Feature design occupies a central position to knowledge representation and classification approaches. As expressed in Section 4.1, the Gestalt's theory [132] is fundamental to understand how our visual cognition systems perceive our surrounding when trying to feed a classifier with information. While many factors make intuitive sense (E.g. Figure 24), it is often very hard to translate them into algorithms.

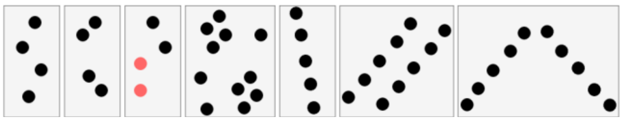


Figure 24. Visual patterns on points from left to right: Not grouped; Proximity criterion; Similarity criterion; Common cluster region; Linear criterion; Parallel criterion: Symmetry criterion.

This gives an edge to deep learning approaches where the emphasis is toward training dataset's constitution rather than feature engineering. In this section, we cover both problematics, i.e. feature- engineering and point-cloud supervised learning, which is further linked to Section 4.3 and 4.5. First, features and methods that work well for extracting relevant information from point clouds are investigated. Then, relevant references and recent works (2015+) that deals with point clouds semantic segmentation are given to the reader. We specifically look at voxel approaches and features that already made their proof over complex point cloud artefacts.

4.2.1 Point cloud feature extraction

In this sub-section, we analyse low-level shape-based approaches that try to extract local descriptors from 3D neighbourhood [134]. We refer initially to the pertinent work of Ghorpade et al. [40] which proposes a review of 2D and 3D shape representation and is a good introduction to get an idea of the landscape of features in use.

The work of Bueno et al. [135] focuses on the detection of geometric key-points and its application to point cloud registration. The authors primarily study data subsampling to keep key points for coarse alignment purposes. These points are obtained using an approach mainly based on features being eigen entropy, change of curvature and planarity. Indeed, they state these provide a good representation in both, visual and mathematical value of the point clouds. This is found in many recent works such as [136], where authors also use local eigen-based features for disaster damage detection through synergistic use of deep learning. The work of Blomley et al. [137] provides larger insights on the common geometric (e.g. eigen-based) covariance features in varying scale scenarios. In 2018, Thomas et al. proposed a semantic classification of 3D point clouds in [39] which also employs eigen-based features as well as colour derived feature. The specificity lies in the definition of a multiscale neighbourhoods which allows the computation of features with a consistent geometrical meaning. The authors in [138] also uses several eigen-based feature, spectral and colour-derived features for the classification of aerial LiDAR point clouds. The features coupled with their approach provides good results, and therefore orient our choice of features toward eigen-based features, for they representativity of local neighbourhood as well as low-knowledge requirement.

Other recent works for learning local structures [139] or local shape properties [90] highlighted the wide acceptation of normals. Shen et al. present in [139] two new operations to improve PointNet [101] – one of the earliest deep learning reference for point cloud semantic segmentation – with a more efficient exploitation of local structures. The first one focuses on local 3D geometric structures. In analogy to a convolution kernel for images, they define a point-set kernel as a set of learnable 3D points that jointly respond to a set of neighbouring data points according to their geometric affinities measured by kernel correlation. The second one exploits local high-dimensional feature structures by recursive feature aggregation on a nearest-neighbour-graph computed from 3D positions. They specifically state that *"As a basic surface property, surface normals are heavily used in many areas including 3D shape reconstruction, plane extraction, and point set registration"* [140–144]. The paper of Song et al. [145] provide a comparison of normal estimation methods, which can also be achieved via neural networks such as PCPNet [90]. In this last article, Guerrero et al. propose a deep-learning method for estimating local 3D shape properties in point clouds. The approach is especially well-adapted for estimating local shape properties such as normals (both unoriented and oriented) and curvature from raw point clouds in the presence of strong noise and multi-scale features. Therefore, we will specifically integrate normal within our workflow, while looking at performance issues during its computation.

Edge-based features have also been investigated in [146] or [147] but their applicability is mostly oriented toward point cloud line tracing. Thus, we confront large performance issues due to analysing geometric properties of each point's neighbourhood, and combining RANSAC [148,149] and angular gap metrics to detect edges. While extended in [150] to contour extraction of large 3D point clouds, we will specifically avoid region growing approaches due to performances limitations.

4.2.2 Semantic segmentation applied to point clouds

The first challenge in pure segmentation frameworks is to obtain group of points which can describe with enough detachment the organization of the data by a relevant clustering. A first approach using relationships while conserving the point-based flexibility is given by the work of Papon et al. [146]. They propose an over-segmentation algorithm using 'supervoxels', an analogue of the superpixel approach for 2D methods. Based on a local k-means clustering, they try and group voxels with similar feature signatures (39-dimensional vector) to obtain segments. The work is interesting because it is one of the earliest to try and propose a voxel-clustering with the aim to propose a generalist decomposition of point cloud data in segments. Son et Kim use such a structure in [151] for indoor point cloud data segmentation. They aim at generating as-built BIMs from laser-scan data obtained during the construction phase. Their approach consists of three steps: region-of-interest detection to distinguish the 3D points that are part of the structural elements to be modelled, scene segmentation to partition the 3D points into meaningful parts comprising different types of elements using local concave and convex properties between structural elements, and volumetric representation. The approach clearly shows the dominance of planar features in man-made environments.

Another very pertinent work is [121] which propose a SigVox descriptor. The paper first categorizes object recognition task following the approach as (1) model-fitting based (starts with segmenting and clustering point cloud followed by fitting point segments); (2) semantic methods (based on a set of rule-based prior knowledge); (3) shape-based methods (shape featuring from implicit and explicit point clusters). They use a 3D 'EGI' descriptor to differentiate voxels extracting only specific values from a Principal Component Analysis (PCA) [152]. The approach proves useful for MLS point clouds, grouping points in object candidates, following the number. Another voxel-based segmentation approach is given in [153,154] using a probabilistic connectivity model. The authors use a voxel structure of which they extract local contextual pairwise-connectivity. It uses geometric "cues" in a local Euclidean neighbourhood to study possible similarity between voxels. This approach is similar to [155] where authors classify a 2.5D aerial LiDAR point cloud multi-level semantic relationships description (point homogeneity, supervoxel adjacency, class-knowledge constraints). They use a feature set among other composed of the elevation above ground, normal vectors, variances and eigen-based features. Another analogous approach can be found in [156] for building point detection from vehicle-borne LiDAR data based on voxel group and horizontal hollow analysis. Authors present a framework for automatic building point extraction including three main steps: voxel group-based shape recognition, category-oriented merging and building point identification by horizontal hollow ratio analysis. This article proposes a concept of "voxel group" where each group is composed of several voxels that belong to one single class-dependent object. Then the shapes of point clouds in each voxel group are recognized and this shape information is utilized to merge voxel group. This article leverages efficiently a sensory characteristic of vehicle-borne LiDAR building data but specialize the approach in consequence.

The references [157,158] are built upon a graph-based over-segmentation methodology composed of a local 3D variation extraction, a graph construction, descriptor computation and edge-wise assignment followed by sequential subgraph criteria-based merging. The used descriptors are mainly RGB, location and normal vectors on top of the fast point feature histogram [159]. While the approach is domain-related, it offers some additional insight on the power of relational approaches between local point patches for the task of semantic segmentation.

However, as show in [86], using a multi-scale voxel representation of 3D space is very beneficial even for complexity reduction of terrestrial lidar data. The authors propose a combination of point and voxel generated features to segment 3D point clouds into homogenous groups in order to study surface changes and vegetation cover. The results suggest that the combination of point and voxel features represent the dataset well, which shows the benefit of dual representations. The work of [134] uses Random Forests for aerial Lidar point cloud segmentation which aim at extracting planar, smooth and rough surfaces, classified using semantic rules. This is interesting to answer specific domains through ontology formalization.

These methodologies contrast with deep learning approaches as they try to solve the semantic segmentation problem by first understanding which set of features/relations will be useful to obtain relevant results. The following methodologies start directly with the data and will learn by themselves how to combine the initial attributes (X, Y, Z, R, G, B …) into efficient features for the task at hand.

Following PointNet [101] and PointNet++ [102] which are considered as a baseline approach in the community, other work applied deep learning to point set input or voxel representations.

The end-to-end framework SEGCloud [160] combines a 3D-FCNN, trilinear interpolation and CRF to provide class labels for 3D point clouds. Their approach is mainly performance-oriented compared to state-of-the-art methods based on neural networks, random forests and graphical models. They interestingly use a trilinear interpolation which add an extra boost in performance enabling segmentation in the original 3D points space from the voxel representation. Another promising approach is given by Landrieu and Simonovsky for large scale Point Cloud semantic segmentation with Superpoint graphs [161]. In the article, the authors propose a deep learning-based framework for semantic segmentation of point clouds. They initially postulate that the organization of 3D point clouds can be efficiently captured by a structure (Superpoint graph), derived from a partition of the scanned scene into geometrically homogeneous elements. Their goal is to offer a compact representation of contextual relationships between object parts to exploit through convolutional network. The approach is similar to Connected Elements [2,108] in essence, through a graph-based representation. Finally, the works of Engelmann et al. in [82,88] provides very interesting performances by including the spatial context into the PointNet neural network architecture [82] or providing an efficient feature learning and neighbourhood selection strategy [88]. These works are very inspiring and have the potential to become de-facto methodologies for a wide variety of application through transfer learning. As such, they are very good methodology for benchmarking semantic segmentation approaches.

In this state-of-the-art review of pertinent related work, we highlighted three different directions that will drive our methodology. First, it is important that we identify key points in a point cloud which can retain a relevant connotation to domain-related objects. Secondly, we noted that for gravity-based scenes, these elements have a space continuity and often feature homogeneity. Third, specifically, man-made scene retain a high proportion of planar surfaces that can host other elements (floor, ceiling, wall …) [117]. Therefore, detecting these constitute a central first step in our methodological framework but must be quick, scalable, robust, reliable and flexible. It is important to note that the global context may be lost if working with relatively small neighbourhood samples.

4.3 MATERIALS AND METHODS

In this section, we describe a point cloud parsing method to extract semantic clusters (Connected Elements [108]), which can be refined in application-dependent classes.

Our automatic procedure is serialized in 7 steps illustrated in Figure 25 and described in the 4 following sub-sections. In Section 4.3.1, we describe the voxel grid constitution. In Section 4.3.2 we cover feature extraction processes for low-level shape descriptors (Section 4.3.2.1) and relational features (Section 4.3.2.2). Then, we provide in Section 4.3.3 a connected-component system using extracted feature sets SF1 and SF2, followed by a point-level refinement within each voxel to obtain Semantic patches. Finally, we propose a graph-based assembly for the constitution of Connected Elements [2] and a classification routine to obtain labelled point data (Section 4.3.4) benchmarked in Section 4.5.

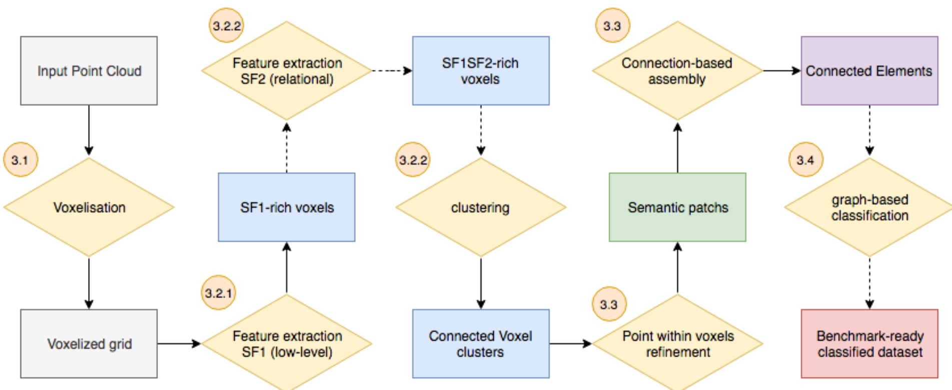


Figure 25. Methodological workflow for the constitution of Connected Elements and knowledge-based classification. A point cloud goes through 7 serialized steps (diamonds) to obtain a fully classified dataset (red square).

4.3.1 Voxelisation grid constitution

Our approach proposes to integrate different generalization levels in both feature space and spatial space. First, we establish an octree-derived voxel grid over the point cloud and we store points at the leaf level. As stated in [69,162], an octree involves recursively subdividing an initial bounding-box into smaller voxels until a depth level is reached. Various termination criteria may be used: the minimal voxel size, predefined maximum depth tree, or a maximum number of sample points within a voxel. In the proposed algorithm, a maximum depth tree is used to avoid computations necessitating domain knowledge early on. To study the influence of the design choice, we study the impact of tree depth selection over performances in Section 4.5, starting at a minimum level of 4. The grid is constructed following the initial spatial frame system of the point cloud to account for complex scenarios where point repartition doesn't follow precisely the axes.

Let p_i be a point in $\mathbb{R}^s$, with s the number of dimensions. We have a point cloud $\mathcal{P} = \{p_i\}_{i=1}^{n}$ with n the number of points in the point cloud. Let $\mathcal{V}_{i,j,k}$ be a voxel of $\mathcal{P}$ identified by a label $\mathcal{L}_i$, containing m points from $\mathcal{P}$.

The cubic volume defined by a voxel entity provides us with the advantage of fast yet uniform space division, and we hence obtain an octree-based voxel structure at a specific depth level. Our approach similarly to [163] is constructed using indexes to avoid overhead. The constituted voxel grid, with the goal of creating Connected Elements discards empty voxels to retain only points-filled voxels. However, for higher end applications such as pathfinding, the voxel-grid can be used as a negative to look for empty spaces. We then construct a directed graph $\mathcal{g}$ defined by a set $\mathcal{v}(\mathcal{g})$ of inner nodes, a set $e(\mathcal{g})$ of edges and a set $\mathcal{v}_e(\mathcal{g})$ of leaf nodes, each representing a non-empty voxel at an octree level, illustrated over a room sample of the S3DIS dataset in Figure 26.

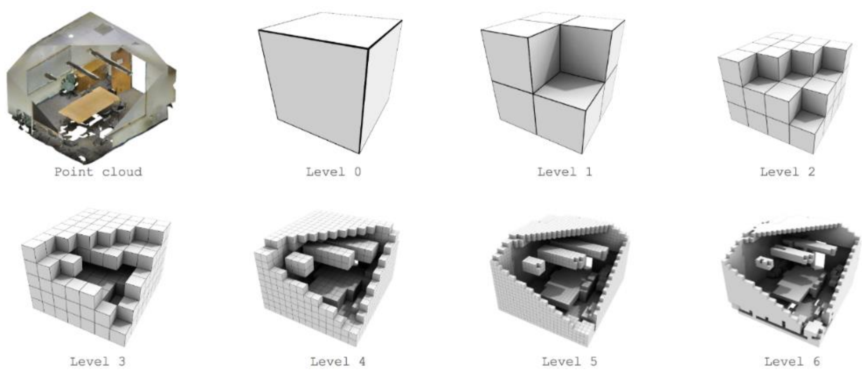


Figure 26. Point Cloud and its extracted voxel structure, where each octree level represents the grid voxels, each subdivided in subsequent 8 voxel children.

Once each point has been assigned to a voxel regarding the defined grid within the $\mathbb{R}^3$ Euclidean space along $\overrightarrow{e_x}, \overrightarrow{e_y}, \overrightarrow{e_z}$, we consider leaf nodes $v_e(\mathcal{G})$ of $\mathcal{G}$ as our representative primitive.

4.3.2 Feature Extraction

As a single object of the resulting feature vector is hardly interpretable [76], we aim at extracting a robust feature set for general semantic segmentation frameworks. To insure interoperable workflows, we used descriptors that were thoroughly studied and made their proof in various works referred in Section 4.2.

Our new voxel-primitive serves as an initial feature host, and acts as a point neighbourhood selection approach. These can then be transferred following the structure of $\mathcal{G}$, permitting feature transfer at every octree depth level extended to the point-storage (Figure 27).

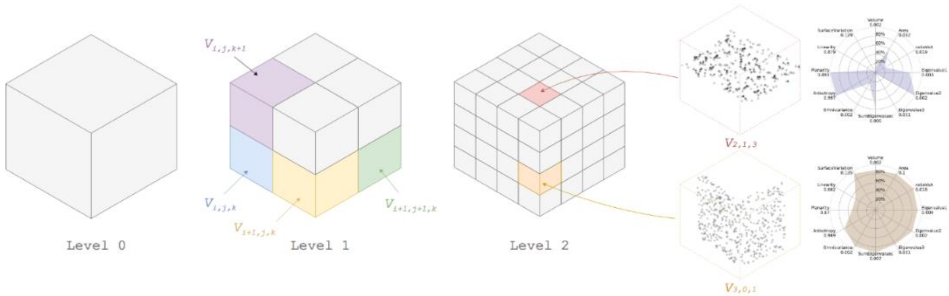


Figure 27. Feature transfer between octree levels. We note that each non-empty node describes a voxel which can then permit a point-level access for example to compute feature sets (Here, a planar voxel and a corresponding SF1 sample, and a transition voxel and its corresponding SF1 sample)

This permits a flexible and unconstrained feature-based point cloud parsing which can process raw data (i.e. pure X, Y, Z Euclidean sets). In the next sub-section 4.3.2.1, we present several low-level shape-based features used to construct our SF1 feature set. Then, we explain our relationship-level feature set (SF2) which permits to leverage local topology and relationships at different cluster levels.

4.3.2.1 Low-level shape-based features (SF1)

The first group of low-level features is mainly derived from Σ, our data covariance matrix of points within each voxel for the low memory footprint and fast calculation, which in our case we define as:

$$\Sigma = \frac{1}{m-1}\sum_{i=1}^{m}(X_i - \bar{X})(X_i - \bar{X})^T \tag{1}$$

where $\bar{X}$ is the mean vector $\bar{X} = \sum_{i=1}^{m} p_i$.

From this high yielding matrix, we derive eigen values and eigen vectors through Singular Value Decomposition [164] to increase computing efficiency, which firstly correspond to modelling our voxel containment by a plane, showing to largely improve performances. We follow a Principal Composant Analysis (PCA) to describe three principal axes describing the point sample dispersion. Thus, we heavily rely on eigen vectors and eigen values as a feature descriptor at this point. Therefore, their determination needs to be robust. This is why we use a variant of the Robust PCA approach presented in the article [5] to avoid miscalculation. We sort eigenvalues $\lambda_1, \lambda_2, \lambda_3$ such as $\lambda_1 > \lambda_2 > \lambda_3$, where linked eigen vector $\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}$ respectively represent the principal direction, its orthogonal direction and the estimated plane normal. These indicators as reviewed in Section 4.2 are interesting for deriving several eigen-based features [138] as following:

$$\lambda_a = (\lambda_1 - \lambda_3)/\lambda_1 \tag{2}$$

$$\lambda_l = (\lambda_1 - \lambda_2)/\lambda_1 \tag{3}$$

$$\lambda_p = (\lambda_2 - \lambda_3)/\lambda_1 \tag{4}$$

$$\lambda_v = \lambda_3 / \sum_{i=1}^{3} \lambda_i \tag{5}$$

$$\lambda_o = \sqrt[3]{\prod_{i=1}^{3} \lambda_i} \tag{6}$$

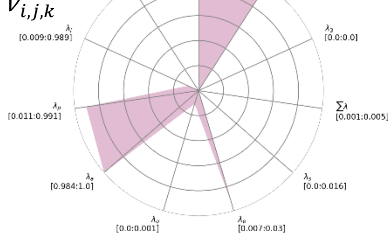
$$\lambda_s = \lambda_3/\lambda_1 \tag{7}$$

$$\lambda_e = -\sum_{i=1}^{3} \lambda_i * \ln(\lambda_i) \tag{8}$$

Where for the voxel $\mathcal{V}_{i,j,k}$, λ_a is its anisotropy, λ_l its linearity, λ_p its planarity, λ_v its surface variation, λ_o its omnivariance, λ_s its sphericity and λ_e its eigen entropy. The first set of eigen-based features is summarized in Table 5.

Table 5. Eigen-based features part of the SF1 feature set

Eigen-based feature	**Description**
$\lambda_1, \lambda_2, \lambda_3$	Eigen values of $\mathcal{V}_{i,j,k}$ where $\lambda_1 > \lambda_2 > \lambda_3$
$\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}$	Respective Eigen vectors of $\mathcal{V}_{i,j,k}$
$\overrightarrow{v_3}$	Normal vector of $\mathcal{V}_{i,j,k}$
λ_a	Anisotropy of voxel $\mathcal{V}_{i,j,k}$
λ_e	Eigen entropy of voxel $\mathcal{V}_{i,j,k}$
λ_l	Linearity of voxel $\mathcal{V}_{i,j,k}$
λ_o	Omnivariance of voxel $\mathcal{V}_{i,j,k}$
λ_p	Planarity of voxel $\mathcal{V}_{i,j,k}$

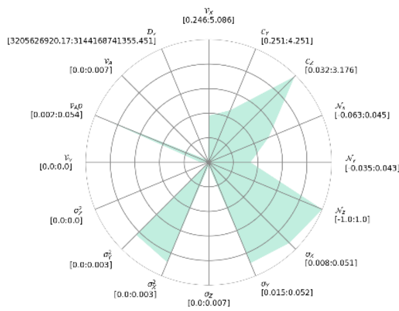


λ_s	Sphericity of voxel $\mathcal{V}_{i,j,k}$
λ_v	Surface variation of voxel $\mathcal{V}_{i,j,k}$

We extract a second geometry-related set of features (Table 6), starting with $\overline{\mathcal{V}_{i_x}}, \overline{\mathcal{V}_{i_y}}, \overline{\mathcal{V}_{i_z}}$ the mean value of points within a voxel $\mathcal{V}_{i,j,k}$.

Table 6. Geometrical features part of the SF1 feature set

Geometrical feature	Description
$\overline{\mathcal{V}_{i_x}}, \overline{\mathcal{V}_{i_y}}, \overline{\mathcal{V}_{i_z}}$	Mean value of points in $\mathcal{V}_{i,j,k}$ respectively along $\overrightarrow{e_x}, \overrightarrow{e_y}, \overrightarrow{e_z}$
${\sigma_{i_x}}^2, {\sigma_{i_y}}^2, {\sigma_{i_z}}^2$	Variance of points in voxel $\mathcal{V}_{i,j,k}$
$\mathcal{V}_{\mathcal{A}p}$	Area of points in $\mathcal{V}_{i,j,k}$ along $\overrightarrow{n_\mathcal{V}}$ ($\overrightarrow{v_3}$)
$\mathcal{V}_{\mathcal{A}}$	Area of points in $\mathcal{V}_{i,j,k}$along $\overrightarrow{e_z}$
m	Number of points in $\mathcal{V}_{i,j,k}$
$V_\mathcal{V}$	Volume occupied by points in $\mathcal{V}_{i,j,k}$
$D_\mathcal{V}$	point density within voxel $\mathcal{V}_{i,j,k}$



The area features $\mathcal{V}_{\mathcal{A}p}, \mathcal{V}_{\mathcal{A}}$ are obtained through a convex hull (Eq. 10) analysis respectively along $\overrightarrow{v_3}$ and $\overrightarrow{e_z}$. The third is the local point density within the segment, which is defined as follows:

$$D_\mathcal{V} = \frac{m}{V_\mathcal{V}} \tag{9}$$

where $V_\mathcal{V}$ is the minimum volume calculated through a 3D convex hull, such as:

$$Conv(\mathcal{P}) = \left\{\sum_{i=1}^{|\mathcal{P}|} \alpha_i q_i \,\middle|\, (\forall i: \alpha_i \geq 0) \wedge \sum_{i=1}^{|\mathcal{P}|} \alpha_I = 1\right\} \tag{10}$$

$$V_\mathcal{V} = \frac{1}{3}\left|\sum_{j=1}^{m} (\overrightarrow{Q_F}.\overrightarrow{n_F}) area(F)\right| \tag{11}$$

In order to prevent outweighing some attributes and to equalizes the magnitude and variability of all features we standardize their values from different dynamic ranges into a specified range. There are three common normalization methods as referred in [86]: Min-max, Z-score, and decimal scaling normalization. In this research we use Min-max method found empirically more computationally efficient to normalize the multiple features F in F_N, normalized feature in a [0: 1] range:

$$F_N = \frac{F - \min(F)}{\max(F) - \min(F)} \tag{12}$$

We combine eigen-based features and geometrical features for an easier data visualization in two separate spider charts (E.g. in Table 5 and Table 6). Then we plot normalized distributions per-voxel category (E.g. in Figure 28) to better understand the variations within features per element category.

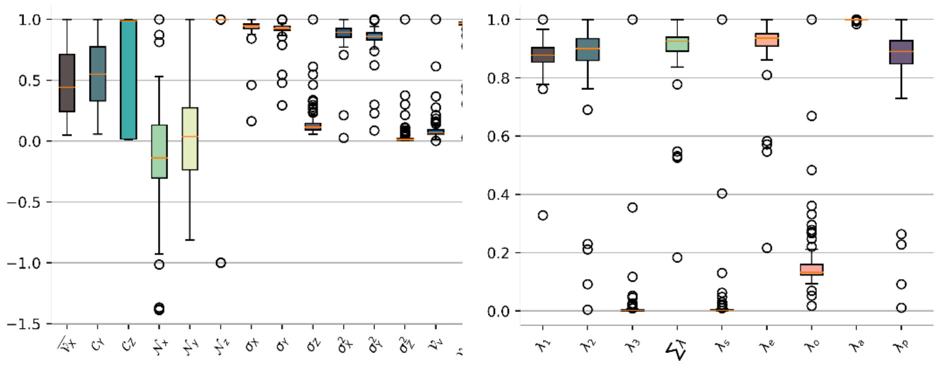


Figure 28. Box plot of primary elements feature variation.

We note that for the example of Primary Elements (mostly planar, described in Section 4.3.3), there is a strong similarity within the global voxel feature sets, except for orientation-related features (Normals, Position, Centroids).

4.3.2.2 Connectivity and Relationship features (SF2)

There are very few works that deals with explicit relationship feature extraction within point clouds. This is mostly justified by the complexity and exponential computation to extract relevant information at the point-level. Thus, the second set of proposed feature set (SF2) is determined at several octree levels. First, for each leaf voxel, we extract a 26-connectivity graph which appoints every neighbour for every voxel. These connectivity are primarily classified regarding their touch-topology [125] which either is vertex.touch, edge.touch or face.touch (Figure 29).

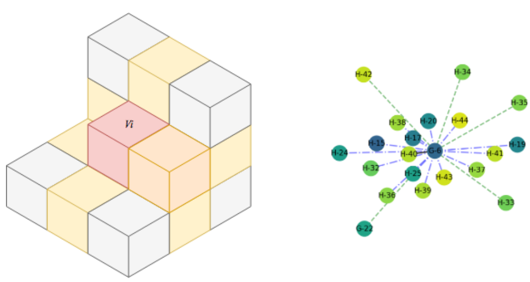


Figure 29. Direct voxel-to-voxel topology in a 26-connectivity graph. Considered voxel $\mathcal{V}_i$ is red, direct connections are either vertex.touch (grey), edge.touch (yellow) or face.touch (orange)

To complement this characterization of voxel-to-voxel topology, each processed voxel is complemented through new relational features. Immediate neighbouring voxels are initially studied to extract F_g (geometrical difference) using the log Euclidean Riemannian metric, a measure of the similarity between adjacent voxels covariance matrices:

$$F_g = \left\| \log \Sigma_{v_i} - \log \Sigma_{v_j} \right\|_F \quad (13)$$

where log(.) is the matrix logarithm operator and $\| . \|_F$ is the Frobenius norm.

If the SF1 feature set is available (non-constrained through computational savings), and depending on the desired characterization, these are favoured for an initial voxel tagging.

To get higher end characterization while limiting the thread surcharge to a local vicinity, we estimate concavity and convexity between adjacent voxels. It refines the description of the graph edge between the processed node(voxel) and each of its neighbour (Algorithm 2. We define $\alpha_{\mathcal{V}}$ the angle between two voxels $\mathcal{V}_i$ and $\mathcal{V}_j$ as:

$$\alpha_{\mathcal{V}} = \overrightarrow{n_{\mathcal{V}_i}}.(\overrightarrow{\Sigma_{\mathcal{V}_i}} - \overrightarrow{\Sigma_{\mathcal{V}_j}}) \quad (14)$$

Algorithm 2 Voxel Relation Convexity/Concavity tagging

Require: A voxel $\mathcal{V}_i$ and its direct vicinity $\{\mathcal{V}_j\}_{j=1}^{26}$ expressed as a graph g.

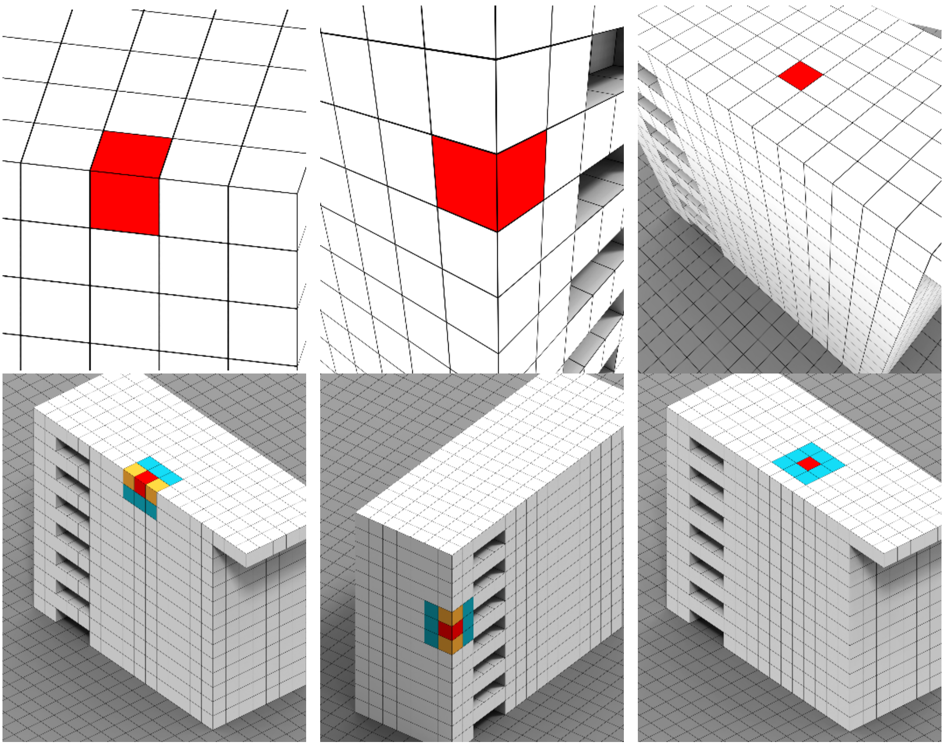
1. **For each** $\mathcal{V}_j \neq \emptyset$ **do**
2. $\alpha_{\mathcal{V}} \leftarrow$ angle between normal of voxels
3. **if** $\alpha_{\mathcal{V}} < 0$ **then**
4. $e_{ij}(g) \leftarrow$ edge between $\mathcal{V}_i$ and $\mathcal{V}_j$ is tagged as Concave
5. **else** $e_{ij}(g) \leftarrow$ edge between $\mathcal{V}_i$ and $\mathcal{V}_j$ is tagged as Convex
6. **end if**
7. **end for**
8. **end**
9. **return** (g)

Third we extract 4 different planarity-based relationships (Figure 30) between voxels which we define as:

- Pure Horizontal relationship: For $\mathcal{V}_i$, if an adjacent voxel $\mathcal{V}_j$ has a $\overrightarrow{v_3}$ colinear to the main direction (vertical in gravity-based scenes), then the edge $e(v_i, v_j)$ is tagged $\mathcal{Hr}$. If two adjacent nodes v_i and v_j hold an $\mathcal{Hr}$ relationship and both $\overrightarrow{v_3}$

aren't colinear, they are connected by a directed edge, $e_d(v_i, v_j)$ where v_i is the starting node. In practice, voxels near horizontal surfaces hold this relationship.

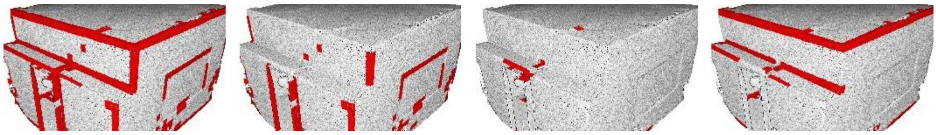
- Pure Vertical relationship: For $\mathcal{V}_i$, if an adjacent voxel $\mathcal{V}_j$ has a $\overrightarrow{v_3}$ orthogonal to the main direction (vertical in gravity-based scenes), then the edge $e(v_i, v_j)$ is tagged $\mathcal{V}e$. If two adjacent nodes v_i and v_j are connected through $\mathcal{V}r$ and both $\overrightarrow{v_3}$ are coplanar but not colinear, they are connected by a directed edge, $e_d(v_i, v_j)$. In the case we are in a gravity-based scenario, they are further refined following $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$ axis. These typically includes voxels near vertical surfaces.
- Mixed relationship: For $\mathcal{V}_i$, if within its 26-connectivity neighbours, the node v_i presents $\mathcal{V}e$ and $\mathcal{H}r$ edges, then v_i is tagged as $\mathcal{M}r$. In practice, voxels near both horizontal and vertical surfaces hold this relationship.
- Neighbouring relationship. If two voxels do not hold one of these former constraining relationships but are neighbours, associated nodes are connected by an undirected edge without tags.



(**a**) (**b**) (**c**)

Figure 30. Relationship tagging in the voxel-space. (a) represent a mixed relationship $\mathcal{M}r$, (b) a pure vertical relationship $\mathcal{V}r$, (c) a pure horizontal relationship $\mathcal{H}r$.

Illustrated on the S3DIS dataset, for a room, this is an example of what the different voxel categories are:



(**a**) (**b**) (**c**) (**d**)

Figure 31. S3DIS points within categorized voxels. (a) Full transition voxels, (b) vertical group of points, (c) horizontal group of points, (d) mixed group of points,

Finally, the number of relationships per voxel is accounted as edge weights pondered by the type of voxel-to-voxel topology, where vertex.touch=1, edge.touch=2 and face.touch=3. We obtain a feature set SF2 as in Table 7:

Table 7. Relational features of the SF2 feature set for 3D structural connectivity

Relational feature	Description
$g_{26}(i)$	Graph of voxel entity i and its neighbours retaining voxel topology (vertex.touch, edge.touch, face.touch)
F_g	Geometrical difference
$g_{26-cc}(i)$	$g_{26}(i)$ retaining Convex/Concave tags.
$g_{26-cc-p}(i)$	$g_{26-cc}(i)$ retaining planarity tags ($\mathcal{H}r, \mathcal{V}r, \mathcal{M}r$).

This is translated into a multi-set graph representation to give a flexible featuring possibility to the initial point cloud. As such, extended vicinity is then possible seed/host of new relationships that permit a topology view of the organization of voxels within the point cloud (E.g. Figure 32).

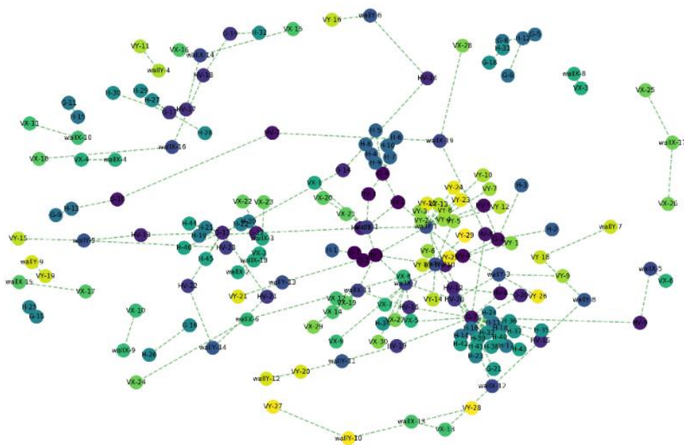


Figure 32. Graph representation within a voxel sample of the point cloud.

These relationships are represented in different groups to extract different feature completing the relationship feature set. Graphs are generated automatically through full voxel samples regarding Category tags and Convex-Concave tags.

4.3.3 Connected Element constitution and voxel refinement

Based on the feature sets SF1 and SF2, we propose a connected-component workflow driven by planar patches. Connected-component labelling is one of the most important processes for image analysis, image understanding, pattern recognition, and computer vision and is reviewed in [165]. Mostly applied for 2D data, we extend it to our 3D octree structure for efficient processing and parallelization compatibility. We study the predominance of planar surfaces in man-made environments and the feature-related descriptor which provides segmentation benefits. The designed feature representations described in Section 4.3.2 are used as a mean to segment the gridded point cloud into groups of voxels that share a conceptual similarity. These groups are categorized within four different entities: Primary Elements (PE), Secondary elements (SE), transition elements (TE) and remaining elements (RE) as illustrated in Figure 33.

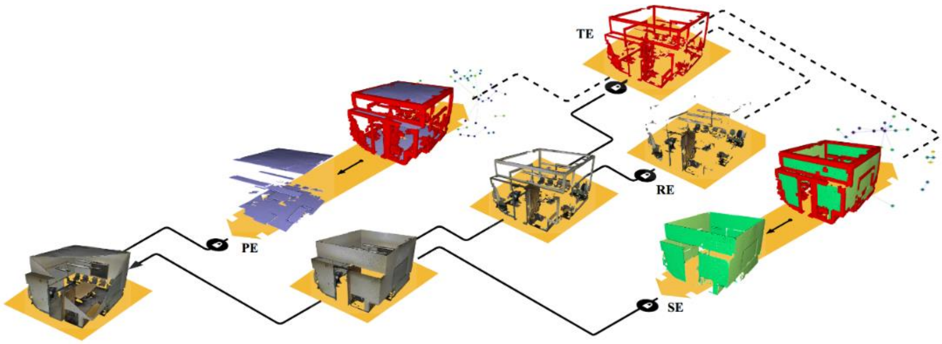


Figure 33. Elements detection and categorization. A point cloud is search for PE, the rest is searched for SE. The remaining from this step is searched for TE, leaving RE. TE permits to extracts graphs through SF2 analysis with PE, SE and RE

We start by detecting the PE using both feature sets. Initially, we group voxels that answer a collinearity condition with the main direction. Due to the normal dispersion in voxel sets (which has no exact collinear match), this condition is translated by comparing the angle of normalized vectors against a threshold:

$$\alpha_v < th_a \quad with \quad \alpha_v = \cos^{-1}\left(\frac{\overrightarrow{v_3(\imath)}.\overrightarrow{v_3(\jmath)}}{\left\|\overrightarrow{v_3(\imath)}\right\|.\left\|\overrightarrow{v_3(\jmath)}\right\|}\right) \tag{15}$$

Using SF2, we then cluster linked nodes through connected-component labelling. PE presents mainly clusters of points which are the main elements of furniture (table top, chair seat …) or ceiling and ground entities.

SE are constituted of voxels which hold $\overrightarrow{v_3}$ orthogonal to the main direction, further decomposed along $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$. As such, they are usually constituted of elements which belong to walls, and horizontal planar-parts of doors, beams …

The "edges" voxels which are within the set of tagged voxels $\{\mathcal{Hr}, \mathcal{Vr}, \mathcal{Mr}\}$ are seeds to constitute TE which are then further decomposed (voxel refinement) in semantic patches with homogeneous labelling depending on their inner point characterization. As such, they play an important role for understanding the relationships between primary, secondary and remaining elements. They are initially grouped based on F_g and clustered in connected-components using $g_{26-cc-p}(i)$ (SF2). The voxels containing "edges" (E.g. in Figure 34) or multiple possible points that should belong to separate objects are further subdivided by studying the topology and features with their neighbouring elements.

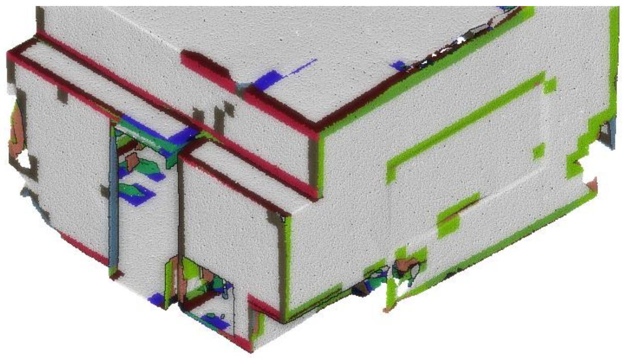
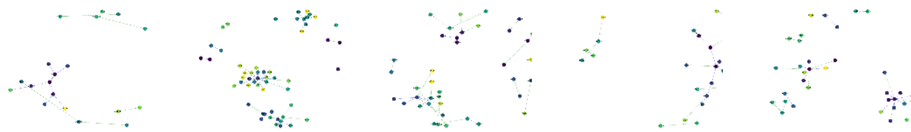


Figure 34. Edges elements to be decomposed in TE and RE

Finally, the remaining voxels are labelled through connected-components as RE, and their SF1-similarity is aggregated as a feature descriptor. For each element within the Connected Elements (CEL) set {PE, SE, TE, RE}, voxel features are aggregated to obtain a global SF1 and SF2 feature set per CEL, updated through voxel refinement. Implementation-wise, CEL are sorted by occupied volume after being sorted per category, Relationships exist between primary, secondary, edges and remaining elements due to their voxel-based direct topology. This proximity is used to refine voxels (thus elements), by extracting points within voxel neighbours of an element ε_i which belong to an element ε_j based on defined SF1 features of ε_i-voxel. In this context, this permits to leverage planar-based dominance in man-made scenes using for example eigen-based features. We thus extract a new connectivity graph between CEL where the weight of relationships is determined using the number of connected voxels. This allows to refine the transition voxels based on their local topology and connectivity to surrounding elements. The global element's features therefore play the role of reference descriptors per segment, and points within targeted voxels for refinement are compared against these. If within a voxel points justify belonging to another Connected Element, then the voxel is split in semantic patches, which each retains a homogeneous CEL label. The final structure retains unique CEL labels per leaf, where leaves called semantic patches are either pure voxel or voxel's leaf.

We obtain a graph-set composed of a general CEL graph, a PE graph, a SE graph, a TE graph, a RE graph, and any combination of PE,SE,TE and RE (e.g. Figure 35) :



(**a**) (**b**) (**c**) (**d**) (**e**)

Figure 35. different graphs generated on voxel categories. (a) CEL graph, (b) PE graph, (c) SE graph, (d) TE graph, (e) RE graph

In order to estimate the impact of designed features, we establish a graph-based semantic segmentation over these CEL described in Section 4.3.4.

4.3.4 Graph-based semantic segmentation

For every CEL in the graph-set, we first employ a multi-graph-cut (set of edges whose removal makes the different graphs disconnected) approach depending on the weight of edges defining the strength of relations, where the associated cut cost is:

$$cut(\varepsilon_i, \varepsilon_j) = \sum_{p\epsilon\varepsilon_i, q\epsilon\varepsilon_j} w_{pq} \tag{16}$$

Where w_{pq} is the weight of the edge between nodes p and q. In order to avoid min-cut bias, we use normalized cut by normalizing for the size of each segment:

$$Ncut(\varepsilon_i, \varepsilon_j) = \frac{cut(\varepsilon_i, \varepsilon_j)}{\sum_{k\epsilon e_{\varepsilon_i}(g)} w_k} + \frac{cut(\varepsilon_i, \varepsilon_j)}{\sum_{k\epsilon e_{\varepsilon_j}(g)} w_k} \tag{17}$$

Where $e_{\varepsilon_i}(g)$ are the edges that touches ε_i, and $e_{\varepsilon_i}(g)$ are the edges that touches ε_j.

Our approach was thought as a mean to provide only a first estimate of the representativity of CELs in semantic segmentation workflows, especially to differentiate big planar portions. As such, the provided classifier is very naïve, and will be subject of many improvements in the near future for a better flexibility and to reduce empirical knowledge. It was constructed for indoor applications. For example, a segment with the largest membership to the ceiling might belong to beam or wall, and a segment with the largest membership to floor might belong to wall or door. To handle such semantic mismatches, the graph g_{CEL}, that was previously constructed is used to refine the sample selection using the following rules and the search sequence starting from the class floor, and is followed by the class ceiling, wall, beam, table, bookcase, chair and door. Once a node is labelled with one of these classes, it is excluded from the list of nodes being considered in the sample selection. The definition of thresholds was directly extracted from knowledge about the dimension of furniture objects from the European Standard EN1729-1:2015. As for the concepts at hand, these were defined regarding the Semantic Web resources, mainly the ifcOWL formalized ontology representing the Industry Foundation Classes application knowledge [166]. It is important to note that furniture (chair, table, bookcases) models were extracted from these rules and then we simulated scan positions to obtain simulated data. Indeed, sensors artefacts produce noisy point clouds which can then slightly change the definition of thresholds. The obtained samples were then looked against 5 objects of the S3DIS to insure consistency with the device knowledge.

1) A node is tagged "floor" when for a primary element $p\varepsilon_i$ and all primary elements $p\varepsilon$:

$$\begin{aligned} \mathcal{V}_{\mathcal{A}}(p\varepsilon_i) \in maximas(\mathcal{V}_{\mathcal{A}}(p\varepsilon)) \;\&\; \sum_k e_{p\varepsilon_i}(g_{p\varepsilon}) \\ \in maximas() \;\&\; Z_{p\varepsilon_i} \in maximas(Z_{p\varepsilon}) \end{aligned} \tag{18}$$

   with $\sum_k e_{p\varepsilon_i}(g_{p\varepsilon})$ being the sum of edge weights of all outgoing edges and incoming edges.

2) The "ceiling" is similar to the "floor" labelling with the difference that $Z_{p\varepsilon_i}$ is search among minimas of $p\varepsilon$.

3) Once all the ceiling and floor segments are identified, the nodes in the graph $g_{s\varepsilon}$ of secondary elements are searched for "wall" segments by first identifying all the nodes that are connected to the ceiling or floor nodes through the edges of the designated relationships. To handle complex cases, the area feature guides the detection through thresholding to exclude non-maxima.

4) To identify "beams", A sub-graph $g_{r-p\varepsilon-s\varepsilon}$ composed of remaining non-classified elements from PE and SE. A connected-component labelling is performed guided by transition elements. It is then searched for nodes that are connected to the ceiling and the walls, which are then classified as "beam" segments.

5) The "table" segments are extracted by the remaining elements of primary elements, if its SF1 feature set presents a correspondence of more than 50% with a sample table object. We note that the predominant factor is the height which is found within 70 and 110 cm from the ground segment. The feature

correspondence is a simple non-weighted difference measure between the averaged SF1 features between the sample and the compared element. The sample element is constructed by following the domain concepts and thresholds as explained previously.

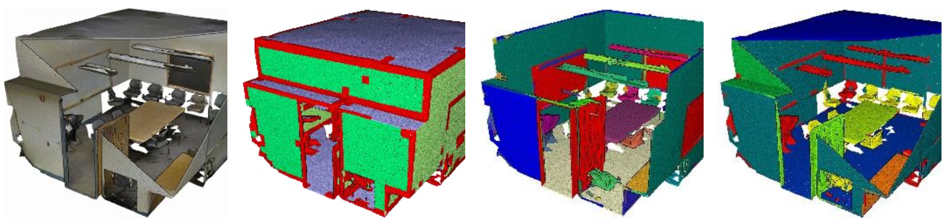
6) From the remaining elements RE, we identify “bookcases” if it presents a direct SF2 connectivity to wall segments and a SF1 feature correspondence of more than 50%.

7) Then, RE and remaining PE are aggregated through connected components and tagged as “chair” if their mean height above ground is under 100 cm.

8) All the unclassified remaining nodes are aggregated in a temporary graph, and a connected-component labelling is executed. An element is tagged as “door” if the bounding-box element’s generalization intersect a wall segment.

9) Every remaining element is classified as “clutter”.

By using the above 9 rules, the ceiling, floor, wall, beam, table, chair, bookcase, door and clutter classes are looked for, going from raw point cloud to a classified dataset as illustrated in Figure 36.



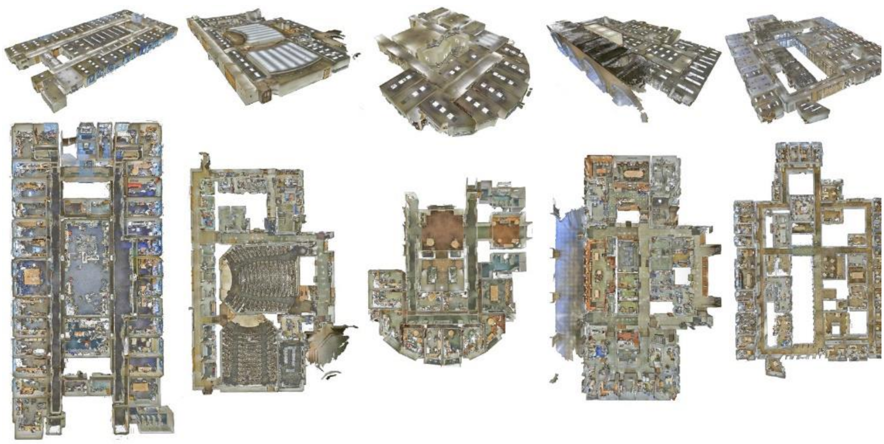
(**a**) (**b**) (**c**) (**d**)

Figure 36. a) raw point cloud; (b) {PE, SE, TE, RE} groups of voxels; (c) Connected Elements; (d) Classified point cloud.

4.4 DATASET

To test our approach, we evaluate feature performance in one application context: 3D semantic segmentation for indoor environment. We used the S3DIS dataset [133] from the Matterport sensor [70]. It is composed of 6 areas each subdivided in a specific number of rooms (Table 8) for a total of 270 sub-spaces [108]. These areas show diverse properties and include 156 offices, 11 conference rooms, 2 auditoriums, 3 lobbies, 3 lounges, 61 hallways, 2 copy rooms, 3 pantries, 1 open space, 19 storage rooms and 9 restrooms. One of the areas includes multiple floors, whereas the rest have one, and is very representative of building indoor spaces. The dataset is very noisy, presents imprecise geometries, clutter and heavy occlusion. During the tests we noted that some points were mislabelled in the ground-truth labels, and that several duplicate points (points where the distance is inferior to 10^{-9} m from one another) add an extra bias. However, it was chosen as it is a big dataset which provides a high variability of scene organization and is currently used for benchmarking new algorithms. It is a very interesting opportunity to evaluate the robustness of our approach and to study the impact of features and their robustness to hefty point cloud artefacts. We remind the readers that the goal is to obtain relevant semantic patches constituting Connected Elements in a Smart Point Cloud Infrastructure.

Table 8. The S3DIS dataset and its 6 areas used for testing our methodology

	Area-1	**Area-2**	**Area-3**	**Area-4**	**Area-5**	**Area-6**
						
#Points	43 956 907	470 023 210	18 662 173	43 278 148	78 649 818	41 308 364
Area (m^2)	965	1100	450	870	1700	935
Rooms (nb)	44	40	23	47	68	48

We consider 9 out of 13 classes in the S3DIS dataset, holding 88.5% of the total number of segments representing both moveable and structural elements. The choice was motivated by the colour-dependence of the remaining classes. Indeed, in this article we focus on a general approach with minimal input, and as such we filtered the initial dataset before computing metrics for every point initially assigned to one of the following classes: column, window, sofa, board. Thus, our approach runs on the full dataset, but we compare only these classes. The repartition is found in Table 9:

Table 9. S3DIS per-area statistics regarding the studied classes.

Method	Ceiling 0	Floor 1	Wall 2	Beam 3	Door 6	Table 7	Chair 8	Bookcase 10	Others 12
Area 1	56	45	235	62	87	70	156	91	123
Area 2	82	51	284	62	94	47	546	49	92
Area 3	38	24	160	14	38	31	68	42	45
Area 4	74	51	281	4	108	80	160	99	106
Area 5	77	69	344	4	128	155	259	218	183
Area 6	64	50	248	69	94	78	180	91	127
Full S3DIS	391	290	1552	215	549	461	1369	590	676

4.5 RESULTS

4.5.1 Metrics

Existing literature has suggested several quantitative metrics for assessing the semantic segmentation and classification outcomes. We define the metrics regarding the following terms extracted from a confusion matrix C of size $n \times n$ (with n the number of labels, and each term denoted c_{ij}):

- True Positive (TP): Observation is positive and is predicted to be positive.
- False Negative (FN): Observation is positive but is predicted negative.
- True Negative (TN): Observation is negative and is predicted to be negative.
- False Positive (FP): Observation is negative but is predicted positive.

Then the following metrics are used:

$$IoU_i = \frac{TP_i}{FP_i+FN_i+TP_i} \text{ equivalent to } IoU_i = \frac{c_{ii}}{c_{ii}+\sum_{j\neq i} c_{ij}+\sum_{k\neq i} c_{ki}} \quad (19)$$

$$\overline{IoU} = \frac{TP}{FP+FN+TP} \text{ equivalent to } \overline{IoU} = \frac{\sum_{i=1}^{n} IoU_i}{n} \quad (20)$$

$$oAcc = \frac{\sum_{i=1}^{n} \frac{TP_i}{FP_i+TP_i}}{n} \text{ equivalent to } oAcc = \frac{\sum_{i=1}^{n} c_{ii}}{\sum_{j=1}^{n} \quad \sum_{k=1}^{n} c_{jk}} \quad (21)$$

$$precision = \frac{TP}{TP+FP}, recall = \frac{TP}{TP+FN}, F_{1-score} = \frac{2TP}{2TP+FP+FN} \quad (22)$$

The Overall Accuracy ($oAcc$) is a general measure on all observation about the performance of the classifier to correctly predict labels. The precision is the ability of the classifier not to label as positive a sample that is negative, the recall is intuitively the ability of the classifier to find all the positive samples, The F_1-score can be interpreted as a weighted harmonic mean of the precision and recall, thus gives a good measure of how well the classifier performs. Indeed, global accuracy metrics are not appropriate evaluation measures when class frequencies are unbalanced, which is the case in most scenarios both in real indoor and outdoor scenes, since they are biased by the dominant classes.

In general, the Intersection-Over-Union (IoU) metric tends to penalize single instances of bad classification more than the F1-score quantitatively even when they can both agree that this one instance is bad. Thus, IoU metric tends to have a "squaring" effect on the errors relative to the F1-score. Henceforth, the F1-score in our experiments gives an indication on the average performance of our proposed classifier, while the IoU score measures the worst-case performance.

4.5.2 Quantitative and qualitative assessments

4.5.2.1 Feature influence

Our first experiment uses SF1 independently and combined with SF2 to highlight performances and influence consequences on a representative sample from the S3DIS dataset. We list in Table 10 the main results regarding timings, number of CEL, elements (PE, SE, TE and RE) extracted as well as global metrics.

Table 10. Analyses of the impact of feature sets over samples of the S3DIS dataset

Method	Zone	Time (min)	CEL number	mIOU	oAcc	F1-score
SF1	Room	0.7	214	0.53	0.73	0.77
	Area 1	42.4	10105	0.35	0.58	0.63
SF1SF2	Room	1.0	125	0.83	0.95	0.95
	Area 1	55.0	5489	**0.47**	**0.75**	**0.75**

We note that SF1SF2 takes 30% longer but permits to obtain 12 IoU points overall for Area-1, as well as 17 overall accuracy points and 12 F1-score points. For some rooms where the connectivity predicates are predominant, we can obtain more than 30 IoU points increase. It also is very important to limit over-segmentation problematics while being versatile enough depending on different application needs. Thus, if we look at both the room and area 1 S3DIS samples, we note that the global number of CEL drops significantly, which permits classifier to reach a more representative detection (E.g. Table 11 gives an SF1SF2 instance detection comparison to ground truth)

Table 11. Quantitative CEL segmentation compared to nominal number of elements per class for both a room (Conference room) and area (area 1)

CEL number	Ceiling 0	Floor 1	Wall 2	Beam 3	Door 6	Table 7	Chair 8	Bookcase 10
Room 1	1	1	4	1	1	1	13	1
Tagged CEL	1	1	4	1	1	1	11	1
Area 1	56	44	235	62	87	70	156	91
Tagged CEL	52	44	146	47	23	67	129	70

We then applied our specific knowledge-based classification approach over both SF1 alone and SF1SF2. The metrics per class over the Area 1 are shown in Table 12.

Table 12. Global per-class metrics concerning the Area-1 of the S3DIS dataset. SF1 alone and combined SF1SF2 are compared.

Global metrics Area-1	Ceiling 0	Floor 1	Wall 2	Beam 3	Door 6	Table 7	Chair 8	Bookcase 10	Clutter 12
SF1 IoU	0.81	0.75	0.61	0.39	0.10	0.24	0.06	0.02	0.14
SF1 Prec.	0.99	0.99	0.84	0.67	0.11	0.96	0.09	0.15	0.32
SF1 Recall	0.82	0.75	0.69	0.48	0.57	0.25	0.14	0.03	0.20
SF1 F_1 score	0.90	0.86	0.76	0.56	0.18	0.39	0.11	0.05	0.24
SF1SF2 IoU	0.95	0.92	0.67	0.49	0.14	0.32	0.32	0.15	0.31
SF1SF2 Prec.	0.98	0.95	0.79	0.88	0.29	0.9	0.69	0.2	0.41
SF1SF2 Rec.	0.97	0.97	0.82	0.53	0.2	0.33	0.37	0.37	0.56
SF1SF2 F_1	0.97	0.96	0.8	0.66	0.24	0.48	0.48	0.26	0.47

If we look at IoU scores, combining SF1 and SF2 permits to obtain between +6 and +26 points (+13 points in average) compared to SF1 alone, which is a notable increase of performances. The highest growth is achieved for the 'chair' class, and the lowest for the 'door' class. The 'chair' detection rate increase is mostly explained by the 3D connectivity information given by SF2 through {PE, RE} isolation and clustering, which permits to overcome SF1 matching limitations due to large varying signatures within voxels. Concerning doors, the low increase is explained by its low SF2 connectivity information as within the S3DIS dataset, door elements don't show any clear 'cuts' with wall elements, and therefore aren't identified clearly within RE. This can be solved by accounting for colour information to better segment the point cloud, or by using the spatial context and empty voxels within wall segments. Also, the high recall score for bookcase shows that the combination permits to better account for the right number of bookcase elements. Overall, while we notice a slight precision score decrease for planar-based classes (ceiling, floor, wall), recall rates largely increases between SF1 and SF1SF2. This highlight the ability of our classifier to better identify all the positive samples. This is translated in F_1-scores which are superior for all classes up to +37 points.

Then we studied the impact of influential factors over the results and performance of the algorithm (experiments were run 50 times each to validate the time output) as show in Figure 37.

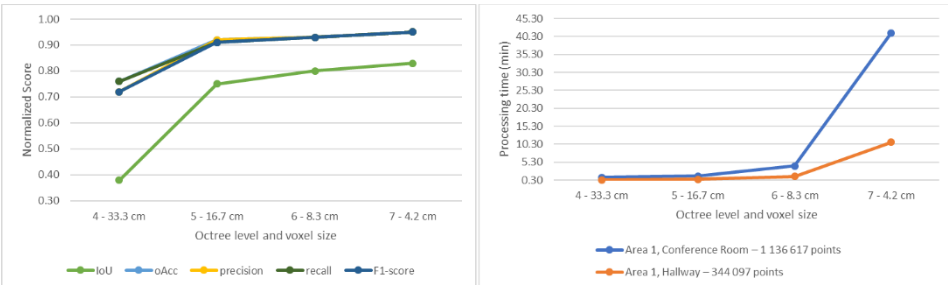


Figure 37. Normalized score and processing time in function of the defined octree level

We observe that our different metrics rise in a similar manner, with a great score increase from octree level 4 to 5 (38 IoU points), and then not a distinctive increase. On the other end, we see an increase in processing time from octree level 5 to 6, and a great increase from octree 6 to 7. This orient our choice toward a base process at octree level 5, sacrificing some score points for an adequate performance.

4.5.2.2 Full S3DIS benchmark

We see that combining both SF1 and SF2 outperform a sole independent use of SF1 feature sets. Therefore, SF1SF2 method is compared against state-of-the-art methodologies. Due to the rise of deep learning approaches, we related our knowledge-based procedure to the best-performing supervised architectures.

We first tested our semantic segmentation approach on the most complex area, Area 5, which holds a wide variety of rooms with varying size, architectural elements and problematic cases. This is our worst-case scenario area. It holds different complex cases that the knowledge-based classification approach struggles to handle, and results can be found in Appendix B. Concerning performances and calculation times for Area-5 (68 rooms), our approaches finishes in 59 minutes (3538.5 seconds) in average (10 test-run) whereas the well-performing SPG approach [161] allows the classification of the Area (78 649 682 points) in 128.15 minutes (7689 seconds). Thus, while results have a large improving margin for non-planar elements, the approach (without parallelization and low optimization) is very efficient. We provide more details in Section 4.5.3.

We then execute our approach on the full S3DIS dataset, including varying problematic cases of which non-planar ceiling, stairs, heavy noise, heavy occlusion, false-labelled data, duplicate points, clutter, non-planar walls (See Appendix B for examples). This is a very good dataset for getting a robust indicator of how well a semantic segmentation approach performs and permitted to identify several failure cases as illustrated in Figure 38.

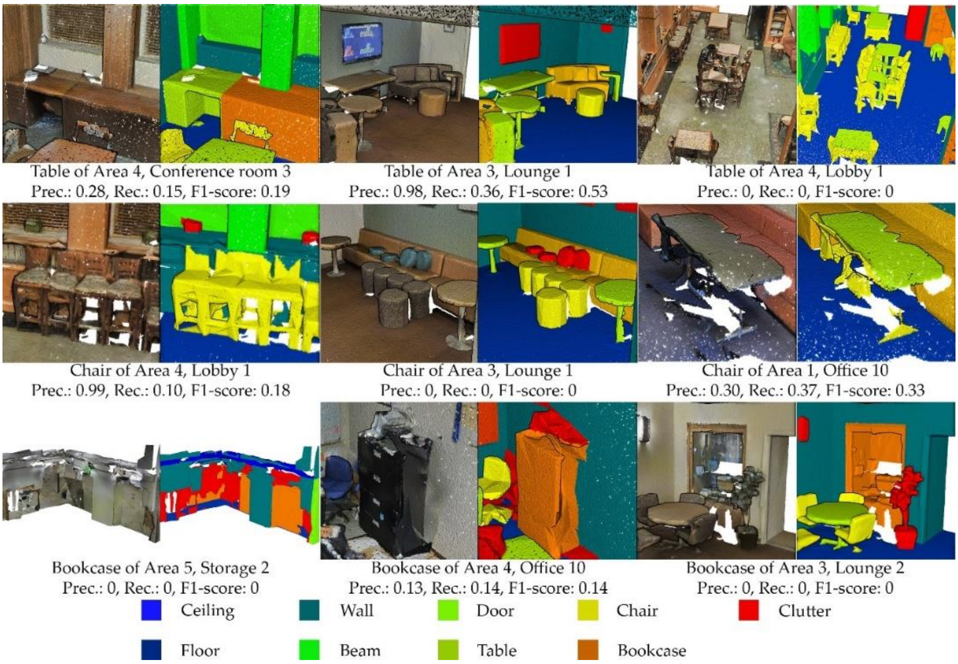


Figure 38. Problematics cases which often include point cloud artefacts such as heavy noise, missing parts, irregular shape geometries, mislabelled data.

We didn't use any training data and our autonomous approach treats points by using only X, Y, Z coordinates. Again, we first use $\overline{IoU}$ metric to get an idea of the worst-case performances achieved by our classifier based on established Connected Elements summarized in Table 13.

Table 13. Benchmark results of our semantic segmentation approach against best-performing deep-learning methods

$\overline{IoU}$	Ceiling	Floor	Wall	Beam	Door	Table	Chair	Bookcase	Clutter
	0	1	2	3	6	7	8	10	12
PointNet [101]	88	88.7	69.3	42.4	51.6	54.1	42	38.2	35.2
MS+CU(2) [82]	88.6	95.8	67.3	36.9	52.3	51.9	45.1	36.8	37.5

SegCloud [160]	90.1	**96.1**	69.9	0	23.1	**75.9**	**70.4**	40.9	42
G+RCU [82]	90.3	92.1	67.9	**44.7**	51.2	58.1	47.4	39	41.9
SPG [161]	**92.2**	95	72	33.5	60.9	65.1	69.5	38.2	51.3
KWYND [88]	92.1	90.4	**78.5**	37.8	**65.4**	64	61.6	**51.6**	**53.7**
Ours	85.4	92.4	65.2	32.4	10.5	27.8	23.7	18.5	23.9

We note that our approach proposes $\overline{IoU}$ scores of 85.4, 92.4 and 65.2 respectively for the ceiling, floor and wall classes. It is within a 3% to 15% range of achieved scores by every state-of-the-art method. This gives enough range for further improvements as discussed in Section 4.6. The 'table' elements present meagre performances explained by looking at (high precision, low recall). Concerning bookcases our approach achieves poorly, partly due to the limitations of the knowledge-based approach. Indeed, the definition of a bookcase in Section 4.3.4 is not very flexible and doesn't allows a search for hybrid structures where clutter on top of a bookcase hides planar patches thus classifying a bookcase as clutter and impacting $\overline{IoU}$ score of both classes. Yet, the ground-truth dataset presents a very high variability and discussable labelling as illustrated in Appendix B. The lowest score achieved concerns doors as identified previously. These elements are often misclassified as clutter, due to their SF1 signature and low SF2 characterization. Overall, our classification approach is comparable to the best deep learning approaches, and the very low computational demand as well as promising improvement flexibility due to the nature of Connected Elements will be further discussed in Section 4.6. Indeed, while the score is in general lower than the best performing deep-learning approaches, this is mainly due to the classification approach.

It is interesting to note that the deep learning architecture in Table 1 make use of colour information, whereas ours solely considers X, Y, Z attributes. A small benchmark is executed to account for this and provided in Table 14.

Table 14. Benchmark results of our semantic segmentation approach against deep-learning methods without any colour information used.

Method	Ceiling 0	Floor 1	Wall 2	Beam 3	Door 6	Table 7	Chair 8	Bookcase 10	Clutter 12
PointNet	84	87.2	57.9	37	35.3	51.6	42.4	26.4	25.5
MS+CU(2)	86.5	94.9	58.8	37.7	36.7	47.2	46.1	30	31.2
Ours	85.4	92.4	65.2	32.4	10.5	27.8	23.7	18.5	23.9

We see that we outperform PointNet when using only X, Y, Z data for ceiling, floor, and wall classes. To better understand where our classifier presents shortcomings, we studied F1-scores per Area and per class to obtain insights on problematic cases and possible guidelines for future works. The analysis can be found in Appendix C.

To summarize SF1SF2 performances, we present in Table 15 and the associated confusion matrix (Figure 39) per class scores over the full S3DIS dataset.

Table 15. Per class metrics for the full S3DIS dataset using our approach

S3DIS class metrics	Ceiling 0	Floor 1	Wall 2	Beam 3	Door 6	Table 7	Chair 8	Bookcase 10	Clutter 12	Average
Precision	0.94	0.96	0.79	0.53	0.19	0.88	0.72	0.28	0.33	0.75
Recall	0.90	0.96	0.79	0.46	0.19	0.29	0.26	0.36	0.47	0.72
F1-score	0.92	0.96	0.79	0.49	0.19	0.43	0.38	0.31	0.39	0.72

We note that we obtain in average a precision score of 0.75, a recall score of 0.72 thus a F_1-score of 0.72. These are relatively good metrics considering the complexity of the test dataset, and the naïve classification approach.

The largest improvement margin is linked to the 'door' and 'bookcase' classes as identified earlier and confirmed in Table 15. While for horizontal planar-dominant classes being ceiling and floor, the F_1-scores of 0.92 and 0.96 give little place for improvement. It orients future work toward problematic cases handling (presented in Appendix B), and irregular structures targeting.

The wall class detection scores of 0.79 gives a notable place for improvements, aiming both at a more precise and coherent classification approach. While table and chair precision are relatively good, their recall rate orients future work to better account for the full number of positive samples ignored with the present classification iteration.

Looking at the normalized confusion matrix (denominator: 695 878 620 points in S3DIS), a large proportion of false positives are given to the clutter concerning all classes, which also demands a better precision in the recognition approach.

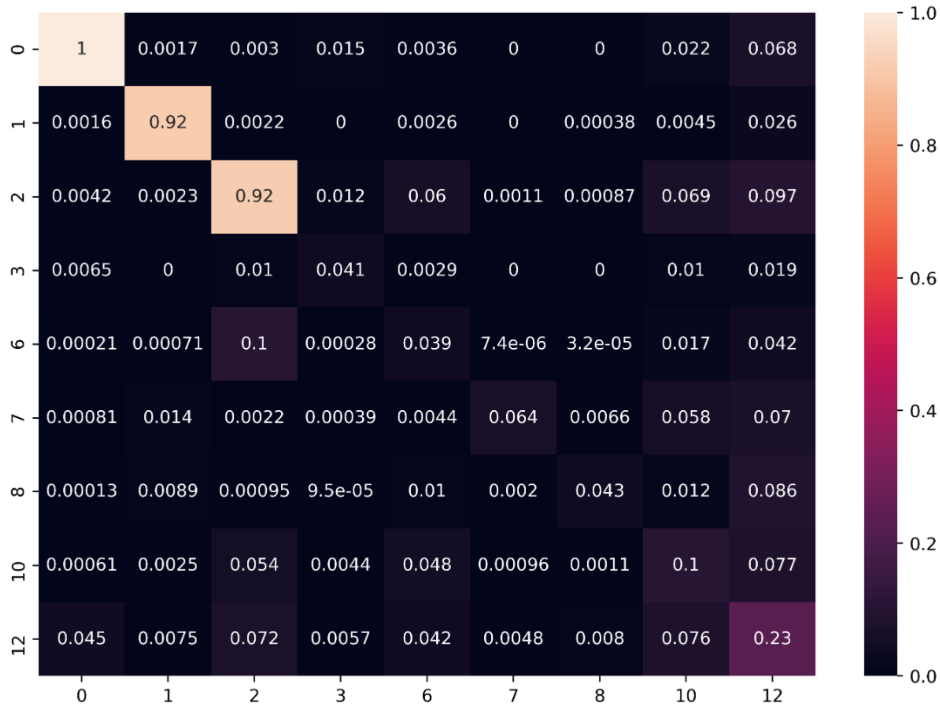


Figure 39. Confusion matrix of our semantic segmentation approach over the full S3DIS dataset

While the above metrics were compared against best-performing deep learning approaches, Table 16 permits to get a precise idea about how good the classifier achieves against the well-performing unsupervised baseline accessible in [133].

Table 16. Overall precision on the full S3DIS dataset against non-machine learning baselines.

Overall precision	**Ceiling 0**	**Floor 1**	**Wall 2**	**Beam 3**	**Door 6**	**Table 7**	**Chair 8**	**Bookcase 10**
Baseline (no colour) [133]	0.48	0.81	0.68	0.68	0.44	0.51	0.12	0.52
Baseline (full) [133]	0.72	0.89	0.73	0.67	0.54	0.46	0.16	0.55
Ours	**0.94**	**0.96**	**0.79**	0.53	0.19	**0.88**	**0.72**	0.2

The used feature sets SF1/SF2 largely outperforms the baseline for the ceiling, floor, wall, table and chair classes permitting satisfying results as illustrated in Figure 40.

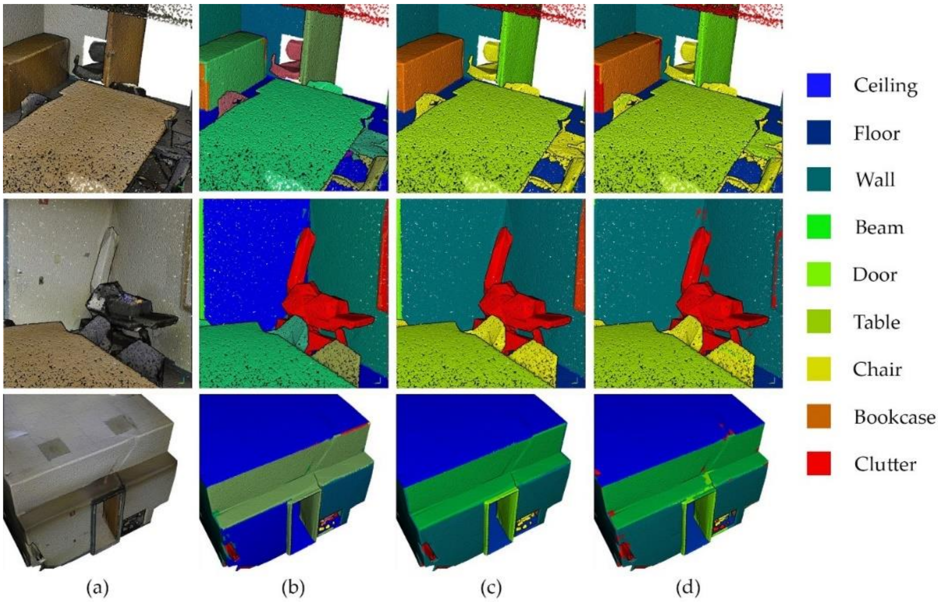


Figure 40. Results of the semantic segmentation on a room sample. (a) RGB point cloud, (b) Connected Elements, (c) Ground Truth, (d) Results of the semantic segmentation

However, we identified issues with the class 'bookcase' and 'door' where our approach performs poorly compared to both the baseline with all features and without the colour. While the door performance is mostly explained by the initial lack of SF2-related connectivity information as stated previously, the latter (bookcase) is partially linked to the variability under which it is found in the dataset and our too specialized classifier (indeed, we consider mostly ground-related bookcases which complicates the correct detection of wall-attached open bookcases). We thus noticed that several points were tagged as bookcase whereas they specifically are desks, or clutter (E.g. Appendix B).

4.5.3 Implementation and performances details

The full autonomous parsing module was developed in Python 3.6. A limited number of libraries were used in order to easily replicate the developing environment thus the experiments and results. As such several functions were developed and will be accessible as open source for further research. All the experiments were performed on a 5 years old laptop with a CPU Intel Core i7-4702HQ CPU @ 2.20Ghz, 16 Gb of RAM and an Intel HD Graphics 4600. As currently standing (no optimization and no parallel processing), the approach is quite efficient and permit to process in average 1.5 million points per minute. This allows offline computing to include in server-side infrastructures. If we compare its performance to a state-of-the-art approach like [161] (2018), our approach is 54% faster, does not necessitate any GPU and does not need any (important) training data.

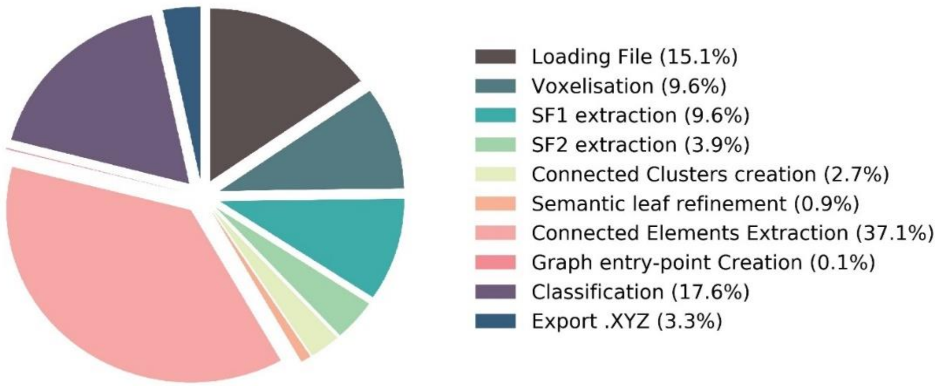


Figure 41. Relative temporal performances of our automatic semantic segmentation workflow

By looking at the relative temporal performances (Figure 41), we note that the first computational hurdle is the creation of Connected Elements. This is mainly explained by the amount of handled points without any parallel computing, which can majorly reduce the needed time. Then, it is followed by the classification approach, but as our main goal was to provide a strong 3D structural connectivity structure for a Smart Point Cloud parsing module, we did not targeted classification performances. Loading/Export times can be reduced if input files are in the .las format. The voxelisation approach and following steps until the semantic leaf extraction can also be parallelized for better performances. In the current version, 1.5 million points per minute are processed on average using the above configuration without any GPU acceleration. It uses around 20% of the CPU and 900 Mb of RAM under full load. As it stands, it is therefore deployable on low-cost server-side infrastructures while giving the possibility to process in average 90 million points per hour.

4.6 DISCUSSION

From the detailed analysis provided in Section 4.5 we first summarize identified strengths in the sub-section 4.6.1 and then we propose 5 main research directions for future work addressing limitations in sub-section 4.6.2.

4.6.1 *Strengths*

First, the presented method is easy to implement. It is independent from any high-end GPUs, and in its current state mainly leverage the processor and the Random-Access Memory (around 1 Gb). This is crucial for a large number of companies that do not possess high-end servers, but rather web-oriented (no GPU, low RAM and intel Core processors). As such, it is easily deployable on a client-server infrastructure, without the need to upgrade the server-side for offline computations.

Secondly, the approach is majorly unsupervised, which gives a great edge over (supervised) machine learning approaches. Indeed, there is currently no need for a huge amount of training data, thus avoiding any time-consuming process of creating (and gathering) labelled datasets. This is particularly beneficial if one wants to create such a labelled dataset, as the provided methodology will speed-up the process by recognizing the main "host" elements of infrastructures leaving mainly moveable elements supervision.

Third, on top of such a scenario, the approach gives acceptable results for various applications that mainly necessitate the determination of structural elements. As such, it can be used for extracting the surface of ceilings, walls or floors if one wants to make digital quotations; it can provide a basis to extract semantic spaces (sub-spaces) organized regarding their function; it can be used to provide a basis for floor plans, cut and section creation…

Fourth, the provided implementation delivers adequate performances regarding the time needed for obtaining results. As it stands, without deep optimizations, it permits offline automatic segmentation and classification, and the data structure provides a parallel-computing support.

Fifth, there is a low input requirements which only necessitate unstructured X, Y, Z datasets, contrary to benchmarked Deep Learning approaches that leverage colour information and provides a complete directed graph of the relations within CELs or classified objects. This information permits reasoning services to use the semantic connectivity information between objects and subspaces for advanced queries using both spatial and semantic attributes.

Finally, the unsupervised segmentation and rule-based classification is easily extensible by improving the feature determination, enhancing the performances or providing a better and more flexible classifier. For example, one can differentiate clutter based on connectivity and proximities to further enhance the classification (E.g., clutter on top of a table may be a computer; clutter linked to the ceiling and in the middle of the room is a light source …). Some of these potentials are addressed as research tracks for future works, as presented in the following sub-section 4.6.2.

4.6.2 Limitations and research directions

First, we note that the new relational features are very useful for the task of semantic segmentation. Plugged to a basic knowledge-based graph, it permits good planar-elements detection such as floor, ceiling and wall. At this point, it is quite useful for the creation of Connected Elements as all the remaining points cover mainly remaining "floating" elements, which can then be further defined through classification routines. This is a very interesting perspective for higher end specialization per application, were the remaining elements are then looped for accurate refinement depending on the level of specialization needed, as expressed in [117]. Future work will also further study learning-based feature extraction such as the ones presented in [167,168] proposing a design of the shape context descriptor with spatially inhomogeneous cells. The parsing methodology can also be extended through other domain ontologies such as the CIDOC-CRM as presented in [4], which highlight the flexibility to different domains.

Secondly, the creation of links between CEL is a novelty which provides interesting perspectives concerning reasoning possibilities that plays on relationships between elements. Indeed, the extracted graph is fully compatible with the semantic web and can be used as a base for reasoning services, and provide counting possibilities such as digital inventories [4] and semantic modelling [5]. Additionally, the decomposition in primary, secondary, transition and rest elements is very useful in such contexts as one can specialize or aggregate elements depending on the intended use and application [108]. Indeed, the approach permits to obtain a precise representation of the underlying groups of point contained within Connected Elements and homogenized in Semantic Patches.

Third, the extended benchmark proved that untrained schemes can reach comparable recognition rate to best-performing deep learning architectures. Particularly, detecting the main structural elements permits to achieve a good first semantic representation of space, opening the approach to several applications. However, the scores for 'floating' CEL (moveable elements) is poor in its current version. Shortcomings are linked to the naïve knowledge-based classifier which lacks flexibility/generalization in its conception and gives place for major improvements in future works. Specifically, it will undergo an ontology formalization to provide a higher characterization and moving thresholds to better adapt the variability in which elements are found in the dataset.

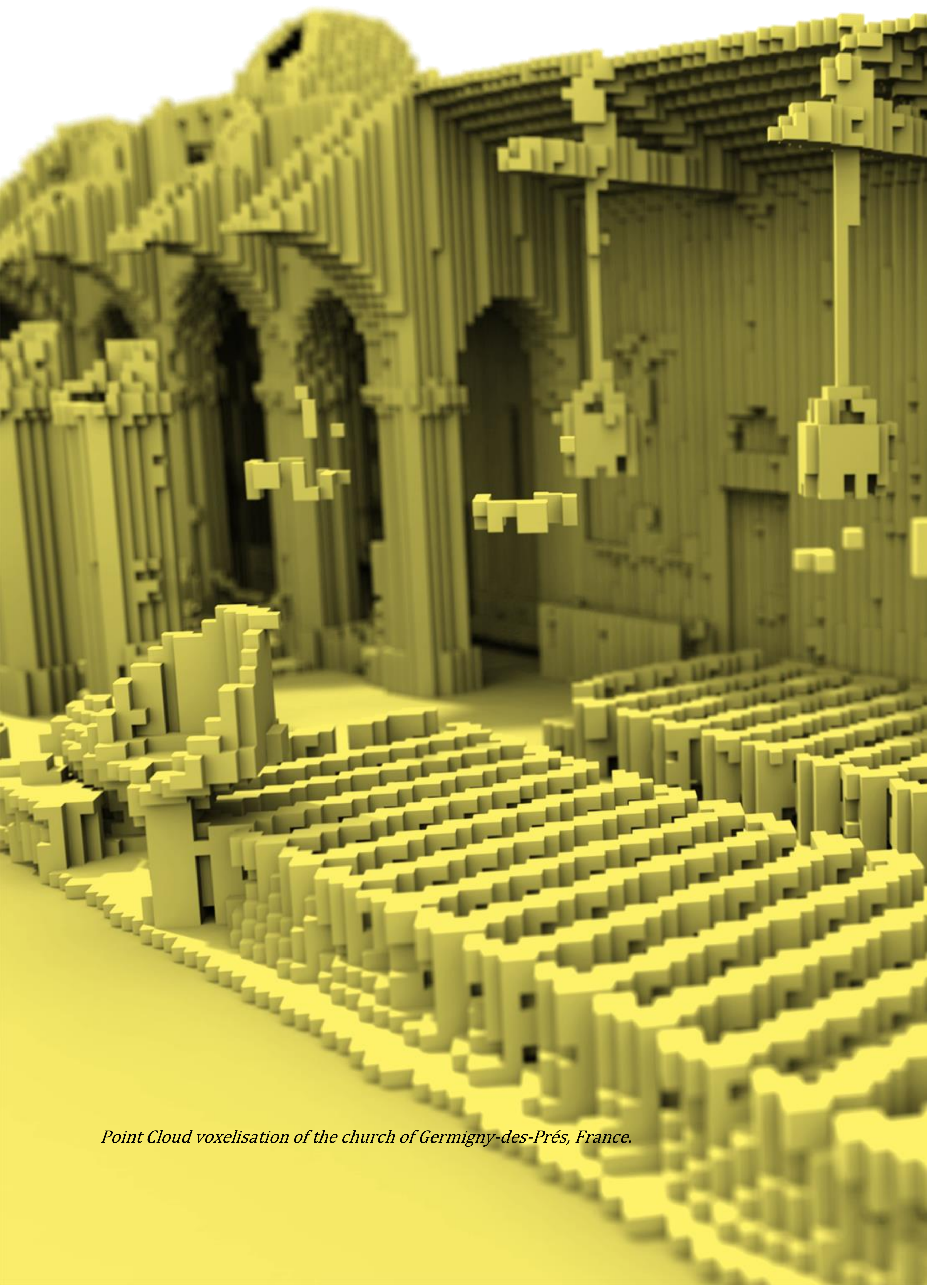
Fourth, some artefacts and performances hurt the approach due to the empirical octree-based voxelization determination and enactment, but as it stands, it provides a stable structure robust to aliasing and block effect at the borders. Further works in the direction of efficient parallel computing will permit both an increase in time performances and deeper depth tree selection (thus better characterization). Also, the octree definition will be looked for variable octree depth depending on pre-define sensor-related voxel leaf size. Other possibilities include using a local voxelated structure such as proposed in [69] to encode the local shape structure into bit string by point spatial locations without computing complex geometric attributes. On the implementation side, while the dependency to voxelization is limited due to the octree structure to allow a constant point density per voxel in average, it will be further studied to avoid exponential time explosion when changing the deepness level. As such, the structure is already ready for parallel computing and it will be studied in future works.

Finally, while our dedicated approach was tested on the S3DIS dataset, it can easily be adapted to other point clouds which provide an additional research direction. The approach will be tested against indoor and outdoor point clouds from different sensors and the classification adapted to account for various well-established classes. As such, a large effort is currently undergoing to create accurate labelled datasets for AEC and outdoor 3D mapping applications, to be shared as open-data.

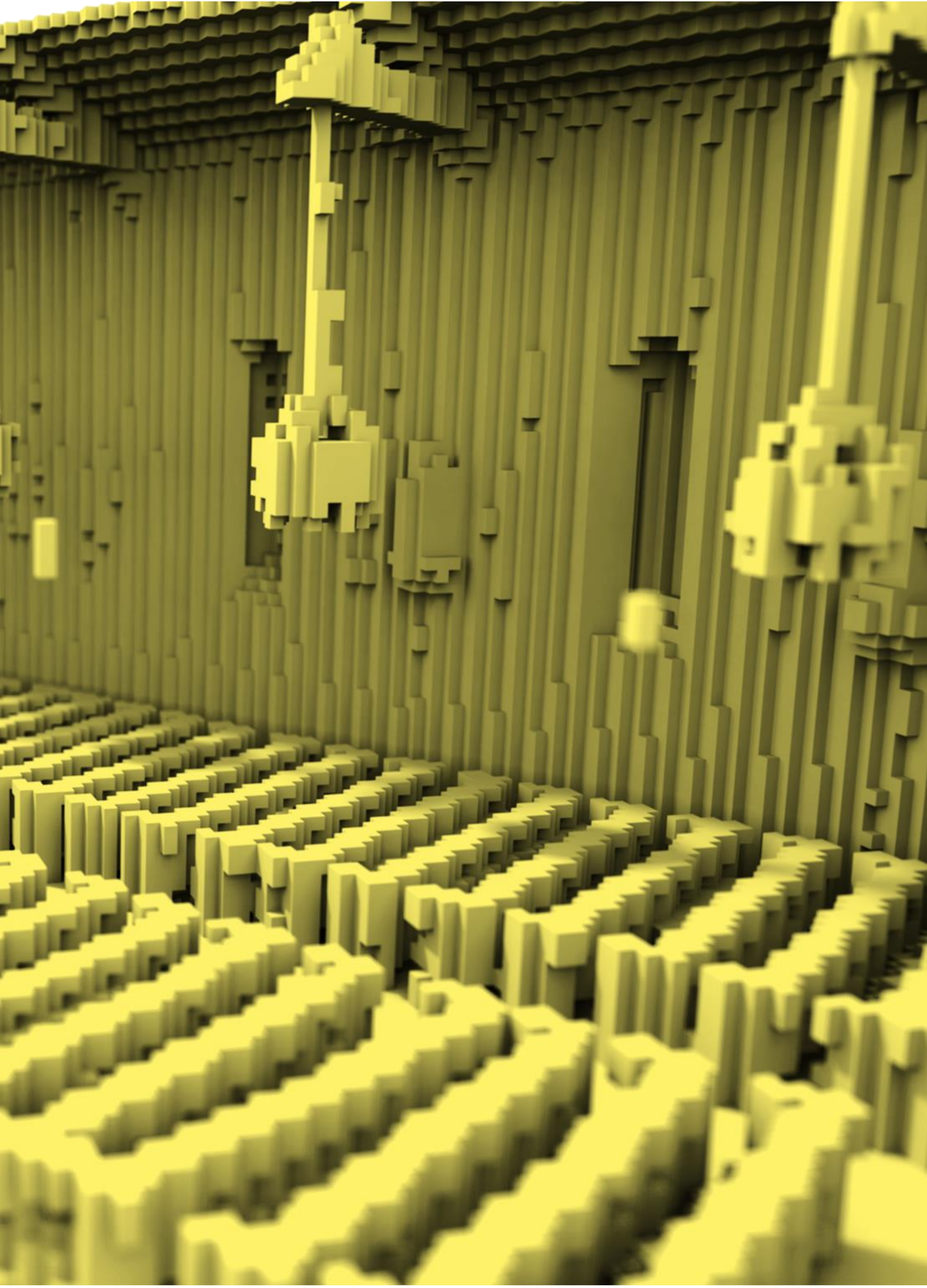
Our focus is driven by a general global/local contextualization of digital 3D environments where we aim at providing a flexible infrastructure which should be able to scale up to different generalization levels. As such, the proposed unsupervised segmentation approach in Connected Elements and Semantic patches acts as a standard module within the Smart Point Cloud Infrastructure and permit to obtain a full autonomous workflow for the constitution of semantically rich point clouds [2].

4.7 CONCLUSIONS

In this article, a point cloud parsing module for a Smart Point Cloud Infrastructure was presented. It provides a semantic segmentation framework that groups points in a voxel-based space where each voxel is studied by analytic featuring and similarity analysis to define semantic clusters, that retain highly representative SF1 and SF2 signatures. This process is conducted regarding an initial connected component from multi-composed graph representations after automatically detecting different planar-dominant elements leveraging their prevalence in man-made environments. A classification approach to automatically detect main classes in the S3DIS dataset and to obtain a measure of performance against best-performing deep learning approaches is provided. While the method is well performing for floor, ceiling and wall classes, extended research is needed if one wants to use the classification as a robust approach for moveable elements detection.



Point Cloud voxelisation of the church of Germigny-des-Prés, France.



CHAPTER 5

Application to archaeology

Chapter 5 – Application to archaeology 120

5.1 Introduction 126
5.2 Digital Reconstruction in Archaeology: A Review 128
5.2.1 Archaeological Field Work 128
5.2.2 Integration of 3D Data 132
5.3 Materials and Methods 135
5.3.1 Point Cloud Data Acquisition and Pre-Processing 136
5.3.2 Knowledge-Based Detection and Classification 139
5.4 Results 147
5.5 Discussion 153
5.6 Conclusions 160

CHAPTER'S PREFACE

The previous chapter 4 permitted to establish a parsing methodology, interoperable enough that it can be used for different applications at different scales. E.g., it was the starting-point of a knowledge-based classification approach for indoor point cloud data, aimed at the detection of specific infrastructure-related elements and assets. In this chapter 5, we will dive into the applicability of the SPCI – specifically the Parsing module and the Point Cloud Database module – to a different domain: archaeology. After largely reviewing the usage of 3D point clouds within archaeological frameworks, we provide a methodology to use formalized knowledge to guide the classification approach through adequate language manipulation. As such, this chapter targets the classification module, the language and query process module and the Visualisation & Interaction module as highlighted in Figure 42.

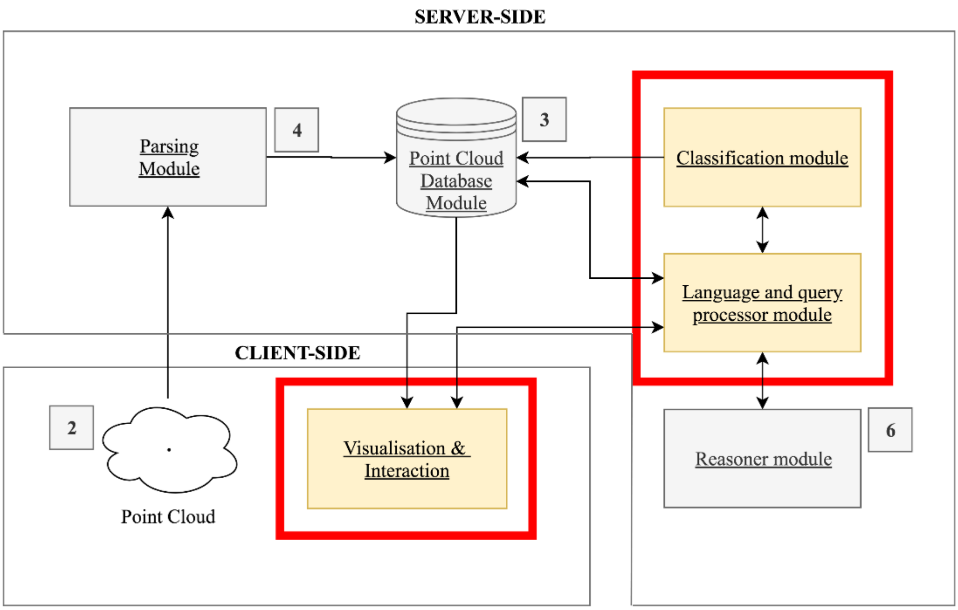


Figure 42. Chapter 5: Extension of the Smart Point Cloud Infrastructure to the archaeological domain

Details about the integration and compatibility of each module are also given. It also permit to test on another scale with different data artefacts (see chapter 2) and other reality capture sensors the integration of dense point cloud data within the SPCI. The following chapter 6 is linked to chapter 4 and presents the non-described modules so far.

Based on Article [4]

3D Point Clouds in Archaeology: Advances in Acquisition, Processing and Knowledge Integration Applied to Quasi-Planar Objects

Abstract: Digital investigations of the real world through point clouds and derivatives are changing how curators, cultural heritage researchers and archaeologists work and collaborate. To progressively aggregate expertise and enhance the working proficiency of all professionals, virtual reconstructions demand adapted tools to facilitate knowledge dissemination. However, to achieve this perceptive level, a point cloud must be semantically rich, retaining relevant information for the end user. In this paper, we review the state of the art of point cloud integration within archaeological applications, giving an overview of 3D technologies for heritage, digital exploitation and case studies showing the assimilation status within 3D GIS. Identified issues and new perspectives are addressed through a knowledge-based point cloud processing framework for multi-sensory data and illustrated on mosaics and quasi-planar objects. A new acquisition, pre-processing, segmentation and ontology-based classification method on hybrid point clouds from both terrestrial laser scanning and dense image matching is proposed to enable reasoning for information extraction. Experiments in detection and semantic enrichment show promising results of 94% correct semantization. Then, we integrate the metadata in an archaeological smart point cloud data structure allowing spatio-semantic queries related to CIDOC-CRM. Finally, a WebGL prototype is presented that leads to efficient communication between actors by proposing optimal 3D data visualizations as a basis on which interaction can grow.

Keywords: point cloud; data fusion; laser scanning; dense image-matching; feature extraction; classification; knowledge integration; cultural heritage; ontology

Published: 30 September 2017
Type: Open Access Article
Publisher: MDPI
Journal: Geosciences
Special Issue: Remote Sensing and Geosciences for Archaeology

5.1 INTRODUCTION

Gathering information for documentation purposes is fundamental in archaeology. It constitutes the groundwork for analysis and interpretation. The process of recording physical evidence about the past is a first step in archaeological study for a better understanding of human cultures. In general, the goal is to derive spatial and semantic information from the gathered and available data. This is verified in various sub-disciplines of archaeology that rely on archaeometry [169]. In this setting, remote sensing is particularly interesting as a means to not only safely preserve artefacts and their context for virtual heritage [170], but also to complement or replace techniques presenting several limitations [171].

An archaeological breakthrough given by this technique is the moving of interpretation from the field to a post-processing step. The possibility to gather massive and accurate information without transcripts interpretation or in situ long presence is a revolution in archaeological workflows. It started with stereo-vision and photogrammetry to derive 3D information, but recent development deepened the representativity of digital 3D data through higher resolution, better accuracy and possible contextualization [172]. The study of materials is often linked with on-site related information, forever lost if not correctly transmitted. Digital preservation is therefore necessary to document a state of the findings, and this at different accessible temporal intervals. Visions shared by [173,174] for the digital documentation and 3D modelling of cultural heritage states that any project should include (1) the recording and processing of a large amount of 3D multi-source, multi-resolution, and multi-content information; (2) the management and maintenance of the 3D models for different applications; (3) the visualization of the results to share the information with other users allowing data retrieval "through the Internet or advanced online databases"; (4) digital inventories and sharing "for education, research, conservation, entertainment, walkthrough, or tourism purposes". In this paper, we propose such a solution.

The information as we see it is mostly 3D: "when we open our eyes on a familiar scene, we form an immediate impression of recognizable objects, organized coherently in a spatial framework" [1]. Therefore, tools and methods to capture the 3D environment are a great way to document a 3D state of the archaeological context, at a given time. Analogous to our visual and cognitive system, 3 steps will condition the completeness of the surveyed object. First, the perception, i.e., how the visual system processes the visual information to construct a structured description of the shape of the object/scene. Second, the shape recognition or how the product of the perceptual treatment will contact stored representations in the form of known objects (it will construct a perceptual depiction that will be a representation of the same nature stored in memory). Finally, the identification (labelling), i.e., when a stored structural representation is activated, it will in turn activate the unit of meaning (concept) that corresponds to it, located in the semantic system. Sensors are the analogue to our perception and aim at extracting the visual stimuli it is sensitive to (spatial information, colour, luminance, movement, etc.). At this stage, neither the information on the shape of the object nor the label is extracted. In the case of 3D remote sensing, the quality of observation is therefore critical to enable high quality and relevant information extraction about the application. As such, the sensory perceptive processing capable of extracting visual primitives must be as objective and complete as possible, making sensors for point cloud generation favourable. Constituted of a multitude of points, they are a great way to reconstruct environments tangibly, and enabling further primitive's extractions (discontinuity, corners, edges, contour,, etc.) as in our perceptive visual system (described in [1]). However, their lack of semantics makes them a bona-fide [127] spatial representation, thus of limited value if not enhanced.

Deriving semantic information is fundamental for further analysis and interpretation. This step is what gives a meaning to the collected data and allows to reason on sites or artefacts. All this information must be retained and structured for a maximum interoperability. In an archaeological context, many experts must share a common language and be able to exchange and interpret data through ages, which necessitate the creation of formalized structure to exchange such data. Multiple attempts were made, and the CIDOC Conceptual Reference Model (CRM) is a formalization that goes in this direction. "It is intended to promote a shared understanding of cultural heritage information by providing a common and extensible semantic framework that any cultural heritage information can be mapped to. It is intended to be a common language for domain experts and implementers to formulate requirements for information systems and to serve as a guide for good practice of conceptual modelling. In this way, it can provide the semantic glue needed to mediate between different sources of cultural heritage information, such as that published by museums, libraries and archives". It is used in archaeology such as in [175] and provides semantic interoperability. Ontologies offer considerable potential to conceptualize and formalize the a-priori knowledge about gauged domain categories [176] that relies primarily on expert's knowledge about real world objects. If correctly aggregated and linked to spatial and temporal data, digital replicas of the real world can become reliable matters of study that can survive through times and interpretations, which reduces the loss or degradation of information related to any site study in archaeology.

However while promising structures and workflow provide partial solutions for knowledge injection into point clouds [117,177], the integration, the maturation state as well as the link between semantic and spatial information is rudimentary in archaeology. Concepts and tools that simplify this process are rare, which complicates the merging of different experts' perceptions around cultural heritage applications. Being able to share and exchange contextual knowledge to create a synergy among different actors is needed for planning and analysis of conservation projects. In this context, we explore ways to (1) better record physical states of objects of interest; (2) extract knowledge from field observations; (3) link semantic knowledge with 3D spatial information; (4) share, collaborate and exchange information.

This paper is structured in a dual way to provide both a background of 3D used techniques in archaeology, and technical details of the proposed point cloud workflow for quasi-planar heritage objects.

In the first part, we carefully review the state of the art in digital reconstruction for archaeology. This serves as a basis to identify research perspectives and to develop a new methodology to better integrate point clouds within our computerized environment.

Secondly, we propose a framework to pre-process, segment and classify quasi-planar entities within the point cloud based on ontologies, and structure them for fast information extraction. The methodology is illustrated on the case of the mosaics of Germigny-des-Prés (France) and then applied to other datasets (façade, hieroglyphs). Finally, the results are presented, and we discuss the perspectives as well as data visualisation techniques and WebGL integration.

5.2 DIGITAL RECONSTRUCTION IN ARCHAEOLOGY: A REVIEW

3D digital exploration and investigations are a proven way to extract knowledge from field observations [178,179]. The completeness and representativity of the 3D data gathered by sensors are critical for such digitalization. Equally, methodologies, materials and methods to "clone" a scene are important for the extensiveness of any reconstruction. The 3D-capturing tools and software drastically evolved the last decade; thus, we review the current state of the art in digital reconstruction for archaeology.

5.2.1 Archaeological Field Work

Even if an increasing number of archaeological contributions deal with 3D and related management of information, archaeologists are still sceptical about 3D technologies and often use manual drawings for cautious observations and first analyses on the field [180]. The literature gravitates around a controversial or diverging hypothesis which illustrates this reticence to adopt new technologies in remote sensing [179,181]. During an empirical recording of monuments or sites, measurements are taken (by hand), taking distances between characteristic points on the surface of the monument. The definition of the coordinates is done on an arbitrary coordinate system on a planar surface of the structures. The method is simple, reproducible and low-cost but limiting factors such as limited accuracy, time demand and necessary direct contact makes it unfavourable in many scenarios including for inaccessible areas. However, archaeologists will often use such an approach over remote sensing to gather insights that are otherwise considered incomplete. The 3D methods are frequently regarded as intricate, expensive and not adapted to archaeological issues [182]. On most sites, for buildings studies or in excavations, the data gathering and acquisition are made with drawings and pictures in 2D. In some cases, 3D can come after the analyses process and is used as a "fancy" means to present results and rebuilt a virtual past. As noticed by Forte [183]: "there was a relevant discrepancy between bottom-up and top-down processes. The phase of data collecting, data-entry (bottom-up) was mostly 2D and analogue, while the data interpretation/reconstruction (top-down) was 3D and digital". However, the new possibilities given by 3D remote sensing extend the scope of possible conservation and analysis for digital archaeology, and can progressively move to post-processing a part of the interpretation process, making the underlying data (if complete) the source on which different reading and conclusions can be mined. Yet, such techniques cannot replace a field presence when complementary semantics (from other senses such as hearing, taste, smell, touch) are necessary.

The different data types from these remote sensing platforms played a vast role in complexifying the diversification in methodologies to derive the necessary information from the data (data-driven). However, the 3D spatial data extracted from the bottom-up layer for most of these techniques are surveyed points, in mass, creating point clouds. They are driven by the rapid development of reality capture technologies, which become easier, faster and incur lower costs. Use cases in archaeology show the exploration and acceptation of new techniques, which are assessed not only in regard to their accuracy, bust mostly in accordance to their fit to a specific context, and the associated costs. Following the categorization defined in [181], we distinguish "(1) the regional scale, to record the topography of archaeological landscapes and to detect and map archaeological features, (2) the local scale, to record smaller sites and their architecture and excavated features, and (3) the object scale, to record artefacts and excavated finds". In their article, the authors reviewed some passive and active sensors for 3D digitization in archaeology at these different scales. They conclude that the principal limiting factor for the use of the different remote sensing technologies reviewed (Synthetic Aperture Radar (SAR) interferometry, Light Detection and Ranging (LiDAR), Satellite/Aerial/Ground imagery, Terrestrial Laser Scanning (TLS), Stripe-projection systems) is the ratio added value of a digital 3D documentation over the time and training that inexperienced users must invest before achieving good results. In their paper, [184] state that 3D recording is the first step to the digitization of objects and monuments (local and object scales). They state that a 3D recording method will be chosen depending on the complexity of the size and shape, the morphological complexity (Level of Detail—LoD), and the diversity in materials. While this is accurate looking purely at a technical replication, other factors such as user experience, available time or budget envelope will constrain the instrument or technique of choice. They propose a 9-criteria choice selection as follows: cost; material of digitization subject; size of digitization subject; portability of equipment; accuracy of the system; texture acquisition; productivity of the technique; skill requirements; compliance of produced data with standards. While this extends the global understanding and 3D capture planning, it lacks a notion of time management (implied in productivity) or constraints in line with contextual laws and regulations (no contact survey only, etc.). Although they separate “accuracy” from “texture acquisition”, both can be related, as well as additional features provided by the sensors (e.g., intensity) that can extend the criterion table.

We note a large discrepancy between scales of the remote sensing and related costs/methods tested for point cloud generation.

At a regional scale, airborne LiDAR is sparsely used in archaeology, mostly as a 2.5 D spatial information source for raster data analysis. It is a powerful tool to analyse past settlement and landscape modification at a large scale. Use cases such as in [185–188] helped remove preconceptions about settlements size, scale, and complexity by providing a complete view of the topography and alterations to the environment, but while it provided new research and analysis directions, the LiDAR data did not leverage 3D point clouds considered too heavy and too raw to provide a source of information.

At both the local scale and the object scale, several use cases exploit active sensoring, specifically terrestrial laser scanners (TLS) using phase-based and time-of-flight technologies [32]. Archaeological applications vary such as in [189] to reconstruct a high-resolution 3D models from the point cloud of a cave with engravings dating back to the Upper Palaeolithic era, in [190] to study the damage that affected the granitic rock of the ruins of the Santo Domingo (Spain), or in [191] for the 3D visualization of an abandoned settlement site located in the Central Highlands (Scotland). More recent procedures make use of TLS to reconstruct the Haut-Andlau Castle (France) [192], or in [193] to map the Pindal Cave (Spain). These showed that to capture fine geometric details, laser-scanning techniques provide geometric capabilities that have not yet been exceeded by close-range photogrammetry, especially when concave or convex forms need to be modelled. Rising from the static concept, Mobile laser scanning (MLS) [33] has scaled up the data rate generation of TLS by allowing dynamic capture using other sensors including GNSS position and inertial measurements for rapid street point cloud generation and public domain mapping. New concepts and technology including Solid State LiDAR and simultaneous localization and mapping (SLAM) have pushed dynamic acquisition for quickly mapping with a lower accuracy the surroundings, extending cases using HMLS (Hand-held mobile laser scanning) [34], MMS (Mobile Mapping System) [35], or more recently MMBS (Mobile Mapping Backpack System) [36]. At the object scale, active sensors namely for active triangulation, structured light and computer tomography for 3D modelling is widely used due to its high precision, and adaptation to small isolated objects [27]. Moving to ground technologies, surveys are precise in detecting sub-surface remains. Different geophysical processing techniques and equipment (such as ground penetrating radar (GPR), magnetometry and resistivity) are usually integrated together, to increase the success rate of uncovering archaeological artefacts, for example in [194] to delineate the extent of the remains of a small town that has been submerged (Lake Tequesquitengo, Mexico).

Passive sensing gained a lot of attention in the heritage community following terrestrial use cases and image crowdsourcing, allowing a wide range of professionals and non-expert to recreate 3D content from 2D poses (exhaustive software list from [16,17,195,18–25]). The rapidly growing interest for light aerial platforms such as UAV (Unmanned Aerial Vehicle) based solutions and software based on multi-view dense image matching [15,71] and structure from motion [13] swiftly provided with an alternative to active sensoring. Use cases for 3D archaeological and heritage reconstruction are found at the object scale through terrestrial surveys [10,172,179,196] and the local scale through light aerial platforms, making this technique a favourable way to obtain quick and colour balanced point clouds. Moreover, the cost and accessibility (hardware and software) of dense-image matching reconstruction workflows have allowed its spread in archaeological studies. For example, in the Can Sadurní Cave (Spain), [197] successfully reconstructed an object via dense-image matching and georeferenced the obtained 3D model using TLS point cloud data of the Cave. They state that capture from different positions is fundamental to generate a complete model that does not lack important information.

While reconstruction accuracy is increasing [26], remote sensing via active sensors is favoured in the industry for local scales. There are discussions in which computer vision would replace LiDAR [192,198]; however, practical cases tend to a merging of both (Reconstruction of the Amra and Khar-anah Palaces (Jordan) [199], the castle of Jehay (Belgium) [200]), and predilection applications for each techniques, combining strength of natural light independence with low-cost and highly visual image-based reconstruction [200]. Particularly in the case of mosaics, decorations and ornaments, the combination of features from sensors generating accurate and complementary attributes permits the overcoming of limits arising from a small set of features. Indeed, use case such as in [201] results in a high richness of detail and accuracy when combining TLS and close range photogrammetry which was not achievable otherwise. Thus, multisensory acquisition provides an interesting method that will be investigated.

The high speed and rate generation of 3D point clouds has become a convenient way to obtain instant data, constituting datasets of up to Terabytes, so redundant and rich that control operation can take place in a remote location. However, they often go through a process of filtering, decimation and interpretation to extract analysis reports, simulations, maps, 3D models considered as deliverables. A common workflow in archaeology concerns the extraction of 2D profiles and sections, 2D raster to conduct further analysis or to create CAD deliverables, particularly looking at ornaments, mosaics or façades. This induces several back and forth movements within the pipeline, and the general cohesion, storage system often lack extensibility. This challenge is particularly contradictory, and a solution to automate recognition such as [202] in the context of mosaics would therefore provide very solid ground for tesserae detection, extended to 3D by combining many more sets of features. This will be specifically studied in Section 5.3.

While all the reviewed literature specifically points out the problems linked with data acquisition and summarize the strength and weakness of each regarding the recorded spatial information, few specifically link additional semantic information. Gathered in situ or indirectly extracted from the observation, the measurements often rely on specific interest points. While this is handy looking at one specific application for one archaeologist, this practice is dangerous regarding the problematic of curators and conservation. Indeed, preserving at a later stage the interpretation through sketches, drawings, painting or text description based on interest points makes any possible data analysis impossible from the raw source. Therefore, 3D point selectivity should not arise at the acquisition step, but in a post-processing manner, to benefit of the flexibility given by 3D data archives, which was impossible before the emergence of automatic objective 3D capturing devices.

We postulate that when designing data processing workflow, specific care must be given to the objectivity linked with the spatial data, which multisensory systems and point clouds specifically answer. As such, they can constitute the backbone of any powerful spatial information system, where the primitive is a 0-simplex [203]. Their handling in archaeology, however, is a considerable challenge (often replace by 3D generalization such as meshes, parametric models, etc.) and thus presents many technical as well as interpretation difficulties.

5.2.2 Integration of 3D Data

As demonstrated in [204], the evolution of remote sensing for archaeological research and the acceptance in archaeology has grown linearly since 1999 looking at the number of publications (Sources SCOPUS, ScienceDirect & Web of Science search engines) related to remote sensing per year. While this provides new possibilities, the reliability and heterogeneity of the spatial information are issues in heritage for the conservation, interoperability and storage of data. 3D GIS linked to archaeological databases have been thought and proposed for the management of this information at different scales and on different type of sites such as large excavated sites [205–208]. In their paper [206], the authors discuss the possibility and the ultimate goal of having a complete digital workflow from 3D spatial data, to efficiently incorporate the information into GIS systems while relying on formal data model. After stating the limits and difficulties of integrating efficiently 3D data (as well as time variations), they interestingly express the domain specifications and formalization through ontologies. Within a knowledge system, standards and procedures are key to warrant the consistent meaning of collective contents and to trace the “history” of the processed data [207]. In their use case, they create a 3D model segmented regarding semantic information to allow the independent manipulation as well as GIS query between elements. They claim to provide new standards in 3D data capture to be usable by all archaeologist, but their method is empirically defined and the integration of knowledge sources is blurry regarding segmentation and semantic injection. Building archaeology is also a field where 3D applications are used mainly for conservations purpose [209–211]. In these contributions, the authors highlight two characteristics of archaeological cases: the heterogeneity of data and the difficulty of processing 3D spatial entities from irregular archaeological objects (artefacts, buildings, layers, etc.). Several solutions have been offered in the mentioned papers and specific software have been designed (see [212] for a relevant 3D GIS use case and [213] for the most recent summary). In these studies, the definition of archaeological facts rests on their representation as raster data, specific point of interest, polygons and 3D shapes, but never the direct source of spatial information: point clouds.

At this step, several criterions should be considered to choose the most suitable spatial data model. Many researchers proposed 3D grid representation (voxels) as the most appropriate data format for handling volumetric entities and visualizing continuous events [214,215]. Generalizing point cloud entities by volume units such as a voxels allows 3D GIS functionalities such as object manipulation, geometry operations and topology handling regarding [125]. Although Constructive Solid Geometry (CSG) and 3D Boundary representations (B-rep) can roughly depict a spatial entity, the level of generalization of the underlying data has an impact on the accuracy and representativity of GIS functionalities' results. Thus, point cloud brings an additional flexibility by giving the possibility to recover the source spatial data information. The limits with available commercial and open source database GIS systems (which are mostly used by archaeologist) made point clouds a secondary support information for primarily deriving 3D model generalizations. This of course limits the conservation potential of archaeological findings, as the interpretation behind data modelling workflows is unique and irreversible (one-way). As such, to our knowledge, no 3D archaeological GIS system is directly based on 3D point cloud. They are rather considered heavy and uninterpretable datasets. Furthermore, the constitution and leveraging of knowledge sources is still limited, with some experiences by manual injection reviewed in [205–208]. This of course constitutes a major issue that needs to be addressed for scaling up and generalizing workflows. The different literature involved in the constitution of 3D GIS delineates the need of standardization, especially regarding the variety of data types. In this direction, one specific use case in [206] demonstrated that the main advantage of the 3D GIS methodology is the link between attribute information to discrete objects defined by the archaeologist. Their implementation is done regarding the CIDOC-CRM ISO 21127 standard and the design patterns from the ontological model of the workflow of the Centre for Archaeology to achieve semantic compatibility. As opposed, the approach presented in [216] allows linking of 3D models of buildings and graph-based representation of terms. It describes its domain-linked morphology to provide new visual browsing possibilities. In this approach, one expert creates a graph for one specific application. This allows the comparison of semantic descriptions manually established by experts with divergent perspectives but lacks extensibility to match general rules. Indeed, while the description flexibility within one field can benefit from this, it can lead to interoperability problems when a formalization needs to be established, especially regarding geometrical properties or for structuring the semantics according to a pattern. Even though there are several works dedicated to ontology-based classifications of the real-world entities, the ontologies developed so far are rarely integrated with the measurements data (physical data). As such, [217] proposes an observation-driven ontology that plays on ontological primitives automatically identified in the analysed data through geo-statistics, machine-learning, or data mining techniques. These provide a great standpoint to semantic injection and will be further studied. In particular, the possibility to specialize the ontology through extensions such as CIDOC CRMba (an extension of CIDOC CRM to support buildings archaeology documentation) or CIDOC CRMgeo (an extension of CIDOC CRM to support spatiotemporal properties of temporal entities and persistent items) provides new solutions for higher interoperability.

The literature review showed a shift at an acquisition phase toward better means to record physical states of an environment, an object. TLS and dense-image matching showed an increase in popularity, and their combination provide new and promising ways to record archaeological artefacts and will thus be investigated. However, both methods generate heavy point clouds that are not joined directly with knowledge sources or structured analogously to GIS systems. Rather, their use is limited to providing a reference for other information and deliverables (2D or 3D). While this is a step forward toward higher quality documentation regarding other reviewed field methods, this is not a long-term solution when we look at the evolution of the discipline, the quantity of generated data and the ensuing ethics. The identified problem concerns the link between domain knowledge and spatial information: it evolves in parallel, partially intersects or is hardcoded and manually injected. Moreover, the flexibility regarding possible analysis is often null due to interpreted documents that force a vision over elements that no longer physically exist, or which were poorly recorded. Therefore, a strong need for ways to integrate knowledge to point clouds is essential. This "intelligence spring" is categorized regarding 3 sources as identified in [117], being device knowledge (i.e., about tools and sensors), analytic knowledge (i.e., about algorithms, analysis and their results) and domain knowledge (i.e., about a specific field of application). Their rapprochement to point clouds is, however, a bottleneck that arises early in the processing workflow. If we want to better integrate point clouds as intelligent environments [100], we must correctly assemble knowledge sources with the corresponding "neutral" spatial information. This relies on different procedures to (1) pre-process the point cloud, (2) detect the entities of interest within the initial point clouds and (3) attach the knowledge to classify and allow reasoning based on the classification. As such, our work proposes to leverage the use of ontologies as knowledge sources, as well as defining a workflow to directly process and integrate point clouds within 3D GIS systems, creating virtual heritage [170]. In the next part, we describe our technical method for integrating knowledge within reality-based point-clouds from TLS and close-range photogrammetry. While the following methodology can be extended to different applications with examples such as in Section 5.5, it is illustrated and applied to quasi-planar objects of interest.

5.3 MATERIALS AND METHODS

The applied workflow of object detection and classification is organized as follows: in the data pre-processing step, the different point clouds are treated using the procedure described in Section 5.3.1 (Step 1, Figure 43). Subsequently, point cloud descriptors as well as object descriptors such as the extent, shape, colour and normal of the extracted components are computed (Step 2, Figure 43) and imported into the next classification procedure using a converter developed in this study (Step 3, Figure 43). In the last step, the objects are classified based on the features formalized in the ontology (Step 4, Figure 43).

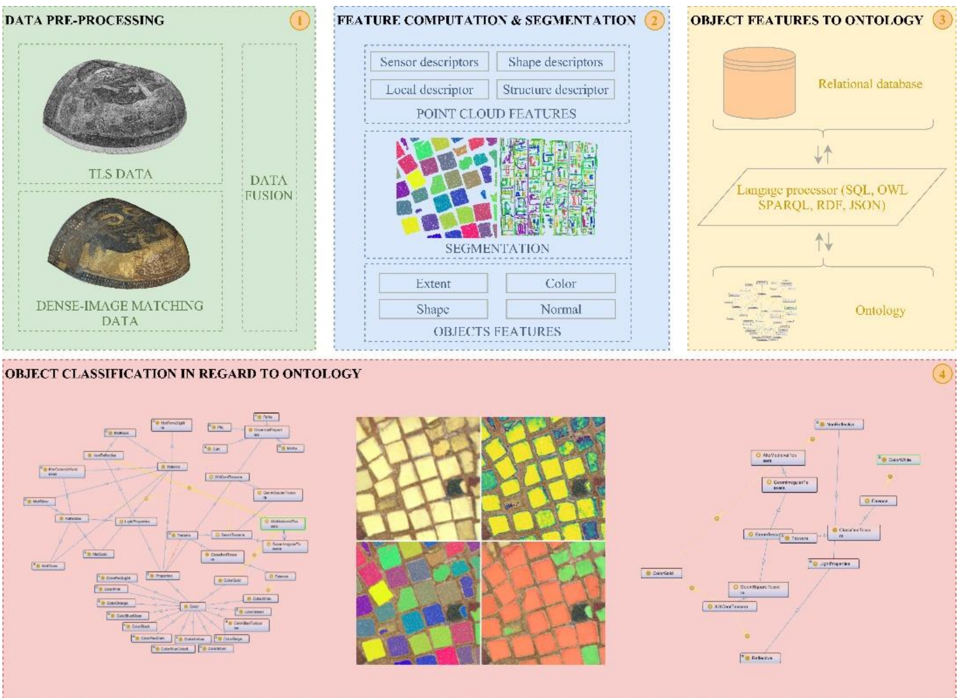


Figure 43. Overview of the methodology developed in this paper. (1) Data pre-processing; (2) Feature computation and segmentation; (3) Object features to ontology; (4) Object classification in regard to ontology

The process to identify features of interest within the signal is the foundation for the creation of multi-scale ensembles from different datasets. The work described in [218] extensively reviews data fusion algorithms defined by the U.S department of Defense Joint Directors of Laboratories Data Fusion Subpanel as "a multilevel, multifaceted process dealing with automatic detection, association, correlation, estimation and combination of data and information from single and multiple sources to achieve refined position and identify estimates, and complete and timely assessments of situations and threats and their significance". The combination of different sensors generating complementary signatures provides pertinent information without the limitations of a single use and creates a multisensory system [219]. Thus, following the postulate of the state of the art, we decide to adopt a multi-sensory workflow for maximizing information (Section 5.2.1)

5.3.1 Point Cloud Data Acquisition and Pre-Processing

A pre-processing step is necessary to obtain a highly representative signal of the value measured as defined in [53]. Indeed, to avoid external influential sources that degrade the information, this step demands adapted techniques to minimize errors including noise, outliers and misalignment. Filtering the data strongly depends on device knowledge [117].

Several sets of data from various contexts were acquired to perform different tests. The Carolingian church located in Germigny-des-Prés (Loiret, France) houses ancient mosaics dating from the 9th century, composed of about one hundred thousand tesserae (the average surface of a tessera is 1 cm^2, square of 1 cm by 1 cm). The preserved works offer a unique opportunity for the study of mosaics and glass. Indeed, the tesserae that composes it are mainly made in this material, which is rare in the archaeological context of the early Middle Ages [220]. However, part of the mosaic was restored in the 19th century; therefore, tesserae are from two periods; thus, we must first distinguish the different tesserae types (based on their age) for accessing alto-medieval glass information. The study could reveal important predicates, considering each tessera taken independently or by analysing different properties, while conjecturing with expert's domain knowledge. The mosaic of the vault culminates at 5402 m above the ground, presenting many challenges for 3D capture from active and passive sensors. The dome is protected, and the limited accessibility tolerates only a light scaffolding, too narrow for the positioning of tripods, illustrating the need to adapt means to the context Figure 44.

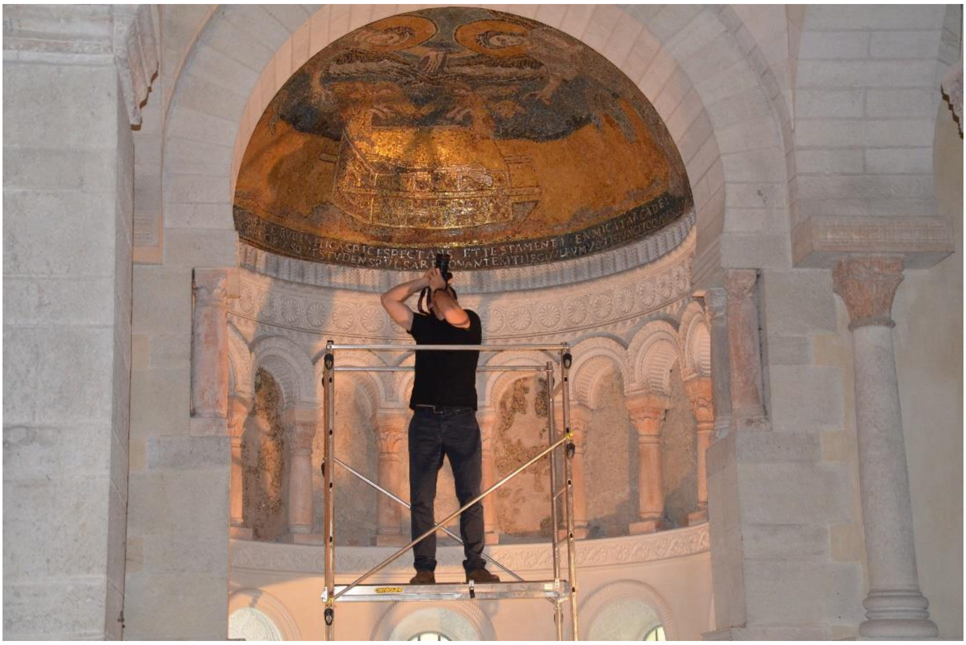


Figure 44. The vault of Germigny-des-prés being captured for dense-image matching processing.

The first sample was acquired using a phase-based calibrated terrestrial laser scanner: the Leica P30. The different scans were registered using 1338 reflective targets, of which 127 were shot by a total station (Leica TCRP1205, accuracy of 3 mm + 2 ppm) and used for indirect georeferencing afterwards. The mean registration error is 2 mm, and the mean georeferencing deviation is 2 mm (based on available georeferenced points measured from the total station). Two point cloud segments of the same zone (mosaic) were extracted: one unified point cloud that includes measurements from 8 different positions with varying range and resolutions, and one high resolution point cloud (HPC) from one optimized position by using an extended mounted tribrach. A comparison emphasized the influence of the angle of incidence and the range over the final resolution, precision and intensity of the point cloud. Thus, we chose the HPC for its higher representativity (Figure 45).

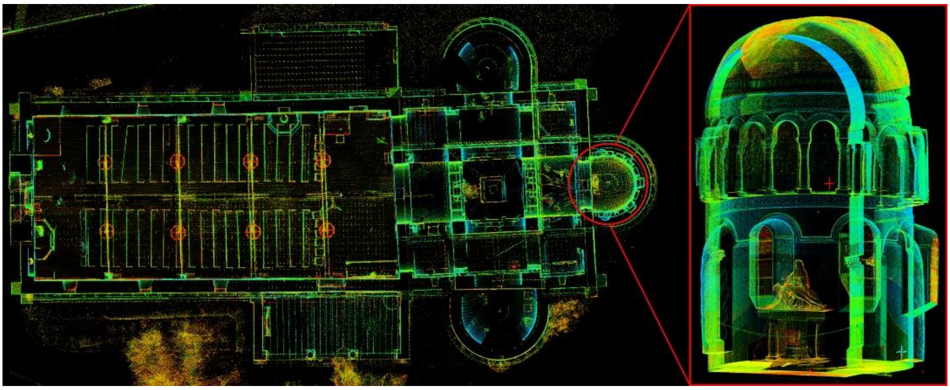


Figure 45. Point cloud of the church of Germigny-des-Prés. Top View (left) and zone of interest (right).

The TLS was operated at 1550 nm for a maximum pulse energy of 135 NJ. Initial filtering was conducted such as deletion of intensity overloaded pixels (from highly retro-reflective surfaces) and mixed pixels to flag problematic multi-peak scan lines and keep the right return via full-waveform analysis. The final accuracy of a single point at 78% albedo is 3 mm. The final HPC is composed of 30,336,547 points with intensity ranging from 0.0023 to 0.9916, and covers solely the mosaic. Several pictures were taken at different positions to obtain a 3D point cloud of the mosaic. These pictures were shot using a Canon EOS 5D mark III camera equipped with a 24–105 mm lens. In total, 286 pictures of 5760 × 3840 in RAW, radiometrically equalized and normalized, were used to reconstruct the photogrammetric point cloud (Figure 46).

Figure 46. Point cloud from a set of 2D poses reconstructed via dense-image matching using the software ContextCapture v 4.4.6, Bentley Systems, Exton, United States [17].

The knowledge around the acquisition methodology provides important information as missing/erroneous data, misadjusted density, clutter and occlusion are problems that can arise from an improper or impossible capture configuration on the scene [38], resulting in a loss of transmitted information or data quality. Combining different sensors with diverse acquisition methodologies allows the overcoming of this challenge and provides a better description of the captured subject through (1) higher quality features (i.e., better colour transcription, better precision, etc.); (2) specific and unique attribute transfer; (3) resolution and scale adaptation, sampling or homogenizing [221]. The knowledge extracted from a device, analytical knowledge or a domain formalisation constitutes the fundamental information repository on which a multi-level data structure is constructed (Section 5.3.2).

The first step is therefore to correctly reference point clouds, known as data registration. The method is derived from previous work to perform accurate attribute transfer [200]. The main idea is that a priority list processing is established and influences data fusion regarding knowledge. When combining different point clouds, their geometry and attributes in overlapping areas are then properly addressed. The complementary information needs to be combined from the different available sources if relevant, keeping the most precise geometry as a structure. Avoiding heterogeneous precisions is essential, leading to point deletion rather than point caching and fusing. Once correctly registered, every point cloud data source goes through a pixel and attribute level fusion (if not previously fused at the sensory level).

5.3.2 Knowledge-Based Detection and Classification

Our approach for object extraction relies on domain knowledge that relays through point cloud features. Segmentation [76] and feature extraction are well studied areas within point cloud processes. However, the integration of knowledge is still rare, with few example of hybrid pipelines [72,222]. Our proposed approach constitute a hybrid method inspired by previous work in shape recognition [148,223–225], region growing pipelines [38,226,227] and abstraction-based segmentation [228–232] relying on 3D connected component labelling and voxel-based segmentation. As such, different features presented in Table 1 constitute the base for segmentation.

Table 17. Point features computed from the point cloud data after data fusion, before segmentation.

Type	Point Features	Range	Explanation
Sensor desc.	X, Y, Z	Bounding-box	Limits the study of points to the zone of interest
	R, G, B [1]	Material Colour	Limited to the colour range that domain knowledge specifies
	I		Clear noise and weight low intensity values for signal representativity

Shape desc.	RANSAC [2]	-	Used to provide estimator of planarity
Local desc.	Nx, Ny, Nz [3]	$[-1,1]$	Normalized normal to provide insight on point and object orientation
	Density [4]	-	Used to provide insights on noise level and point grouping into one object
	Curvature	$[0,1]$	Used to provide insight for edge extraction and break lines
	KB [5] Distance map		Amplitude of the spatial error between the raw measurements and the final dataset
Structure desc. [6]	Voxels	-	Used to infer initial spatial connectivity

[1] [233]; [2] [149]; [3] [234]; [4] [234]; [5] Knowledge-based; [6] [117].

The point cloud data processing was implemented using the programming interfaces and languages MATLAB, Python, SQL, SPARQL, OWL, Java as well as the C++ Library CCLib from CloudCompare [230] and the software Protégé [63].

First, the point cloud is segmented regarding available colour information by referring to the database table containing float RGB colour ranges for each material composing the dataset. Then, the gap is enhanced by superimposing intensity values over colour information (this allows us to refine and better access point filtering capabilities) as in Equation (1).

$$R_e = R \times I,\ G_e = G \times I, B_e = B \times I \tag{1}$$

A statistical outlier filter based on the computation of the distribution of point to neighbour distances in the input dataset similarly to [235] is applied to obtain a clean point cloud. This step can be avoided if the colour range and the datasets are perfectly in line.

The segmentation developed is a multi-scale abstraction-based routine that decomposes the 3D space in voxels at different abstraction levels and by constructing an octree structure to speed up computations. The three-dimensional discrete topology (3DDT) proposed by [236] generates a voxel coverage by intersection with another representation model (parametric or boundary) of an object. This is possible by playing on the different configurations of voxel adjacencies. A voxel has 6 neighbour voxels by one face, 18 neighbour voxels by a face or an edge and 26 neighbours by a face, an edge or a vertex. Our approach is based on a 26-connectivity study that groups adjacent voxels if not empty (i.e., voxels containing points). It is conditioned by the analytical knowledge where the density information constrains the initial bounding-box containing the point cloud. An initial low-level voxel structure is then computed, retaining the number of points as attribute. Let $v_I \in \mathbb{R}^3$ be a voxel. Let v_I be its neighbour voxel. We define V_{CEL} as the connected element (segment) as in Equation (2):

$$\forall\ v_i \in R^3, \exists\ v_j = n(v_i) \mid V_{CEL} = \left[v_I, v_j\right] \leftrightarrow v_j \neq \emptyset \tag{2}$$

where $n(v_i)$ is the neighbour voxel of and v_i, V_{CEL} is the group segment from a 26-connectivity adjacency study.

The topological grouping also permits us to clean the remaining noise N from difficult colour extraction regarding the equations Equations (3) and (4). Let $p_n \in \mathbb{R}^3$ be the *n*-th point of V_{CEL}. There exists P_{CEL} as follows:

$$\forall\, p_1, \ldots, p_n \in \mathbb{R}^3, \exists\, P_{CEL} \mid P_{CEL} = \{p_1, \ldots, p_n\}\ ^1 \tag{3}$$

$$P_{CEL} = N \leftrightarrow S_{Number\_CEL} < d(P_{CEL}) \times \min(S_m)\ \&\ S_{Size\_CEL} < \min(V_m)\ ^2 \tag{4}$$

[1] where p is a point (x, y, z) in space, [2] where N is the remaining noise, S_{Number_CEL} is the number of points in P_{CEL}, $d(P_{CEL})$ is the point density of P_{CEL}, $\min(S_m)$ is the minimum of the surface of the material S_m, S_{Size_CEL} is the voxel volume occupancy of the CEL, $\min(V_m)$ is the minimum of the volume of the material V_m; therefore, N is the group composed of fewer points than the knowledge-based assumption from the density achievable from the sensor, the minimum surface of the object and the minimum volume of the object.

Then, our multi-scale iterative 3D adjacency algorithm at different octree levels recursively segments under-segmented groups (detected by injecting analytical knowledge regarding minimum bounding-box size of processed material as in Equation (4)), refining the voxel-based subdivision until the number of generated voxels is inferior to the density-based calculation of estimated voxel number. When subgroups dimensions correspond to material's available knowledge, segments are added to the "Independent Tesserae" segments. Otherwise, a convolution bank filter is applied regarding the longest side of the calculated best fit P.C.A Bounding Box. For absorbent materials or objects sensitive to the sensor emitting frequency (implies low intensity, thus high noise), the 3D distance map as in Table 1 is used to detect points that belong to each object of interest. The accuracy of the extracted segments is assessed by ground truth manual counting of different samples.

Then, on each detected segment, every point is projected on the RANSAC best fit plane, and we extract the 2D outline through the convex hull of the projected points p_I to constrain the plane. Let P_{pCEL} be the projected points of P_{CEL} onto the best fit plane. Then, we obtain $\mathrm{Conv}(P_{pCEL})$ as in Equation (5):

$$\mathrm{Conv}(P_{pCEL}) = \left\{ \sum_{I=1}^{|P_{pCEL}|} \alpha_I \times p_i \mid \forall\, p_I \in P_{pCEL}, \forall\, \alpha_I \geq 0) : \sum_{I=1}^{|P_{pCEL}|} \alpha_I = 1 \right\} \tag{5}$$

where $\mathrm{Conv}(P_{pCEL})$ is the convex polygon of P_{CEL} as a finite point set (x, y) in R^2 (x, y, z). It can be extended to 3D, nD if necessary.

We calculate the compactness (CS) and complexity (CP) of the generated polygon in regard to the work of [237], as well as its area, and its statistically generalized (gaussian mixture) colour (Table 2).

Table 18. Segment features computed from the extracted segments.

Type	Point Features	Range	Explanation
Sensor generali-zation	Xb, Yb, Zb	barycentre	Coordinates of the barycentre
	Rg, Gg, Bg [1]	-	Material unique colour from statistical generalization
	I	-	Intensity unique value from statistical generalization
Shape desc.	CV [2]	-	Convex Hull, used to provide a 2D shape generalization of the underlying points
	Area	-	Area of the 2D shape, used as a reference for knowledge-based comparison
	CS, CP	[0,1]	Used to provide insight on the regularity of the shape envelope
Local generalization	Nx, Ny, Nz [3]	[−1,1]	Normalized normal of the 2D shape to provide insight on the object orientation

[1] [233]; [2] Convex-hull; [3] [234].

The final classification of the delineated objects is based on the available and constituted domain ontology of point cloud features for archaeology. The idea behind the ontology is that the integrated cultural information from a variety of sources is brought together into an integrated environment where we can ask broader questions than we can ask from individual pieces.

Ontologies can be expressed using different knowledge representation languages, such as the Simple Knowledge Organization System (SKOS), the Resource Description Framework (RDF), or the Web Ontology Language 2 (OWL2) specification. These languages contrast in terms of the supported articulateness. The SKOS specification, for instance, is widely used to develop thesauri, the CIDOC CRM is mainly used for describing heritage sites, the Basic Formal Ontology [238] at a higher conceptualization level to incorporates both 3D and 4D perspectives on reality within a single framework. The OWL2 ontology language is based on the Description Logics (DL) for the species of the language called OWL-DL. DL thus provides the formal theory on which statements in OWL are based and then statements can be tested by a reasoner. The OWL semantics comprise three main constructs: classes, properties and individuals. Individuals are extensions of classes, whereas properties define relationships between two classes (Object Properties), an individual and a data type (Data Properties).

We used the OWL and the RDF languages to define ontologies for their high flexibility and interoperability within our software environment (Protégé & Java). As for the study of mosaics, the ontology is set upon the point cloud data and its attributes, thus indirectly leveraging domain ontologies. Indeed, sensor related knowledge is needed to understand the link between features and their representativity. The following meta-model is formalised in UML and provides a conceptual definition for implementations. We therefore used the model to provide a clear vision and comprehension of the underlying system, but the ontology creation slightly differs from privilege performances; therefore, adaptations are made at the relation scheme modelling level.

The characterization (knowledge representation and data modelling) in Figure 47 is a Level-2 domain meta-model, that can plug to a Smart Point Cloud structure [108]. The general idea is that different hierarchical levels of abstraction are constituted to avoid overlapping with existing models and to enhance the flexibility and opening to all possible formalized structure. The core instruction is that the lower levels are closer to a domain representation than higher levels (level-0 being the higher level), but they impose their constraints. The overall structure is a pyramidal assembly, allowing the resolution of thematic problems at lower levels with reference to constraints formally imposed by the higher levels.

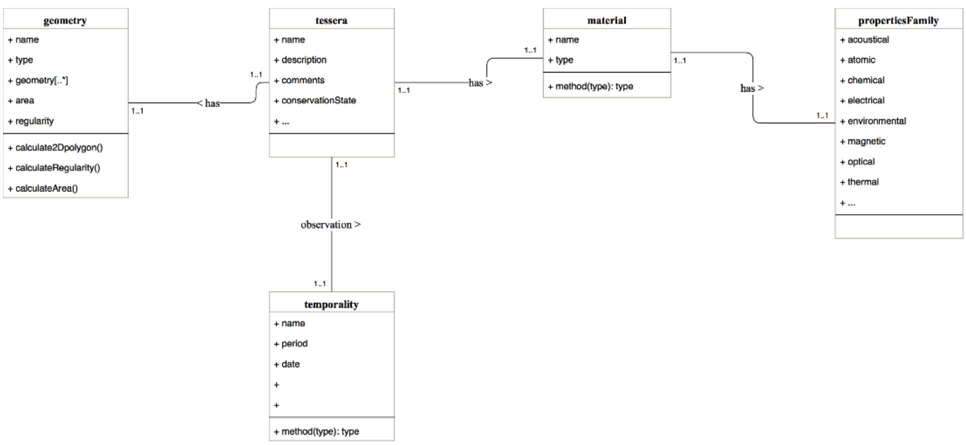


Figure 47. UML meta-model of the ontology. Tesserae have one or multiple geometries, which are characterized by their regularity (determined by the ontological reasoning framework), and an area. Tesserae also have a temporality (characterized as a time interval, being placed at early Middle Ages or during a restoration at the 19th century) and different materials. These materials retain various properties including light sensitivity.

The ontology implementation was structured as triplets. Each triplet corresponds to a relation (subject, predicate and object), which expresses a concept. The end goal is to reason based on the constituted ontology to extract information about the tessera geometric regularity, its material and temporal classification (ClassifiedTessera) as in Figure 48.

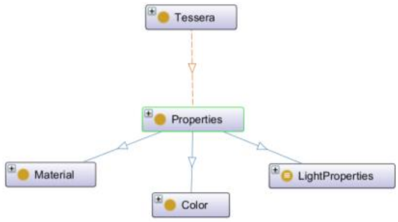


Figure 48. Sub-ontology for the classification of point cloud tessera objects. Blue arrows represent links regarding the tree structure (these are "subClassOf"). The oranges links represent the "hasProperty" relationship that we created to describe the relationship between a Tessera and its properties. It is a simple relationship from domain (Tessera) to Range (Properties).

The ontology is then populated with the domain knowledge as detailed in Figure 49, and the different predicates are established to obtain a final classification of the point cloud. Note that a tolerance of 20% regarding the definition of geometries was used to allow relative variations within one tesserae family. Analogously, any quasi-planar object may be substituted and described by the afore-mentioned properties, thus extending the provided ontology.

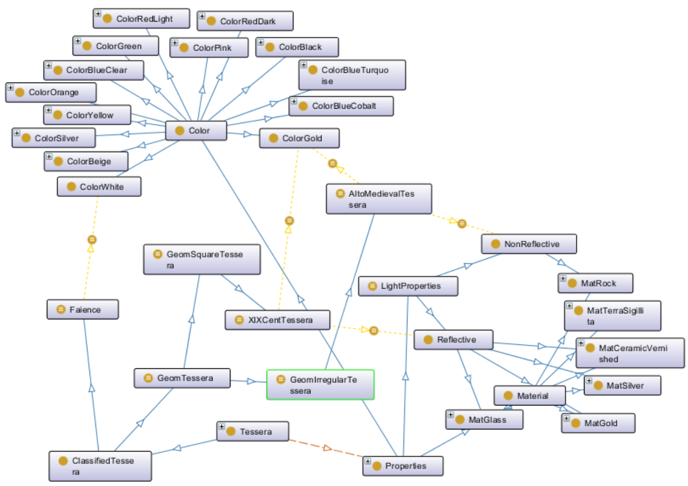


Figure 49. Detailed ontology for the classification of the mosaic's point cloud. The yellow lines are the links of sub-assumptions, of reasoning. They are in fact links of equivalence between a class and its definition.

The different results allow to classify the point cloud, after determining the regularity of the 2D outline regarding different constraints (examples in Table 3).

Table 19. Example of tessera classification using RDF constraints.

RDF Triple Store	Effect
((CS some xsd:double[> "1.1"^^xsd:double]) or (CS some xsd:double[< "1.05"^^xsd:double])) and (CP some xsd:double[> "4.0E-4"^^xsd:double])*	Tessera is irregular *(1)*
(CP some xsd:double[<= "4.0E-4"^^xsd:double]) and (CS some xsd:double[>= "1.05"^^xsd:double, <= "1.1"^^xsd:double])	Tessera is square
(1) and (hasProperty some ColorGold) and (hasProperty some NonReflective) and (Area some xsd:double[<= "1.2"^^xsd:double])	Tessera is alto-medieval
(hasProperty some ColorWhite) and (Area some xsd:double[>= "16.0"^^xsd:double, <= "24.0"^^xsd:double])	Material is Faience

The domain knowledge including size, geometry and spatial distribution leads to object classification. For enhancing its interoperability, the developed DSAE (Digital Survey-based Architectural Element) ontology can directly be extended using the well-established CIDOC-CRM formal ontology. Indeed, the CIDOC-CRM is purely descriptive, and does provide only "factual" tests (a node is linked to an arc, which is linked to another node). The provided DSAE ontology can reason based on complex declaration of conditions (such as AND, OR, ONLY, etc.), thus is much more structured than the CIDOC-CRM, and allows to reason. As such, it permits automatic classification that can be plugged to the CIDOC-CRM enabling archaeologists to better understand the underlying point cloud data. In the case of tesserae, each tessera material is then considered as a E57 Material specialization which comprises the concept of materials (Specialization of E55 Type), LightProperties and Color can be seen as S9 Property (it describes in a parametric way what kind of properties the values are) and Area, CP and CS as SP15 Geometry attributes (which comprises the union of geometric definitions and the linked declarative places) from the extension CIDOC-CRMgeo (based on the ontology GeoSPARQL), as in Figure 50.

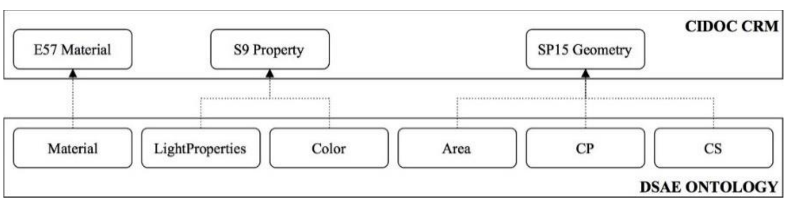


Figure 50. Connectivity relationship between the CIDOC-CRM and the DSAE ontology to extend interoperability and allow descriptive knowledge for archaeologist to be included.

Finally, semantic information is transferred to the point cloud that can be used for information extraction. Once extracted semantics have been successfully linked to the spatial information, we address structuring for interaction purposes. The data structuration is made in regard to [108]. The main idea is that the structure is decomposed in three meta-models acting at three different conceptual levels to efficiently manage massive point cloud data (and by extension any complex 3D data) while integrating semantics coherently. The Level-0 describes a meta-model to efficiently manage and organise pure point cloud spatial data information. The Level-1 is an interface between the level-0 and the level 2 (specific domain-based knowledge). As such, the data integration methodology relies on incorporating the point cloud data in the Smart Point Cloud data structure [108] in regard to the workflows described in section 5.3. The structuration therefore follows the object decomposition, where points of each object are grouped together to form world objects (i.e., Independent Tesserae) once concepts and meaning have been linked. This constitutes the entry point of the ontology which acts as a Level-2 specialization to inject relevant knowledge. To facilitate the dissemination of information, query results from specific queries need to be visualized properly. For users to access and share a common viewpoint result of a semantic query, we enhanced the approach in [239] by applying over each object (i.e., tessera) one unique colour per instance for each class (e.g., faience pieces); all non-requested tesserae are coloured in black as in Figure 51.

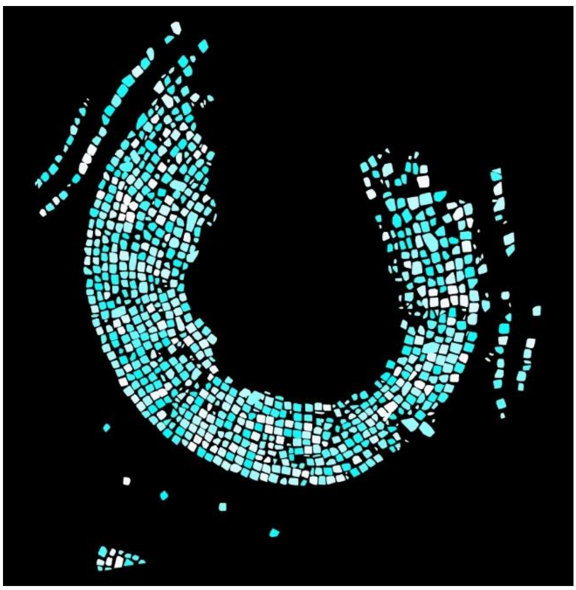


Figure 51. Unique colourisation of a group of golden tesserae (bottom-up view).

For each class of object, we compute a bounding box and we locate its centre. The bounding box centre becomes the centre of the sphere on which the camera will move to determine the optimal camera position. The coordinates of the camera on the sphere are computed according to the following formulae [239]:

$$X = x_{center} + r \times \cos(\phi) \times \cos(\Theta), Y = y_{center} + r \times \sin(\phi), Z = z_{center} + r \times \cos(\phi) \times \sin(\Theta)^{1} \quad (6)$$

[1] where X, Y and Z are the camera coordinates, x_{center}, y_{center} and z_{center} are the coordinates of the centre of the sphere, r is the radius of the sphere, ϕ is the vertical angle, Θ is the horizontal angle.

From each camera position, we compute the number of visible tesserae from the user request observed in the produced image. Since each instance of one sort of tesserae is coloured uniquely, the algorithm performs by counting the number of different pixels colours. Hence, the number of distinct colours in the image corresponds to the number of tesserae seen from this camera position. The camera position that maximises the view of requested tesserae corresponds to the optimal viewpoint. If two camera locations present the same number of observed tesserae, we apply a maximisation criterion regarding the pixels to determine the optimal camera position.

5.4 RESULTS

We tested the method on different samples from different zones of the mosaic to identify the influence of the segmentation and the classification in different scenarios, as well as another point cloud from terrestrial laser scanner captured in Jehay (Belgium). To assess the quality of the segmentation, knowledge-based tessera ground truth was extracted from the point cloud and compared to the segmentation method extracts. Results (Table 4) show an average 95% segmentation accuracy for point cloud gold tesserae, 97% for faience tesserae, 94% for silver tesserae and 91% for coloured glass.

Table 20. Segmentation accuracy of tesserae samples.

Tesserae	**Segmentation Number of Points**		**Accuracy**
	Ground truth	Tesserae C.	
	Gold		
Sample NO. 1	10,891	10801	99%
Sample NO. 2	10,123	11,048	91%
Sample NO. 3	10,272	10,648	96%
Sample NO. 4	11,778	12,440	94%
	Faience		

Sample NO. 1	27,204	28,570	95%
Sample NO. 2	23,264	22,978	99%
Sample NO. 3	23,851	24,440	98%
Sample NO. 4	22,238	22,985	97%
	Silver		
Sample NO. 1	1364	1373	99%
Sample NO. 2	876	931	94%
Sample NO. 3	3783	3312	88%
Sample NO. 4	1137	1098	97%
	C. Glass		
Sample NO. 1	1139	1283	87%
Sample NO. 2	936	1029	90%
Sample NO. 3	821	736	90%
Sample NO. 4	598	625	95%

The tesserae recognition pipeline including segmentation, classification and information extraction was conducted over 3 different representative zones of the point cloud to be exhaustive and to be able to count manually each tessera for assessing the results. In the first zone containing 12,184,307 points, three types of tesserae were studied: 138 Gold tesserae from the 19th century renovation (NG), 239 ancient gold (AG) and 11 faience pieces (FT) (Figure 52). The automatic segmentation correctly recognized all FT (100% accuracy) and 331 golden tesserae (GT) (88% accuracy), remaining ones being 5% of under-segmentation (in groups of 2/3 tesserae), 7% of tesserae not detected. The classification correctly labelled respectively 100% FT, 98% NG, and 99% AG.

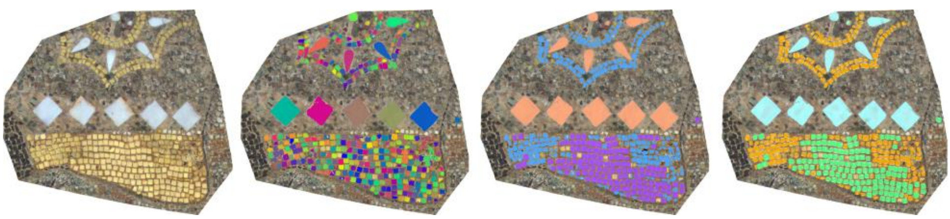


Figure 52. Zone 1: Classification workflow of tesserae in Zone 1. From left to right: Colour point cloud; abstraction-based segmented point cloud; classified point cloud; 2D geometry over point cloud.

In the second zone containing 12,821,752 points, 313 gold tesserae (195 NG and 118 AG) and 269 silver tesserae (ST) were processed. In total, 284 (91%) golden tesserae were correctly segmented, of which 93% were correctly labelled NG and 95% AG, and 93% of ST were correctly segmented, of which 87% were correctly labelled. The third larger sample composed of 34,022,617 points includes 945 gold tesserae and 695 CG (coloured glass) tainted in black. The other tesserae in the sample had an insufficient resolution for ground truth generation. In total, 839 (89%) golden tesserae were correctly segmented, of which 86% were correctly labelled NG and 95% AG. Concerning CG, (494) 71% were correctly segmented, and 98% were correctly labelled. While classification results are very high, segmentation is heavily influenced by the quality of the data; hence, CG shows lower results because of its harsh sensor representation (tesserae are not easily discernible).

Globally, 59,028,676 points and 2610 tesserae were processed; 2208 (85%) were correctly detected and segmented, of which 2075 (94%) were correctly labelled (Table 5).

Table 21. Recapitulation of tesserae detection results.

ID	Tesserae		Segmentation		Classification		Res.
	Type	**Nb**	**Nb**	**%**	**Nb**	**%**	**Nb**
1	NG	138	331	88%	131	98%	7
	AG	239			196	99%	43
	FT	11	11	100%	11	100%	0
2	NG	155	284	91%	128	93%	27
	AG	158			139	95%	19
	ST	269	249	93%	216	87%	53
3	NG	396	839	89%	297	86%	99
	AG	549			471	95%	78
	CG	695	494	71%	486	98%	209
	Total	2610	2208	85%	2075	94%	535

The full workflow was also conducted over a point cloud acquired using a TLS at a local scale to detect specific stones and openings of the façade of the castle of Jehay (Belgium). The point cloud comprises around 95 million points and has an uneven density due to the acquisition set-ups. The segmentation allowed us to correctly detect calcareous stones as in Table 6, as well as openings regarding the surface of reference (best fit plane through convolutional bank filter) and the full limestone bay frames.

Table 22. Segmentation accuracy of the façade of the castle of Jehay over calcareous stones.

Elements	**Segmentation In Number of Points**		**Accuracy**
	Ground truth	Method	
	Calcareous Stones		
Sample NO. 1	37,057	35,668	96%
Sample NO. 2	30,610	27,100	88%

Sample NO. 3	34,087	32,200	99%
Sample NO. 4	35,197	30,459	86%

The same reasoning engine was used based on the DSAE ontology. The DSAE-based classification first studied the material Limestone (related to the property colour, same as S9 from CIDOC-CRM) and the geometry regularity (related to SP15 attribute from geometry) in regard to CS, CP and Area (in the case of 3D objects, the area was extended to a volume feature by taking into account every spatial dimension.), then differentiated openings through dimension-based predicates (SP15 Geometry) as presented in Figure 53. The CIDOC-CRM and its extension CIDOC-CRMba [240] provide an added descriptive value for archaeologists that can be directly plugged as in Figure 53.

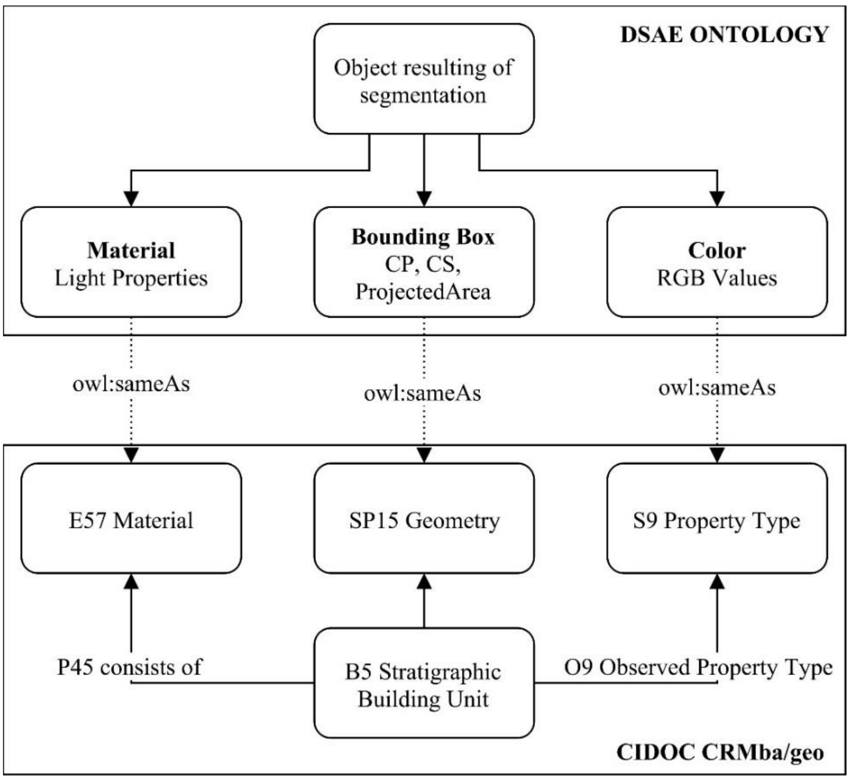


Figure 53. DSAE ontology and plugged CIDOC-CRM + CIDOC-CRMba for the detection of objects of interest: calcareous stones, openings and limestone bay frames.

The detected segments are classified, with 85% accuracy for independent calcareous stones, and 100% for woodworking openings (differentiated by size and geometric regularities) and limestone bay frames. The results over the Renaissance façade recognition pipeline are illustrated in Figure 54. We notice the fine detection for each element and the irregularity for some stones due to the uneven quality of the point cloud colorization. Calcareous stones classification was largely impacted by the segmentation inaccuracy within certain zones that led to over-segmentation and thus incorrect labels due to shape irregularity. These influential factors are discussed in Section 5.5.

Figure 54. Ontology-based classification of the South-South-West façade of the castle of Jehay. From left to right: the façade studied; the result of the segmentation (stones only); the result of the full segmentation; the result of the classification for quasi-planar objects of interest.

It is interesting to note that the DSAE ontology can be further used for distinguishing wide woodworking openings from smaller ones based on Area (or Volume) properties, and their geometric regularity. However, their label is considered weak for archaeological purposes, and extending the ontology as in Section 5.5 would provide a better automatic characterization for archaeological analysis.

The established data infrastructure gravitate around a client-server protocol that allows maximum flexibility and extensibility in regard to the 4 prerequisites of digital archaeology as defined in [173,174]. The platform can scale up to multiple simultaneous connexions and handles multi-source datasets. Every client that connects to the server as in Figure 55 benefits of functionalities from both the ontology reasoner and SQL statements (e.g., in Section 5.4).

The implementation was made using PostgreSQL DBMS enhanced with plugins (PostGIS and pgPointCloud). The software Protégé alongside the programming toolkit JENA (Java) was also used to create and link ontologies.

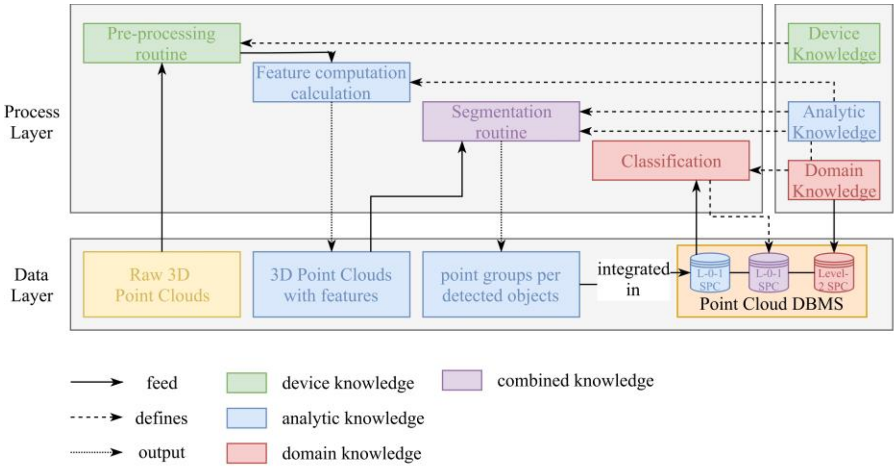


Figure 55. Server-side data management system. Point clouds go through different processing steps regarding Section 5.3, and point groups based on the definition of objects regarding domain knowledge are constituted and populate the SPC database.

As for the client-side, it was constructed to be as open and accessible as possible. As such, the World Wide Web is a democratized way to share and exchange information. It constitutes a long-term means to collaborate, and is independent of the location which is very important considering the need to be able on site to work with digital copies. Indeed, an application accessible anywhere and by multiple users at the same time is key for an archaeological 3D platform. Thus, we implemented the application in WebGL, a JavaScript API for rendering 3D graphics within any compatible web browser. We used Three.js, a cross-browser JavaScript library which uses the WebGL framework and enhances it. By a simple interaction with the GUI, the users can access and share a common viewpoint result of a semantic query. Figure 56 presents the optimal viewpoint for two classes of tesserae similarly to [177].

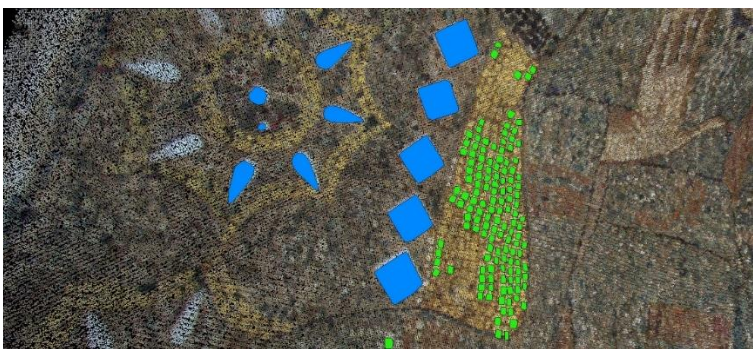


Figure 56. Query result in the WebGL prototype of the optimal viewpoint for faience pieces and gold tesserae in an extracted zone through SQL query.

The complete workflow therefore us allows to (1) pre-process multi-sensory point cloud data, (2) compute features of interest, segment and classify the point cloud according to domain knowledge formalized in ontologies; (3) structure the data in a server-side SPC point cloud 3D GIS; (4) disseminate the information through a client-side app built upon WebGL with a specific visual processing engine to provide optimal viewpoints from queries.

5.5 DISCUSSION

The democratization of TLS and dense-image matching in archaeological workflows makes them a preferred way to record spatial information. Point clouds are very interesting for their objectivity and flexibility in interpretation processes. If the acquisition is complete, they transcript every visual element that was observed on the field. However, other components that can arise to our other senses such as mechano-reception (touching, hearing) or chemo-reception (taste, smell) are not captured by these remote sensors. However, their integration and link to the point clouds can be important as they constitute another source for better comprehension of the observed subject. Sensors that can capture such information as objectively as possible would be another step toward a possible better acquisition automatization. Today, archaeologists rely mostly on field-work to extract necessary information from human senses, eventually with the use of other sensors to detect additional patterns (e.g., x-fluorescent characterization in Germigny-Des-Prés). Exploring combination of multisensory surveys with sensor-level data fusion provides a great opportunity for further research and to keep a record of a more complete context. Indeed, archaeological studies deal with more and more information including archaeological observations but also data coming from other sciences (e.g., geology, chemistry, physics, etc.) and all these must be organized and considered together for an optimal understanding of the site. To avoid loss of information, recording of the fact and interpretation must be integrated in the same process [183] and specific tools should be investigated.

Regarding spatial information, 3D point clouds constitute a very exhaustive source for further archaeological investigations. However, their lack of integration in workflows narrows the possibilities and interpretation work. We identified their main weakness to propose better handling and combinatory potential between different information sources: how to coherently aggregate semantics, spatial (and temporal information in a later stage). With respect to the number of observations (points), autonomous processing is very important. When dealing with thousands of archaeological objects of interest (composed of millions of points) in a scene (composed of billions of points), manually segmenting and classifying would be a very time consuming and an error prone process. In this paper, we presented an effective approach to automate tesserae recognition from terrestrial laser scanning data and dense image-matching. Knowledge-based feature constraints are defined to extract gold, silver, coloured glass and ceramic tesserae from a hybrid point cloud.

Then convex hull polygons are fitted to different segment separately. Knowledge is introduced again to generate assumptions for problematic parts. Finally, all polygons, both directly fitted and assumed, are combined to classify and inject semantics into the point cloud. Tests on three datasets showed automated classification procedures and promising results (Figure 57).

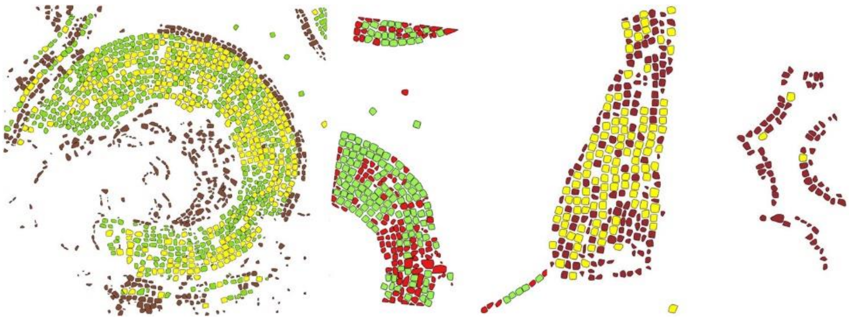


Figure 57. Classification results over the different zones of the mosaics.

The developed method tackles data quality challenges including heterogeneous density, surface roughness, curvature irregularities, and missing, erroneous data (due to reflective surfaces for example). We see that in zones where the colour quality is good and blur is low, classification results exceeds 95% accuracy. However, the method is very sensitive to 3D capture conditions and representativity such as colour, intensity, resolution and sharpness. Therefore, segmentation will fail when the input data does not allow correct feature extraction and abstraction-based connectivity estimation. More complete tesserae knowledge will help to better understand and detect complex shapes and patterns. While the classification results using domain knowledge are promising, the full point cloud labelling scheme could be enhanced by improving specifically the segmentation step. The data quality influences the final results. As illustrated, a challenge is brought about by varying densities and poor point-feature quality that can lead to over-segmentation when predominant features rely on point-proximity/density criterions. While this is not an issue for dense point clouds that describe continuous surfaces, it can constitute a hindrance for heterogeneous density or uneven datasets. Equally, colour/intensity that create imprecise colorization/featuring leads to rough classification. A solution would be to move the colour-based segmentation to the DSAE ontology to provide new discriminative possibilities. Also, the combination of dense image matching with laser data and 3D distance map improves the outline generation in a later stage, and allows a better shape estimation (Figure 58). Yet, an efficient registration is mandatory for accurate results. To improve the classification results, the segmentation can be enhanced using a watershed algorithm as well as obtaining higher representativity colour attribute for example. These are research directions that will be investigated. Also, to improve the robustness of the segmentation, a region-growing from a seed point located at every centroid of each detected connected element potentially provides a solution to under-segmentation, and investigations are necessary in this direction.

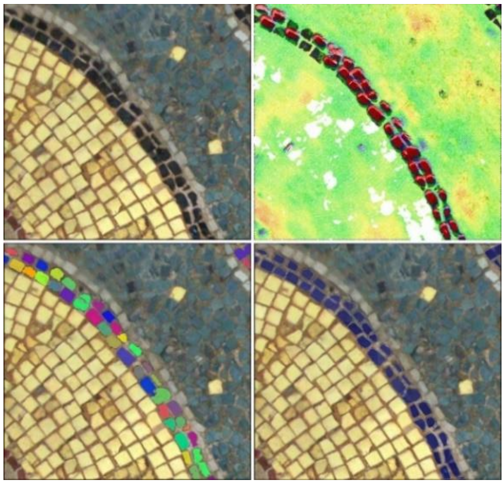


Figure 58. Classification and semantization of dark coloured glass.

The constituted ontology provides a reasoning engine based on available information that can be further enhanced to integrate new triple stores. As such, an acquisition campaign using a portable X-Ray Fluorescent device was carried out to quantify the relative quantity of chemical component within some tesserae. Integrating this semantic information could provide new reasoning capabilities such as detecting every gold tesserae that contain a quantity X of Plumb.

The method will also be refined and extended to the full point cloud by implementing a machine learning framework using obtained labelled data as training data. First results are encouraging using supervised classification [241], and other approaches such as reinforcement learning will be investigated for they high reasoning potential and complementarity to ontologies. However, the computer memory-demand of point clouds may impose a link to 2D projective raster's and to leverage existing training datasets (e.g., DeepNet).

The data structure relies on PostgreSQL RDBMS while indirectly integrating ontology reasoning results. It allows specific queries over the classified point cloud to extract spatial, semantic or a combination of both information. The blend of SPARQL and SQL allows us to combine efficiently the strength of both the relational database structuration and block-wise storage capabilities with the powerful reasoning proficiencies provided by ontologies. Different queries are therefore available, which are big leap forward regarding point cloud processing for archaeology (e.g., in Table 7).

Table 23. Example of queries over the point cloud.

Language	RDF Triple Store	Effect
SPARQL	PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#> PREFIX npt: <http://www.geo.ulg.ac.be/nyspoux/> SELECT ?ind WHERE { ?ind rdf:type npt:AltoMedievalTessera } ORDER BY ?ind	Return all alto-medieval tesserae (regarding initial data input)
SQL	SELECT name, area FROM worldObject WHERE ST_3Dintersects(geomWo::geometry, polygonZ::geometry);	Return all tesserae which are comprised in the region defined by a selection polygon and gives their area
SPARQL & SQL	SELECT geomWo FROM worldObject WHERE ST_3Dintersects(geomWo::geometry, polygon2Z::geometry) AND area > 0,0001; PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> PREFIX npt: <http://www.geo.ulg.ac.be/nyspoux/> SELECT ?ind WHERE { ?ind rdf:type npt: XIXCentTessera } ORDER BY ?ind	Return all renovated tesserae in the region 2 where the area is superior to 1 cm^2

However, while the temporal integration was inferred, only static intervals and fixed point in time were treated. Better integration such as continuous data or the storage and reasoning over datasets covering one location at different time intervals has yet to be further investigated. Indeed, new descriptors emerging from change detection could provide new insights and possibilities for cultural heritage conservation.

The proposed methodology (described in Section 5.3) was as general as possible to be extended to other use cases, at the object and local scales. It provides a potential solution for bringing intelligence to spatial data, specifically point clouds as seen in [108]. For example, we tested a point cloud from dense-image matching captured in Denmark (Ny-Calsberg Museum) and processed using Bentley ContextCapture (Figure 59). It constitutes an interesting object scale dataset where the interest lies in deciphering the hieroglyphs.

Figure 59. 3D point cloud of the statue of the Egyptian priest Ahmose and his mother, Baket-re. Diorite. C.1490–1400 BC. 18th Dynasty. New Empire. Ny Carlsberg Glyptotek Museum. Copenhagen. Denmark. From left to right: 3D point cloud; feature extraction and segmentation; 3D visualization.

The methodology was applied, and each hieroglyph was successfully detected independently. As the spatial context is conserved, we can locate the relative position of each hieroglyph regarding the others, and using a lexicon or a structured ancient hieroglyph ontology, each sign could be detected by shape matching (e.g., RANSAC), and a reading order extracted as in [242]. Thus, the methodology is suitable to reason from information extraction, and possibilities are very encouraging. Deepening the classification through well-established ontologies such as CIDOC-CRMba as illustrated in Section 5.4 is possible, and the extension to other use cases requires us to identify specific specializations and the level of detail within the tree depth. If we consider the Renaissance façade of the castle of Jehay, the CIDOC-CRM ontology as well as the CIDOC-CRMba and the CIDOC-CRMgeo add flexibility for moving deduction capabilities from the analytic part to the ontology. This is very interesting as it maximizes the DSAE reasoning capabilities instead of determining analytically discriminative features (such as bounding-box "is contained in" relationship from coordinates). As an example, the classification of the façade can be related to the specialization levels from B1 to B5 of the CIDOC-CRMba, and directly plugged as in Figure 53. Then, specifically looking at full limestone bay frames (same as B5 Stratigraphic Building Unit), an element that is contained within a limestone bay frame is classified as an empty section regarding Figure 54. The topological relations are introduced with the use of the well-known GeoSPARQL ontology to allow the detection of openings based on a AP12 "contains" relationship.

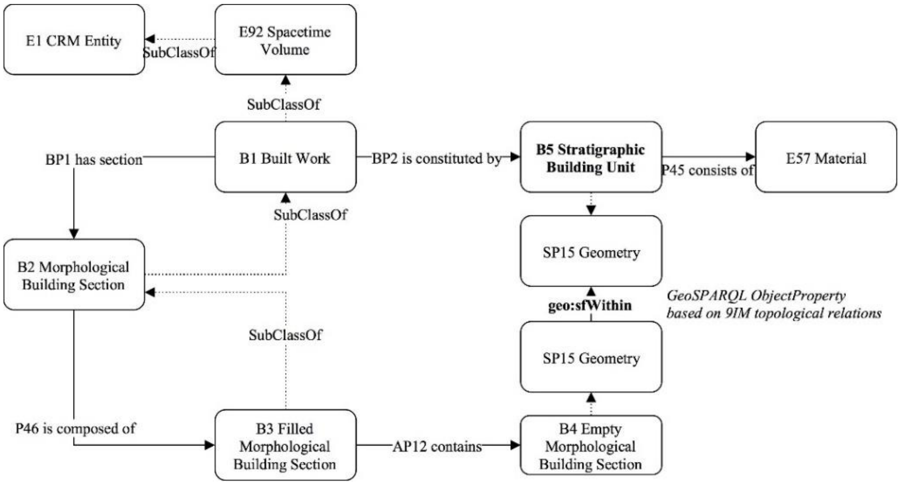


Figure 60. CIDOC-CRM ontology for the detection of objects of interest: calcareous stones and openings. Considering the castle of Jehay (B1), it has a building section (BP1) being the studied façade (B2), composed of different elements such as calcareous stones (B3), embrasures (B3) and openings (B4).

Therefore, by integrating attributes such as Color and ProjectedArea of the different elements (as well as topological "is Within" test), the ontology can be used for reasoning. Based on general axioms, it semantically recognises building parts of a façade as in Table 8.

Table 24. Classification of elements based on numerical attributes and topological relations.

Language	Equivalent To Definition	Effect
OWL (Protégé)	(hasProperty some ColorLimestone) and (hasProperty some NonReflective) and (CP some xsd:double[<= "4.0E-4"^^xsd:double]) and (CS some xsd:double[>= "1.05"^^xsd:double, <= "9.0"^^xsd:double]) and (ProjectedArea some xsd:double[>= "0.05"^^xsd:double, <= "0.4"^^xsd:double])	Defines an element as a BayFrame
OWL (Protégé)	(not (hasProperty some ColorLimestone)) and (sfWithin some BayFrame) and (BoundingBox some xsd:double[>= "2.9"^^xsd:double, <= "3.5"^^xsd:double])	Defines an element as a DoorSection

It is interesting to note that further reasoning is made possible due to extended knowledge over Renaissance-style mullioned windows. Indeed, double mullioned openings are a complex architectural element present over this façade. They are 2 mullioned openings where the separation by stones is inexistent. Each can be described regarding CIDOC CRMba as: 1 frame (B5) and 6 openings (B4 Empty morphological Building Section). Thus, an extended ontology can recognize double mullioned windows and a reason such as the one presented in [243] would provide extended automatization.

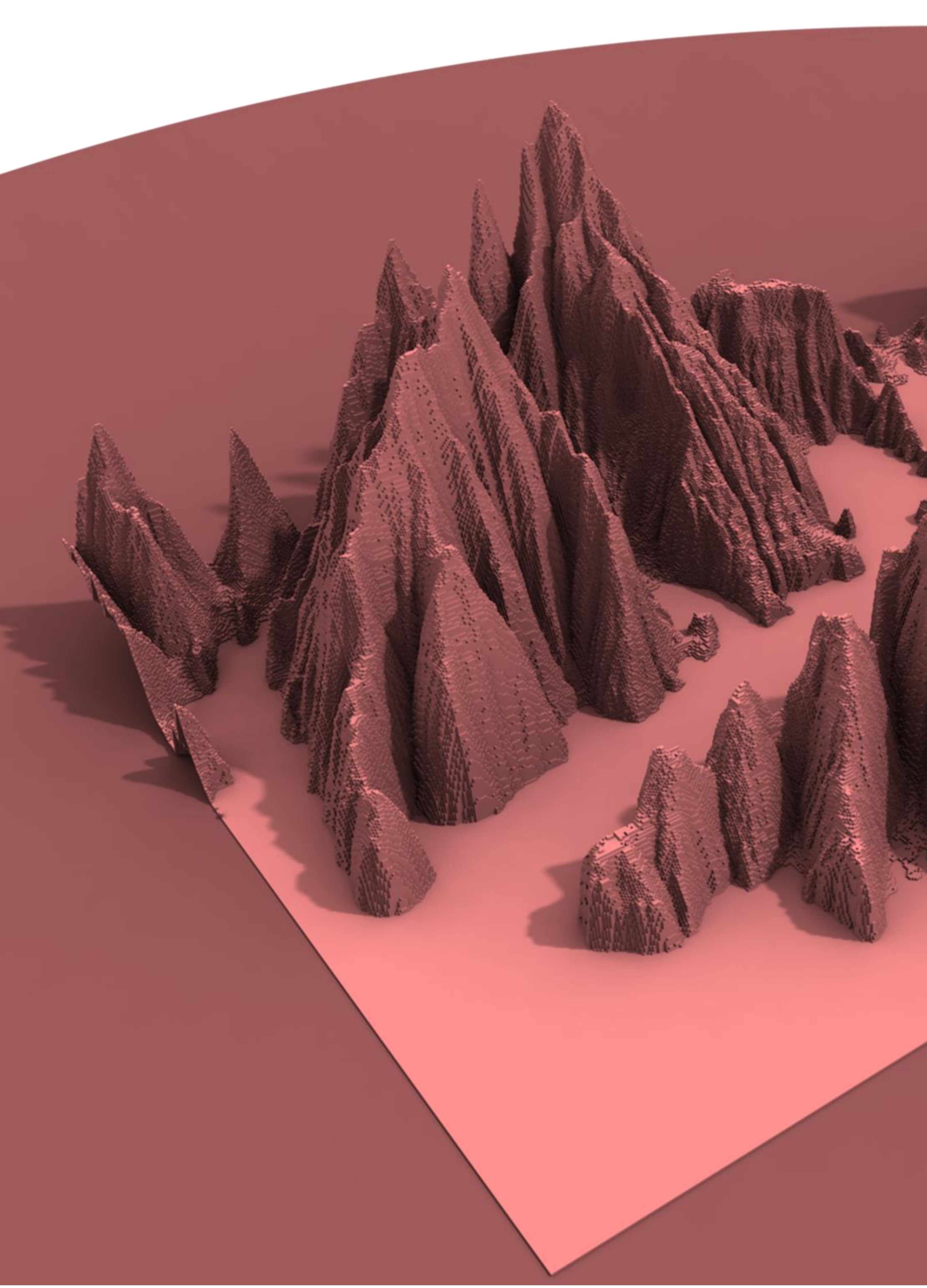
The final step for the visualization and presentation of the results is to share and distribute the information to other users and relies on virtual environments with specific interaction. The perception in 3D spaces is a dynamic phenomenon and concerns firstly behaviours and effects [244]. Data visualization is important to explore the data, to obtain some idea of what they contain, and therefore, to develop some intuitions about how to go about solving a problem from that data, determining what features are important and what kinds of data are involved. Visualization is also important when looking at the output of data science systems: data summarization for creating useful exploratory statistics, essential to understanding what was collected and observed. Although used before for tackling models and algorithms to avoid missing crucial information, data visualization is important for translating what might be interpretable only to a specialist for a general audience. In the context of point cloud, semantics and domain can highly influence the type of rendering used to directly transmit the correct information in a correct way to the end user. New ways of interacting with the data—Virtual reality, augmented reality, real time exploration and collaboration, holograms—are redefining possible interactions and exploration. [245] list different surface representations that can be used to represent and use a point cloud, including parametric modelling, implicit and simplicial representation, approximated and interpolated surfaces. The time-consuming task of accurate 3D surface reconstruction from point cloud requires many steps of pre-processing, topology determination, triangular mesh generation, post-processing and assessing. For example, [246] propose two simplification algorithms for LoD generation by decimating and simplifying meshes, thus reducing accuracy and quality. The development of an Internet browser-based solution allows maximum flexibility regarding theses identified problems, including data indexation vis-à-vis [61] to provide streaming capabilities independently of the size of the dataset.

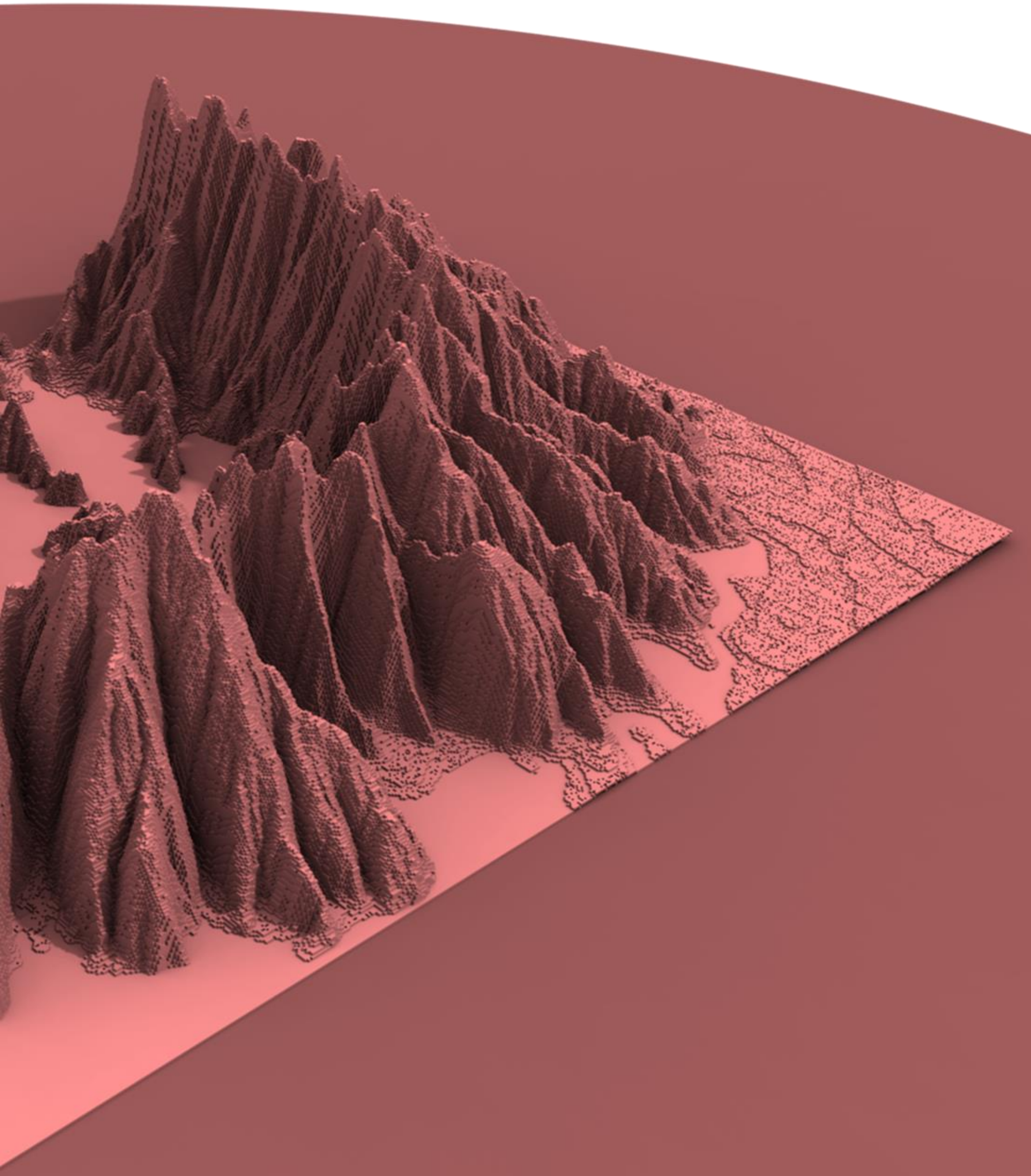
Based on the algorithm developed by [239], we manage the 3D viewpoint so as to determine an optimal position and orientation of the camera for the visualisation of three kinds of tesserae distinguished by their material: faience, gold and silver. Through the previous steps of recognition and semantization described, we are now able to exploit the semantically rich point cloud data structure [117] to visualise efficiently the different sorts of tesserae. To achieve this, we performed a pre-processing step, totally transparent to users, in which we compute the optimal camera positions on a 3D COLLADA model of the mosaic which is constituted of the minimum convex hull of each tesserae information stored in the database. This technical implementation will be enhanced to enable more direct integration of the geometry generalizations from the database. The algorithm looks at the pixels of the computational display which avoids the under-object recognition phenomenon. It also allows us to directly work on the final rendering of the 3D model which already integrates the use of an algorithm to process hidden faces. Finally, it can be used on any kind of 3D data structure (vector, raster or point cloud). It is worth mentioning that additional viewpoints could be computed which depends on the initial query. For instance, we can calculate multiple optimal camera positions for one specific sort of tesserae, depending on a needed surface, distance to rotation center, density estimate, etc. The latest could be particularly interesting for the golden tesserae since they are quite scattered in space. Furthermore, we can also investigate the impact of the statistical parameter used when two viewpoints present the same number of objects (maximum, average, etc.).

To integrate the semantically rich point cloud and the viewpoint management of queried tesserae, we developed web software using jQuery, Three.js, Potree (an Open Source JavaScript library for point cloud rendering) and tween.js. The platform includes a tool to directly allow semantic extraction and visualisation of pertinent information for the end users. It enables efficient information relay between actors. The web application is accessible on any HTML5-compatible browser. It enables real time point cloud exploration of the mosaics in the Oratory of Germigny-des-Prés, and emphasises the ease of use as well as performances. However, the integration of a natural language processor would allow us to extend the possibilities for users to formulate queries that are translated into SQL and SPARQL analogues.

5.6 CONCLUSIONS

In this paper, we first reviewed the state of the art in digital archaeology. We pointed out gaps in the integration of spatial information with semantic components and the limited management of 3D point clouds within 3D GIS. The recording and processing of 3D multi-source complex data were addressed, as well as their management, conservation, visualization and presentation for different users. In this paper, we propose a new solution to integrate archaeological knowledge within point cloud processing workflows. Specifically, we decompose point clouds regarding available features and estimated geometric properties that generate ontologies to classify and reason based on information extraction. We developed a data-driven ontology for point cloud analysis to facilitate interoperability to other formal ontologies such as the CIDOC-CRM, and applied the workflow over different point clouds. Quasi-planar objects (doors, windows, tesserae, calcareous stones, hieroglyphs) were successfully detected, and an HTML-5 cross-platform web application was created to facilitate the knowledge dissemination such as ancient mosaic located in the oratory of Germigny-des-Prés. Then, we extracted the necessary requested information from the semantically rich point cloud data to efficiently visualise user's request based on computed optimal camera positions and orientations that maximise the visibility of requested objects (e.g. tesserae). Then, the optimal viewpoints are dynamically rendered to users through the platform on which interactions can grow.





Point Cloud voxelisation of Raster Digital Elevation Model from Lidar data, Canada.

CHAPTER 6

Knowledge integration & representation

Chapter 6 - Knowledge integration & representation 165

6.1 Introduction 170
6.2 Related work 171
6.2.1 Methods for 3D point cloud geometric modelling 172
6.2.2 Instance-based object recognition and model fitting 175
6.2.3 Knowledge integration (KI) for object-relationship modelling 177
6.3 Materials and Methods 179
6.3.1 Knowledge-base structuration 180
6.3.2 Instance-based characterization, feature extraction and description refinement 185
6.3.3 Procedural instance 3D modelling 193
6.3.4 3D Aggregation for scene modelling 196
6.4 Results 198
6.4.1 Datasets 198
6.4.2 Comparisons 198
6.4.3 Computation time 204
6.4.4 Limitations 206
6.5 Perspectives 207
6.6 Conclusions 209

CHAPTER'S PREFACE

The previous chapters 4 and 5 dealt with automatic segmentation and domain-related object recognition. All the extracted knowledge is then structured within the SPCI along point data. In this chapter 6, we provide a way to leverage such a structure for the tasks of automated reasoning, specifically 3D part-modelling. As such, we target the language and query processor module as well as the reasoner module as illustrated in Figure 61.

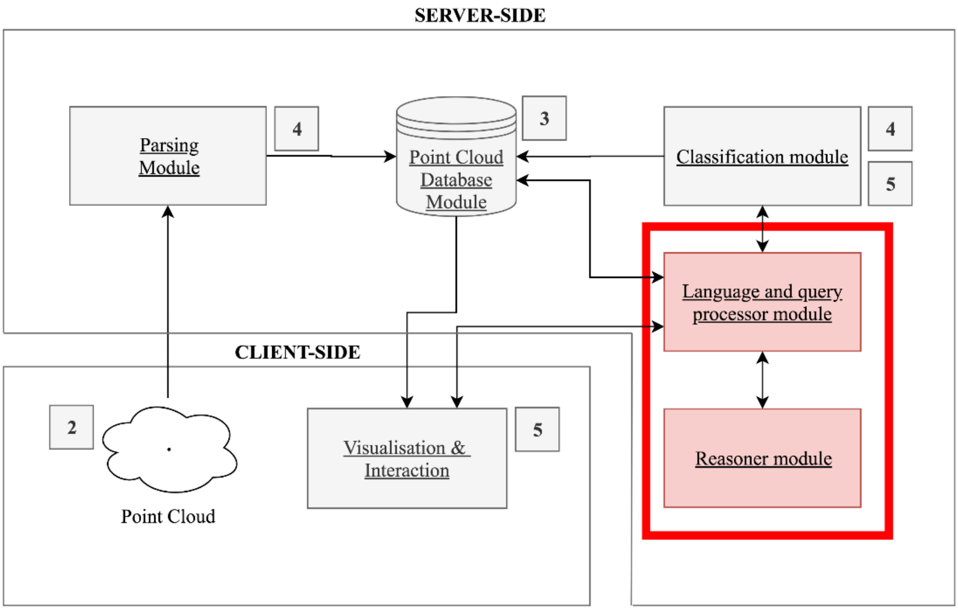


Figure 61. Chapter 6: The Smart Point Cloud reasoning possibilities.

In this chapter, we start with an indoor classified point cloud previously processed (details presented in Chapter 4), and we show a domain-related part segmentation to refine the object toward the application at hand. On top, the reasoner permits to automatically extract different 3D representations mode, which can be useful for interoperable workflow and direct connexion to domains such as BIM and asset management.

Based on Article [5]

3D Point Cloud Semantic Modelling: Integrated Framework for Indoor Spaces and Furniture

Abstract: 3D models derived from point clouds are useful in various shapes to optimize the trade-off between precision and geometric complexity. They are defined at different granularity levels according to each indoor situation. In this article, we present an integrated 3D semantic reconstruction framework that leverages segmented point cloud data and domain ontologies. Our approach follows a part-to-whole conception which models a point cloud in parametric elements usable per instance and aggregated to obtain a global 3D model. We first extract analytic features, object relationships and contextual information to permit better object characterization. Then, we propose a multi-representation modelling mechanism augmented by automatic recognition and fitting from the 3D library ModelNet10 to provide the best candidates for several 3D scans of furniture. Finally, we combine every element to obtain a consistent indoor hybrid 3D model. The method allows a wide range of applications from interior navigation to virtual stores.

Keywords: 3D modelling; 3D point cloud; ModelNet10; PCA; Point Cloud Database; cognition systems; feature extraction; procedural modelling; semantic segmentation; voxel

Published: 05 September 2018
Type: Open Access Article
Publisher: MDPI
Journal: Remote Sensing
Special Issue: 3D Modelling from Point Clouds: Algorithms and Methods

6.1 INTRODUCTION

3D point cloud geometric depiction is addressed through many 3D modelling techniques to best represent underlying shapes. This need is driven by applications in numerous industries[16] for tasks such as structural deformation scenarios [247,248], quality and progress control [249] or even for asset creation in the entertainment business [250,251] . To extend this range of applications, the data mining and processing research communities focus on adding additional information to the 3D model through semantic descriptors [252]. This in turn leads to more advanced uses of 3D virtual data of which many indoor scenario benefits. For instance, we employ 3D semantic models to plan/monitor emergency routes [253–257], for serious gaming [258,259], 3D waves propagation simulations [260–263], localization of safety-relevant features [113], virtual museums [264] or product lifecycle management [265]. More recently, the field of robotics demonstrated a great interest in these enhanced 3D geometries for their creation [65,266], scan planning [267] or for 3D autonomous indoor navigation [268–270] related to transportation and mobility problematics [271]. But bridging the gap between point cloud data, 3D models and semantic concepts is a very complicated task which usually requires a good knowledge of the specific applicative domain.

Our contribution is an attempt to narrow this gap by leveraging formalized knowledge and expert systems [272] based on point clouds. We want to take advantage of computer reasoning over semantic representations of our environment. However, this demands very challenging knowledge processing due to the heterogeneity of application domains and various 3D representations.

Our approach addresses Knowledge Extraction (KE), Knowledge Integration (KI) and Knowledge Representation (KR) [2] to better assimilate point cloud data with various quality [70]. Indeed, most software and tools existing in our computerized environment were developed to work primarily with 3D models. The landscape in standards, practices and usages is mostly established for these representations. This motivates a flexible and modular infrastructure of which point clouds can be the starting point [4,54,108,117] to allow interoperable and two ways exchanges (from and to the point cloud enrichment frameworks). If enhanced with additional information (geometry/topology/semantics), they could be used for deriving more representative 3D shapes and provide a higher compatibility with 3D modelling workflows. As such, we need new methods that can directly derive application-driven 3D semantic representations while conserving interoperability

[16] Main industries object of this study includes: Architecture, Construction, Engineering, and Facility Management (AEC/FM); Risk Assessment and Emergency planning; Simulations; Marketing; Entertainment; Robotics; Transportation and mobility

over centralized semantics. This hypothesis considers that point clouds are semantically-rich[17] and efficiently organized for various processing tasks.

While the present paper is based on the Smart Point Cloud (SPC) Infrastructure [108], it can be replicated over any segmented dataset that benefits from different sets of attribute information. The main idea is that based on a 3D point cloud describing an indoor environment, we can extract 3D models of each object instance regarding an application ontology and combine them to generate tailored 3D representations suited for specific indoor scenarios. To provide a multi-LoD[18] framework for different utilisations, we support the modelling process by a characterization mechanism that can deepen the geometric analysis of shapes. We study the structuration and reasoning aptitude of ontologies to pull contextual information enhancing the modelling fit. Thus, we explore the possibility to leverage formalized knowledge to recreate occluded area and infer non-existent geometry. Our approach is extended with a 3D shape-matching approach through data mining using the 3D library ModelNet10 [273]. The end goal of this contribution is to derive a comprehensive 3D model extended with semantic information. In this paper, we focus on indoor reconstructions and asset management.

In the first part (Section 6.2), we review significant related work studying 3D point cloud modelling approaches and several use cases on which they were successfully employed. Driven by this state of the art, we then present in Section 6.3 our designed 3D reconstruction framework following a part-to-whole outline. We finally show the results (Section 6.4) looking at precisions and performances for different datasets. From a critical analysis, we provide the main findings, the limitations and the perspectives (Section 6.5) that the approach brings to 3D point cloud modelling of building's interior.

6.2 RELATED WORK

In this section, we briefly review most recent methods for point cloud modelling. We organize the related works in (6.2.1) recent methods for geometric reconstruction from indoor point clouds; (6.2.2) instance-based recognition, featuring and model fitting; (6.2.3) knowledge integration for object relationship modelling.

---

[17] A point cloud is considered semantically-rich if it contains semantics linking group of points together such as segment or class information

[18] LoD stands for Level of Detail. In this paper, it is considered as a set of discrete 3D shape representations of a geometry with a narrowed precision and is not employed regarding an establish formalism.

6.2.1 Methods for 3D point cloud geometric modelling

3D reconstruction from indoor point cloud data gravitates around different approaches for automatic modeling with different granularities. The chosen method is often guided by the application needs in term of precision, resolution, complexity and completeness. For example, semantic model utilizations [274] include creating as-built models for the monitoring of construction processes whereas visually-appealing virtual models of historical sites enable immersive experiences. While the latter emphasize high quality visuals, semantic applications often rely on approximate reconstructions of the global scene which conveys the object arrangement. Both semantic and virtual indoor 3D models can be extended using precise metric information to provide key information for public buildings or to assist indoor navigation. Several work address these different characteristics through shape representation which has been extensively studied in the last century. Generally, we look for perceptually important shape features on either the shape boundary information or the boundary plus interior content as noted in [275].

We can primarily distinguish explicit versus implicit shape representations. Explicit representations translate the shape of an object (e.g. a triangle mesh), while implicit representations indirectly encode the shape using a set of features (histograms, normal, curvature …). Explicit representations are well suited for modeling 3D objects, whereas implicit representations are most often used for 3D object recognition and classification [276]. In most reviews, point cloud modelling approaches are categorized regarding the type of representation, the type of input data or the type of algorithms used. In this section we will study the algorithms depending on the context and the available information similarly to F. Remondino in [277].

3D reconstructions that make solely use of the spatial attributes within point cloud data are found in many works. Delaunay-based methods are quite common in this area, and we invite the reader to study [278] for a comprehensive survey of these methods. These approaches place rather strong requirements on the data and are impractical for scanned real-world scenes containing significant imperfections. Also, it is often necessary to optimize the polycount (total number of triangular polygons it takes to draw the model in 3D space) for memory efficiency. As such, quad meshing [279] can lighten the representation and smoothness. A practical example of Boundary-Representation (B-Rep) can be found in Valero et al. [280] for the reconstruction of walls. While these are interesting for the low input requirements, we investigate techniques more fitted toward dealing with challenging artifacts such as occlusion. Berger et al. [37,281] propose an exhaustive state of the art in surface reconstruction from point clouds. They reviewed thirty-two point cloud modelling methods by comparing their fit to noisy data, missing data, non-uniform sampling, outliers, but also their requirements in term of input features (normal, RGB data, scan data …) and shape class (CAD, indoor, primitives, architectural …). While surface smoothness approaches such as tangent planes, Poisson and Graph-Cut [37] can quickly produce a mesh, they often lack robustness to occlusion and incompletion. Sweeping models, primitive instancing, Whitney regular stratification or Morse decompositions [282] may be used for applications in robot motion planning and generally provide a higher tolerance for missing data.

By looking at the flexibility to generalize, the storage facility and applications scenarios, primitives are good candidates for indoor modelling. Indeed, as noticed by the authors in [283], parametric forms are *"mathematically complete, easily sampled, facilitate design, can be used to represent complex object geometries and can be used to generate realistic views"*. They describe a shape using a model with a small number of parameters (e.g. a cylinder may be represented by its radius, its axis, and the start and end points). They can also be represented non-parametrically or converted through the process of tessellation. This step is used in polygon-based rendering, where objects are broken down from abstract primitive representations to meshes. As noted by authors in [284], indoor environments are often composed of basic elements, such as walls, doors, windows, furniture (chairs, tables, desks, lamps, computers, cabinets) which come from a small number of prototypes and repeat many times. Such building components are generally formed of rigid parts whose geometries are locally simple[19]. An example is given in the work of Budroni and Boehm [285] where authors model walls by fitting CAD primitives. Furthermore, although variability and articulation are central (a door swings, a chair is moveable or its base rotates), such changeability is often limited and low dimensional. Thus, simple shapes are extensively used in the first steps of as-built modelling due to their compactness and the low number of parameters that allows efficient fitting methods [286]. For more complex shapes, explicit parametric representations are still available (e.g. Bézier curves, B-spline, NURBS) but they are mostly used as design tools. Since their control points cannot easily be inferred from point cloud data, these representations are rarely used in shape analysis. An example of parametric collection is given by Lee et al. [287]. They propose a skeleton-based 3D reconstruction of as-built pipelines from laserscan data. The approach allows a full automated generation of as-built pipelines composed of straight cylinders, elbows, and tee pipes directly fitted. While the method provides good results, its specificity restrains a possible generalization. Fayolle and Pasko [288] highlight the interaction potential given by parametrized objects within indoor environments. Indeed, through a clever binary (CSG[20]) or n-ary (FRep) construction tree structure, they store object-relations such that the user could modify individual parts impacting the entire logic of the object construction including its topology. Such a parametrized model reconstruction is required in many fields such as mechanical engineering or computer animation. Other relevant works highlight the papers of Fathi et al. [274]

[19] they consist of surfaces that are well approximated by planar, cylindrical, conical and spherical shapes.

[20] Constructive Solid Geometry

for civil infrastructure reconstruction or Adan and Huber [289] which provides a 3D reconstruction methodology (wall detection, occlusion labelling, opening detection, occlusion reconstruction) of interior wall surfaces which is robust to occlusion and clutter. Both results are a primitive-based assembly based on these surfaces.

While parametric assemblage gives a lot of flexibility, for some cases such as highly complex shapes, there is a need of low geometric modelling deviations. In such scenarios, non-parametric representations such as polygonal meshes are employed to better fit underlying data. However, the lack of compactness of these representations limits their use especially when dealing with large point clouds. Hence, using a combination of both representations is advisable when a global representation is required. In such approaches, parametric representations are usually used as local representations and decomposed into parts (e.g using CSG to represent each part with one or more geometric primitives). In contrast, triangle meshes are flexible enough to be used as global representations, since they can describe free-form objects in their entirety [276]. For example, Stamos et al. [290] present such a 3D modeling method on a church environment by combining planar segments and mesh elements. Using a different approach, Xiao and Furukawa [264] propose the "Inverse CSG" algorithm to produce compact and regularized 3D models. A building is sliced and for each slice different features (free space constraint, line extraction, iterative 2D CSG model reconstruction) are extracted, then stacked and textured to obtain a 3D model of walls. The method is interesting for its approach leveraging 2D features and its noise robustness, but it will not process complex structures, furniture or non-linear walls. Other hybrid approaches introduced knowledge within workflows to try to overcome main challenges, especially missing data. A significant work was brought by Lafarge et al. [291] which develops a hybrid modeling process where regular elements are represented by 3D primitives whereas irregular structures are described by mesh-based surfaces. These two different types of 3D representation interact through a non-convex energy minimization problem described in [292]. The approach successfully employed for large outdoor environments shows the benefits of leveraging semantics for better point cloud fitting. The authors in [293] present a reverse engineering workflow based on a hybrid modeling approach also leveraging knowledge. They propose a linear modelling approach through cross-section of the object by fitting splines to the data and then sweeping the cross-section along a trajectory to form the object model. To realize such operation, they first extracted architectural knowledge based on the analysis of architectural treaties to be used for guiding the modelling process. An example illustrated in [72] provide a method to reconstruct the boundary of buildings by extracting walls, doors, roofs and windows from façade and roofs point clouds. The authors use convex or concave polygons adjusted to different features separately. Interestingly, we note that the author use knowledge to generate assumptions for the occluded parts. Finally, all polygons are combined to generate a polyhedron model of a building. The approach is interesting for its whole-to-part consideration which leverage knowledge for optimizing the ratio approximation/compactness.

While triangulation and hybrid modelling are successfully used for various indoor scenarios, we notice that the most prominent module is the parametric modelling method. Its fit to both B-Rep and volumetric modelling accompanied by its high flexibility in representativity at different granularity levels will thus be further investigated in Section 6.3. Also, we notice that in some works, the use of knowledge permits to better describe shapes when used within the modelling workflow. We will thus investigate the literature for KI and KR in Section 6.2.3.

6.2.2 Instance-based object recognition and model fitting

Man-made objects populating indoor scenes often have low-degree of freedom and are arrangements of simple primitives. Beneath representation and compression efficiency [290–292], there is a real need to model independently different objects of interest that can in turn host different relationship information. The process of instance-based object recognition is required for identifying objects with a known shape, or objects that are repeated throughout a facility. The predominance of primitive forms in these environments gives specific shape descriptors a major control over the implicit representation. These can be categorized as geometric feature descriptors, and symmetric feature descriptors. In many works we find that planar detection plays a predominant role for the detection of elements in KE, specifically segmentation workflow.

Geometric feature descriptors: As such, predominant algorithms for geometric featuring in scientific literature are RANSAC [140,147–149,221,226,294–296], Sweeping [297], Hough [224,298–300] and PCA [121,152,226,301–304]. The authors [226,305] provide a robust PCA approach for plane fitting. The paper of Sanchez [295] makes primarily use of RANSAC to detect most building interiors, that may be modeled as a collection of planes representing ceilings, floors, walls and staircases. Mura et al. [306] partitions an input 3D model into an appropriate number of separate rooms by detecting wall candidates then study the layout of possibility by projecting in a 2D space the scenarios. Arbeiter et al. [75] present promising descriptors, namely Radius-Based Surface Descriptor (RSD), Principal Curvatures (PC) and Fast Point Feature Histograms (FPFH). They demonstrate how they can be used to classify primitive local surfaces such as cylinders, edges or corners in point clouds. More recently, Xu et al. [307] provide a 3D reconstruction method for scaffolds from a photogrammetric point cloud of construction sites using mostly point repartitions in specific reference frames. Funkhouser et al. [308] also propose a matching approach based on shape distributions for 3 models. They pre-process through random sampling to produce a continuous probability distribution later used as a signature for each 3D shape. The key contribution of this approach is that it provides a framework within which arbitrary and possibly degenerate 3D models can be transformed into functions with natural parameterizations. This allows simple function comparison methods to produce robust dissimilarity metrics and will be further investigated in this paper.

We find that using other sources of features both from analytical workflow as well as domain knowledge can contribute heavily into better segmentation workflows or for guiding 3D modelling processes.

Symmetry feature descriptors: Many shapes and geometrical models show symmetries: isometric transforms that leave the shape globally unchanged. If one wants to extract relationship graphs among primitives, symmetries can provide valuable shape description for part modelling. In the computer vision and computer graphics communities, symmetry has been identified as a reliable global knowledge source for 3D reconstruction. The review in [309] both at a global and local scale provide valuable insights on symmetry analysis. It highlights the ability of symmetries to extract features better describing furniture using a KR, specifically an ontology. In this paper, the extracted symmetric patches are treated as alphabets and combined with the transforms to construct an inverse-shape grammar [310]. The paper of Martinet et al. [311] provides an exhaustive review of accurate detection of symmetries in 3D Shapes. These are used for planar and rotational symmetries in Kovacs et al. [312] to define candidates symmetry planes for perfecting CAD models in reverse engineering. The paper is very interesting for its ability to leverage knowledge about the sape symmetries. Adan et Huber [289] list façade reconstruction methods, mainly based on symmetry study and planar patches reconstitution and then proposes a method that can handle clutter and occluded areas. While it can achieve a partial indoor reconstruction of walls and openings, it necessitates the position and each scan which is treated independently by ray-tracing. All these features play an important role for implicit geometric modelling, but also in shape matching methods.

Feature-based shape matching: For example, symmetry descriptors are used to query a database for shape retrieval in [313]. The authors in [314] propose a model reconstruction by fitting from a library of established 3D parametric blocks or Nan et al. [315] make use of both geometric and symmetric descriptors to best fit candidates from a 3D database, including a deformable template fitting step. However, these methods often include a pre-segmentation step and post-registration phase, which highly condition the results, and constrain the methodology to perfect shapes without outliers or missing data. Following this direction, F. Bosché [316] proposes a method using CAD model fitting for dimensional compliance control in construction. The approach is robust to noise, and includes compliance checks of the CAD projects with respect to established tolerances for validating the current state of construction. While the approach permits significant automation, it requires a coarse registration step to be performed by manually defining pair points in both datasets. To automate the registration of models with candidates, the Iterative Closest Point (ICP) method [317] is often used for fine registration, with invariant features in [318], based on least squares 3D surface and curve matching [319], non-linear least squares for primitive fitting [320] or using an energy minimization in graph [321]. The recent works of Xu et al. in [322] and [323] present a global framework where co-segmented shapes are deformed by scaling corresponding parts in the source and target models to fit a point cloud in a constrained manner using non-rigid ICP and deformation energy minimization. These works are foundations and provide research directions for recovering a set of locally fitted primitives with their mutual relations.

We have seen that knowledge extraction for indoor scenario is mostly driven by three categories of features namely geometric, symmetric and for shape matching. On top, as noticed by [284], object-relationships among the basic objects of these scenes satisfy strong priors (e.g. a chair stands on the floor, a monitor rests on the table) which motivates the inclusion of knowledge through KI for a better scene understanding and description.

6.2.3 Knowledge integration (KI) for object-relationship modelling

3D indoor environments demand to be enriched with semantics to be used in applications depicted in Section 6.1. It led to the creation of standards such as LADM [324], IndoorGML [325] or IFC (Industry Foundation Class) [326], that were motivated by utilizations in the AEC industry, navigation systems or land administration. Indeed, the different models can deal with semantically annotated 3D spaces and can operate with abstract spaces, subdivision views and have a notion of geometry, topology maintaining relationship between objects. The choice toward one model or another is mainly guided by the usage and its integration within one community. Therefore, semantically rich 3D models provide a great way to extend the field of application and stresses new ways to extract knowledge *a priori* for a fully autonomous Cognitive Decision System (CDS). The CDS can in turn opens new solutions for industries listed in the Global Industry Classification Standard [327].

In their work, Tang et al. [276] separate this process in geometric modelling, object recognition, and object relationship modelling. Whereas a CAD model would represent a wall as a set of independent planar surfaces, a BIM model would represent the wall as a single, volumetric object with multiple surfaces, as well as adjacency relationships between the wall and other entities in the model, the identification of the object as a wall, and other relevant properties (material characteristics, cost ...). These includes topological relationships between components, and between components and spaces. Connectivity relationships indicate which objects are connected to one another and where they are connected. Additionally, containment relationships are used to describe the locations of components that are embedded within one another (e.g. a window embedded within a wall).

Ochmann et al. [224] present an automatic reconstruction of parametric walls and opening from indoor point clouds. Their approach reconstructs walls as entities with constraints to other entities, retaining wall relationships and deprecating the modification of one element onto the other. The authors of [81,267,328,329] extend the processes of parametric modelling of walls and openings for BIM modelling applications. More recently, the paper of Macher et al. [298] presents a semi-automatic approach for the 3D reconstruction of walls, slabs and openings from a point cloud of a multi-storey building, and provides a proof of concept of OBJ to IFC manual creation. While these approaches contributed to new possibilities for walls and slabs semantic modelling, the object-relationship is limited to topological relationships.

The work of Fisher in [330] introduces domain knowledge of standard shapes and relationships into reverse engineering problems. They rightfully state that there are many constraints on feature relationships in manufactured objects and buildings which are investigated in this paper. Indeed, for a general workflow, one must provide a recovery process even when data is very noisy, sparse or incomplete through a general shape knowledge. Complete data acquisition (impossible in practice for some situations) through inference of occluded data permits the discovery of shape and position parameters that satisfy the knowledge-derived constraints. Formalizing knowledge would therefore be useful to apply known relationships when fitting point cloud data and get better shape parameter estimates. It can also be used to infer data about unseen features, which orients our work to consider ontologies. In this area, the work of Dietenbeck et al. [252] makes use of multi-layer ontologies for integrating domain knowledge in the process of 3D shape segmentation and annotation. While they provide only example and a manual approach over meshes, they describe an expert knowledge system for furniture at three conceptual layers, directly compatible with the three meta-models of the Smart Point Cloud Infrastructure introduced in [108] and extended in [2]. The first layer corresponds to the basic properties of any object, such as shapes and structures whereas the upper layers are specific to each application domain and describe the functionalities and possible configurations of the objects in this domain. By using domain knowledge, the authors perform searches amongst a set of possible objects while being able to suggest segmentation and annotation corrections to the user if an impossible configuration is reached. This work will be further investigated within our workflow. Using a different approach, Son and Kim [151] present a semantic as-built 3D modeling pipeline to reconstruct structural elements of buildings based on local concavity and convexity. They provide different types of functional semantics and shapes with an interesting parameters calculation approach based on analytic feature and domain knowledge. These works are fundamental and greatly illustrate the added benefit of leveraging knowledge.

KI and KR are an important part of any intelligent system. In this review we noticed that the use of ontologies provided an interesting addition to knowledge formalization and permits to better define object-relationships. Coupled with a KE approach treating geometric, symmetric and shape matching features, they could provide a solid foundation for procedural modelling based on a 3D point cloud.

Therefore, we develop a method (Section 6.3) inspired by these pertinent related works.

6.3 MATERIALS AND METHODS

In this section, we present a global framework for modelling pre-segmented/classified indoor point cloud data. The approach is divided into four steps A, B, C, D (as illustrated in Figure 62) respectively described in the four sub-sections 6.3.1, 6.3.2, 6.3.3 and 6.3.4. The methodology follows a part-to-whole design where each instance is treated separately before aggregation to reconstruct a semantically rich global 3D model.

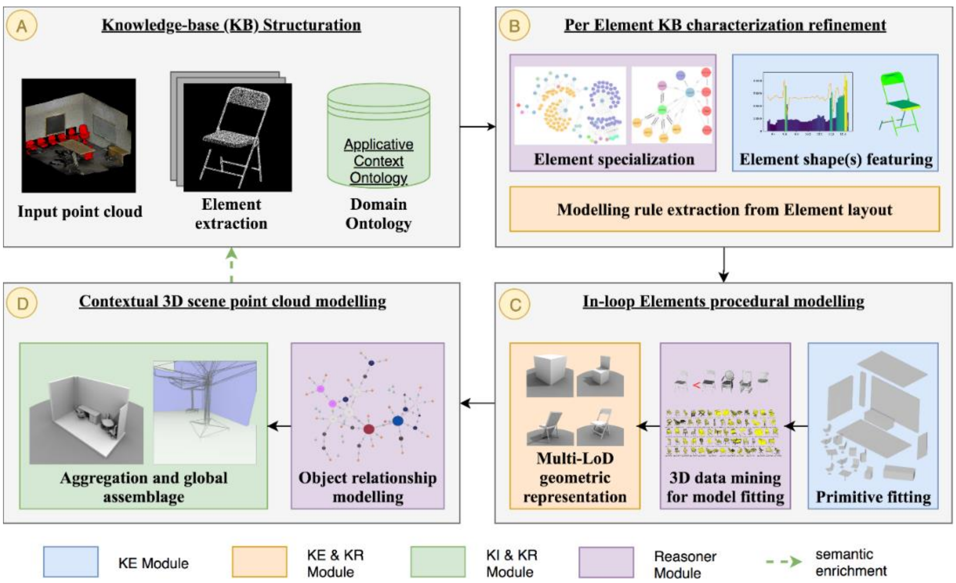


Figure 62. Global workflow for modelling indoor point cloud data. Our approach takes as an input semantically-rich point cloud[21], and process it regarding knowledge-based processes.

In the next sub-section (6.3.1), we describe how semantics are integrated within point cloud data and we provide details on the design of a multi-LoD ontology and its interactions (Figure 62: A).

---

[21] The data is extracted from the Smart Point Cloud Infrastructure which host several semantic information as well as a specific data structuration which makes point cloud handling and enrichment easier.

6.3.1 Knowledge-base structuration

Indoor 3D point clouds that host semantic information such as segments and classes are the starting point of our methodology to generate semantic models, which give insight to the morphology and geometry of building's interiors. For this purpose, we leverage the flexibility given by the Smart Point Cloud (SPC) Infrastructure [2] to consider point cloud data at different granularity levels (expressed in the Tower of Knowledge[22] concept [331]). The SPC conceptual model [108] permits to integrate annotated 3D point clouds, can operate with abstract space definitions and can handle multiple geometric representation. Its structure handles datasets at three levels as illustrated in Figure 63.

Figure 63. In red the handled geometry: **(a)** point level; **(b)** patch level; **(c)** object level.

At the point level (lowest level) the geometry is sparse and defines the lowest possible geometric description. While this is convenient for point-based rendering [332] which can be enhanced through deep learning [333] leading to simpler and more efficient algorithms [334], geometric clustering covers applications identified in Section 6.1. At the patch level, subsets of points are grouped to form small spatial conglomerates. These are better handled in a Point Cloud Database Management System (PC-DBMS) using a block-scheme approach which gives additional hints on the spatial context. At the object level, patches are grouped together to answer the underlying segmentation or classification approach. These three geometric levels are management within the PC-DBMS module (Figure 64) and directly integrate semantics (segment, class, function …) and space (abstract, geometric …) information. While 3D modelling approaches solely based on spatial attributes can leverage both the SPC point and patch levels, using the additional information linked to the object level extend the range of shape representations, thus applications.

Moreover, the SPC provides enough elasticity to centralize knowledge for decision-making scenarios. Its conception allows a mapping between domain specializations such as formalized IFC-inspired ontologies to provide additional reasoning possibilities. As a first step toward 3D point cloud modelling and to permit knowledge-based reasoning for indoor applications, we tailored the SPC Server-Side

[22] The Smart Point Cloud permits to reason solely spatially, semantically, but also includes functionality description and descriptors characterization, following the 4 levels of the Tower of Knowledge.

Infrastructure as in Figure 64 with an applicative context ontology (ACO) for efficient KE, KI and KR.

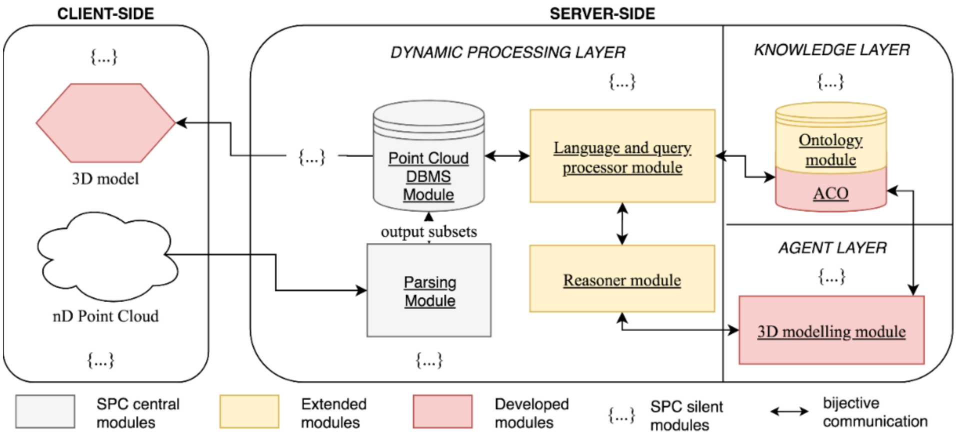


Figure 64. The Smart Point Cloud [2] tailored Infrastructure for 3D modelling.

In this infrastructure, we added a two-way mapping expert 3D modelling module (agent layer) to link 3D geometries to point data with semantic enrichment. The expert system consumes the semantically-rich point cloud data as well as the ACO and will be further described in Sections 6.3.2, 6.3.3 and 6.3.4.

To achieve semantic injection, we construct a multi-LoD IFC-inspired ontology (see Supplementary Materials). It provides knowledge about the object shapes, knowledge about the identities of objects and knowledge about the relationships between objects. Integrated in the knowledge layer (Figure 64) of the SPC, this ACO is used to best describe the morphological features of indoor elements in an interoperable manner. This new knowledge-base is mainly used for refining the definition of objects within the SPC (Section 6.3.2), inferring modelling rules (Section 6.3.2) and providing clear guidelines for object-relationship modelling through the reasoner module (Section 6.3.4). For example, if the considered point cloud dataset benefits of additional information handled by the SPC such as the "space" definition, the reasoner will permit higher characterization (e.g. if the object is within an "office room", high chances that the chair is a "desk chair" with rolls). Moreover, the local topology available through the SPC provides additional information that is crucial regarding CSG-based modelling or for our object-relationship modelling approach (e.g. if a chair is topologically marked touching the floor, occlusion at its base can be treated accordingly).

In a will to pool our ontology in the Web of Linked Data, each object concept, if it exists, is defined as an extension of the DBpedia knowledge base. Providing such a link is an important interoperable feature when it comes to an object that is referred as a DBpedia resource (e.g. for a chair object, http://dbpedia.org/page/Chair.). The classes hierarchy as illustrated in Figure 65 is split in three main conceptual levels.

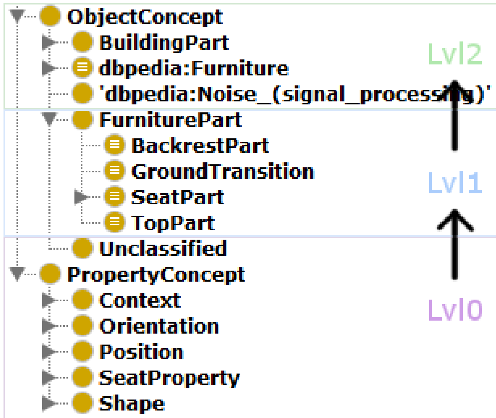


Figure 65. Classes hierarchy in the ACO

The level 0 is the first level of classified elements and defines properties established on point cloud features (shape, orientation ...). These features are stored in the knowledge base as datatype properties. Datatypes properties are relations pointing from entities to literals (characters string, numbers ...). The values describe different bounding boxes characteristics: (length, width ...).

The level 1 englobes Sub-Elements (SE) that are part of an Aggregated-Element (AE), defined in level 2. All entities that are not described within this hierarchy are categorized as "Unclassified". Modelling rules of higher definition levels restrict the lower levels inference through hierarchical constraints, similarly to [252]. These are strictly based on Sub-Elements, except for the topology relations between building parts and/or furniture.

This consistency check permits the construction of complex definitions but maintains a modularity for interoperability purposes. Semantic definitions of pre-labelled objects are extracted from the ACO ontology such as depicted in Figure 66. Input variables are nothing more than labels of the retrieved object capsuled into a JSON object part of the language processing module (Figure 64).

Labels are then used to set SPARQL queries on a graph structure to query the ontology.

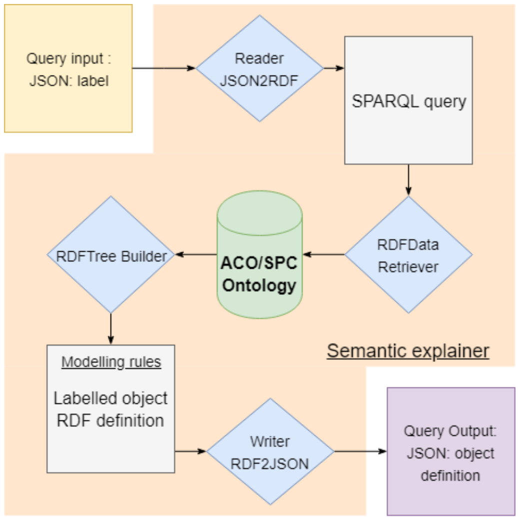


Figure 66. Framework for the extraction of an object definition from ACO to be readable by other modules of the SPC Infrastructure

The ontology-based extraction of object's definition is the reverse-process of classification. It is thus possible to extract mandatory elements of objects for reconstruction and to guide their modelling. As each level is independent, reconstruction can be made at different LoD considerations.

Graph-mining is guided on relation types between elements of each level. E.g. a "chair" is composed of a chair back (BackrestPart) located somewhere above[23] a seat (oneSeat), and some ground transition parts (GroundTransitionPart) under it.

Ontologies as XML-written files are well suited for hierarchical structure. A dedicated parser finally allows structuring the answer in a JSON-file and communicating with the object characterization step (Section 6.3.2).

Listing 1 shows the results of querying ACO/SPC Ontology about the "KitchenChair" label. It is structured as a hierarchical tree of characteristics where each level is detailed by its sub-levels. In this tree, each line specifies a triple that refers to the object description (e.g. "hasNormale some PerpendicularOrientation" specify that the normal need to be perpendicular to the main orientation for this specific Sub-Element).

---

[23] This information is extracted from gravity-based topology analysis described in Section 6.3.2

Listing 1 Kitchen Chair semantic definition extracted from the ACO

Label queried: "KitchenChair"

1. Detected element : *http://www.geo.ulg.ac.be/nyspoux/SPC#KitchenChair*
2. hasPart min 4 http://www.geo.ulg.ac.be/nyspoux/SPC#GroundTransition
3. hasNormale some PerpendicularOrientation
4. hasIntersectionPerpendicular some http://www.geo.ulg.ac.be/nyspoux/SPC#FurniturePart
5. hasAverageHistogram maxExclusive 0.25
6. hasContext some http://www.geo.ulg.ac.be/nyspoux/SPC#ContextKitchen
7. hasPart some http://www.geo.ulg.ac.be/nyspoux/SPC#OneSeat
8. hasWidth maxInclusive 0.45
9. hasThickness maxInclusive 0.45
10. hasLength maxInclusive 0.45
11. hasNormale some MainOrientation
12. hasPart some http://www.geo.ulg.ac.be/nyspoux/SPC#BackrestPart
13. hasNormale some PerpendicularOrientation
14. hasIntersectionPerpendicular some http://www.geo.ulg.ac.be/nyspoux/SPC#SeatPart
15. hasNormale some MainOrientation
16. hasAverageHistogram minInclusive 0.75
17. hasFlatness maxExclusive 0.5
18. hasElongation minInclusive 0.5

The reasoning module depends on the language processing module (Figure 64). As the OWL formalism is a Description-Logic-based language, it allows logical consequence inferences from a set of stated axioms.

We describe asserted facts in both the terminological box (TBox) and assertional box (ABox). TBox is constituted of classes definitions in the ACO ontology whereas individuals are populating the ABox. Because of editing rules and restrictions on class definitions, those boxes are inferred and constitute our structure for knowledge discovery on logical reasoning.

Listing 2 provides a simple example of inference based on both TBox and ABox. While TBox rules define the conditions for classifying an object as a Wall, ABox specify that three objects exist. The first object is constituted of the two others as a collection and these two subparts are defined as WallSurfaces.

Listing 2 Simple example of inferences on TBox and ABox

TBox rules and restrictions:

1. :Wall owl:equivalentTo (hasPart min 2 :WallSurface)

ABox population:

1. :SubElement1 a :WallSurface
2. :SubElement2 a :WallSurface
3. :Object hasPart :Collection(:SubElement1 and :SubElement2)

Inferences:

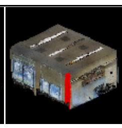
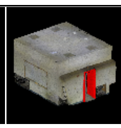
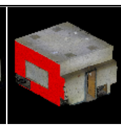
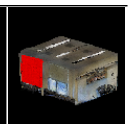
1. :Object a :Wall

Our approach was though for indoor point clouds, thus the ACO retain information for the following elements: beams, ceilings, floors, chairs, columns, doors, tables, walls and windows. These were primarily chosen due to existing workflows providing robust recognition. The elements are considered correctly categorized but it is not necessary to filter point cloud artefacts. In the next sub-section (6.3.2) we present the second step (Figure 62: B) of our workflow which aim at generating modelling rules holding each element specificity.

6.3.2 Instance-based characterization, feature extraction and description refinement

While the SPC-integrated point cloud holds a minima class or segment information, these are not necessarily optimal regarding indoor applications. The ACO permits to deepen the classification of considered classes presented in Table 25 through the Web of Linked Data. For each class, every element is extracted independently from the point cloud and considered for instance characterization.

Table 25. Objects to be modelled from the S3-DIS point cloud dataset [133] highlighted in red, identified by a major primitive (Cuboid, Plane, or CSG model assembly) and a category being Normal-Element (NE), Sub-Element (SE) or Aggregated-Element (AE)

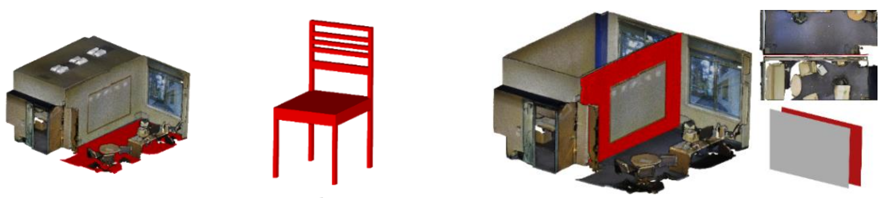
Beam	CeilingS	Chair	Column	Door	Table	WallS	Window
							
Cuboid	Plane	CSG model	Cuboid	CSG model	CSG model	Plane	Cuboid
NE	SE	AE	NE	AE	AE	SE	NE

Let p_i be a point in $\mathbb{R}^s$, with s the number of dimensions. We have a point cloud $\mathcal{P} = \{p_i\}_{i=1}^{n}$ with n the number of points in the point cloud. Let $\mathcal{E}_i$ be an element of $\mathcal{P}$ identified by a label $\mathcal{L}_i$, containing m points from $\mathcal{P}$. Let $\mathscr{g}$ be a directed graph defined by a set $\mathscr{v}(\mathscr{g})$ of inner nodes, a set $e(\mathscr{g})$ of edges and a set $\mathscr{v}_e(\mathscr{g})$ of end nodes. Each edge is oriented regarding a specified topology relation between one or multiple nodes (its ends). Then, three processing cases arise from the definition of elements (Figure 65) where:

(a) $\mathcal{E}_i$ is an element that is described in $\mathscr{g}$ as an end $v_e(\mathscr{g})$ (the final level of the elements decomposition, e.g a wall, a beam…). These elements are directly identified as Normal-Elements (NE) in the ACO Level 2 (Table 25) and refer to a DBpedia resource. Such a case does not necessitate a part segmentation, and therefore allow the expert system to directly address the modelling phase (Section 6.3.3).

(b) When $\mathcal{E}_i$ is an element described in $\mathscr{g}$ as a combination of multiple Sub-Elements (therefore is in the category of Aggregated-Element from the SPC), it goes through part-segmentation (Section 6.3.2) before modelling the entity. This segmentation is guided by inferred rules extracted from the ACO ontology. For instance, if an object is labelled as a "kitchen chair", its definition will specify that segmentation need to find an upper-part, which is a backrest, a middle-part, which is a oneSeat, and some ground transition parts, at least 3 for the example of the "kitchen chair".

(c) When $\mathcal{E}_i$ is a Sub-Element as in the ACO and refer to an Aggregated-Element, the point cloud subset goes through an aggregation step before modelling. This step is guided by semantic definitions and symmetric operations to find (recreate) other Sub-Elements of the Aggregated-Element. For instance, when a WallSurface is considered as $\mathcal{E}_i$, its parallel wall surface will be search for in the SPC database, to constitute the Aggregated-Element "Wall".



(a) (b) (c)

Figure 67. (a) $\mathcal{E}_i$ is a Normal-Element; **(b)** $\mathcal{E}_i$ is an Aggregated-Element, it goes through part-segmentation; **(c)** $\mathcal{E}_i$ is a Sub-Element.

The main processes of this second step are described in **Algorithm 1** part of the global workflow.

Algorithm 1 Element characterization, featuring and generation of modelling rules (Figure 62: B)

Require: A point cloud $\mathcal{P} \in \mathbb{R}^s$, decomposed in n $\mathcal{E}_i$ each with a label $\mathcal{L}_i$

1. **For each** $\mathcal{E}_i$ **do**
2. $C_{\mathcal{E}_i} \leftarrow$ characterization of $\mathcal{E}_i$ (NE, SE or AE)
3. **if** $C_{\mathcal{E}_i} == 'AE'$ **then**

4.   specialization of $\mathcal{E}_i$ through part-segmentation
5.   **end if**
6. ${\mathcal{E}_i}^+ \leftarrow$ Add Bag of features, Object relationship information and Contextual semantics
7. $\mathcal{R} \leftarrow$ Generation of modelling rules through ACO
8. **end for**
9. **end**
10. **return** $({\mathcal{E}_i}^+, \mathcal{R})$

While NE and SE's characterization avoid part-segmentation, AE goes through the case **(b)** for a higher representativity. Looking at considered AE classes (chair, door, table), the "chair" class provides the highest variability for testing and will thus be used as the main illustration of AE specialization and shape featuring (Figure 62: B). We decompose this mechanism (**Algorithm 1**: line 4) in three sub-steps.

Sub-Step 1. Pose determination of 3D shapes: we use a robust variant of Principal Component Analysis (PCA) inspired by Liu and Ramani [152], to compute the principal axis of points composing $\mathcal{E}_i$ (**Algorithm 1.1**). The eigen vector with the largest value in the covariance matrix is chosen as the first estimate of the principal direction.

Algorithm 1.1 Robust Principal Axis Determination (RPAD)

Require: A point cloud object $\mathcal{E}_i \in \mathbb{R}^3$ filtered for considering only spatial attributes along $\vec{x}, \vec{y}, \vec{z}$, max_$iteration$ the maximum number of iteration (by default: 1000),

1. $\wp \leftarrow$ inliers candidates from Statistical Outlier Identification Filter [335] applied to $\mathcal{E}_i$
2. **while** $j <$ max_$iteration$ **do**
3. $\mathscr{o}, \overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3} \leftarrow$ origin and three principal axes of $\wp$ through PCA
4. $r \leftarrow \mathcal{E}_i - \wp$, remaining points as a matrix in $\mathbb{R}^3$, r_i being a point $p_i \in r$
5. $res(r_i) \leftarrow$ the distance between r_i and its projection onto $\overrightarrow{e_1}$
6.   **if** $|res(r_i)| < |mean(res) + \sigma|$ **then**
7. $\wp.append(r_i)$
8.   **end if**
9. **end while**
10. $\overrightarrow{e_2}, \overrightarrow{e_3} \leftarrow$ Update by a 2D RPAD over $\mathcal{E}_i$ projection onto the plane defined by $\overrightarrow{e_2}, \overrightarrow{e_3}$
11. **end**
12. **return** $(\mathscr{o}, \overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3})$

We provide an additional refinement layer leveraging georeferenced datasets and gravity-based scenes by constraining the orientation of Sub-Elements (**Algorithm 1.2**):

Algorithm 1.2 Gravity-based constraints for Sub-Elements $\mathcal{E}_{sub-i}$

Require: A point cloud Sub-Element $\mathcal{E}_{sub-i} \in \mathbb{R}^3$ and $\mathcal{o}, \overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}$, the output of **Algorithm 1.1**

1. $\mathcal{o}_s, \overrightarrow{s_1}, \overrightarrow{s_2}, \overrightarrow{s_3} \leftarrow$ origin and three principal axes of $\mathcal{E}_{sub-i}$ through RPAD
2. $\overrightarrow{p_{s1}} \leftarrow$ projection of the main axis $\overrightarrow{s_1}$ onto the plane defined by the normal $\overrightarrow{e_3}$ from $mainElement(\mathcal{E}_i)$ as defined in ACO
3. $\alpha_1, \alpha_2 \leftarrow$ angle respectively $(\widehat{\overrightarrow{e_1}, \overrightarrow{p_{s1}}})$ and $(\widehat{\overrightarrow{e_2}, \overrightarrow{p_{s1}}})$
4. **if** $\alpha_1 \leq \alpha_2$ **then**
5. $\quad \overrightarrow{s_2} \leftarrow R_z(\alpha_1)\overrightarrow{s_2}$ and $\overrightarrow{s_3} = \overrightarrow{s_1} \times \overrightarrow{s_2}$
6. **elif**: $\overrightarrow{s_2} \leftarrow R_z(\alpha_2)\overrightarrow{s_2}$
7. **end**
8. **return** $(\mathcal{o}_s, \overrightarrow{s_1}, \overrightarrow{s_2}, \overrightarrow{s_3})$

Sub-Step 2. The 2nd sub-step in the part-segmentation process is to extract several shape features which guide the process. Every point p_i composing $\mathcal{E}_i$ is processed following **Algorithm 1.3**:

Algorithm 1.3 Histogram and bin featuring of an element $\mathcal{E}_i$

Require: A point cloud object $\mathcal{E}_i \in \mathbb{R}^3$ filtered for considering only spatial attributes X, Y, Z along $(\vec{x}, \vec{y}, \vec{z})$ axis, X_e, Y_e, Z_e the spatial attributes along principal directions $(\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3})$

1. $\mathcal{B} \leftarrow$ bin gridded elementary subset of $\mathcal{E}_i$, by default an octree-based voxel [4]
2. $done \leftarrow \emptyset$
3. **for** each $\mathcal{B}$ **do**
4. $\quad$ **if** $\mathcal{B}.index$ **is not in** $done$ **then**
5. $\quad\quad done \leftarrow done.append(\mathcal{B}.index)$
6. $\quad\quad \mathcal{C}_{\mathcal{B}} \leftarrow$ Coordinates $(X_{\mathcal{B}c}, Y_{\mathcal{B}c}, Z_{\mathcal{B}c})$ of the center of $\mathcal{B}$ as a vector in $\mathbb{R}^3$
7. $\quad\quad \mathcal{F}_{(\overrightarrow{e_2}, \overrightarrow{e_3})} \leftarrow$ count number of bin $\mathcal{B}$ with same $(Y_{\mathcal{B}c}, Z_{\mathcal{B}c})$ along $\overrightarrow{e_2}, \overrightarrow{e_3}$, and
   different $X_{\mathcal{B}c}$ along $\overrightarrow{e_1}$
8. $\quad\quad \mathcal{F}_{\overrightarrow{e_1}} \leftarrow$ count number of bin $\mathcal{B}$ with same $X_{\mathcal{B}c}$ along $\overrightarrow{e_1}$, and different $(Y_{\mathcal{B}c}, Z_{\mathcal{B}c})$ along $\overrightarrow{e_2}, \overrightarrow{e_3}$
9. $\quad$ **end if**
10. **end for**

---

11. **end**
12. **return** $(\mathcal{F}_{(\overrightarrow{e_2},\overrightarrow{e_3})}, \mathcal{F}_{\overrightarrow{e_1}})$ as a dictionary with the $C_{\mathcal{B}}$ as a key, and $\mathcal{F}_{(\overrightarrow{e_2},\overrightarrow{e_3})}, \mathcal{F}_{\overrightarrow{e_1}}$ as values

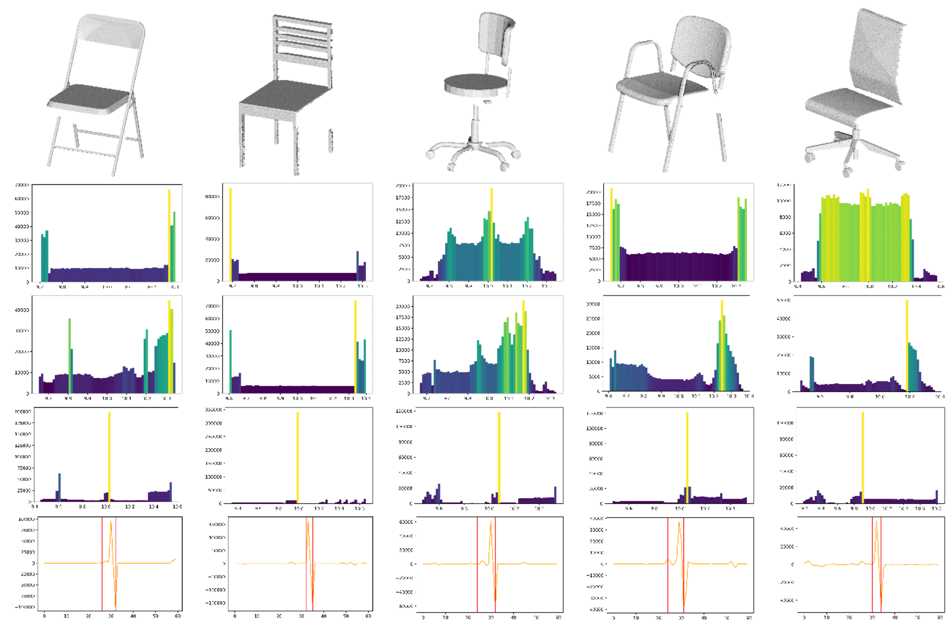
Outputs of **Algorithm 1.3** are used as initial shape descriptors for studying local maxima. This is done through a gradient approach with different neighbourhood to avoid over/under segmentation. The gradient $\vec{\nabla}(f)$ is computed using central difference in the interior and first differences at the boundaries:

$$\vec{\nabla}(f) = \frac{\delta f}{\delta x}\vec{\imath} + \frac{\delta f}{\delta y}\vec{\jmath} \tag{1}$$

$$\delta_h[f](x) = f\left(x + \tfrac{1}{2}h\right) - f\left(x - \tfrac{1}{2}h\right), \text{ with } h = 1 \text{ (interior)} \tag{2}$$

$$\Delta_h[f](x) = f(x+h) - f(x) \text{ (boundaries)} \tag{3}$$

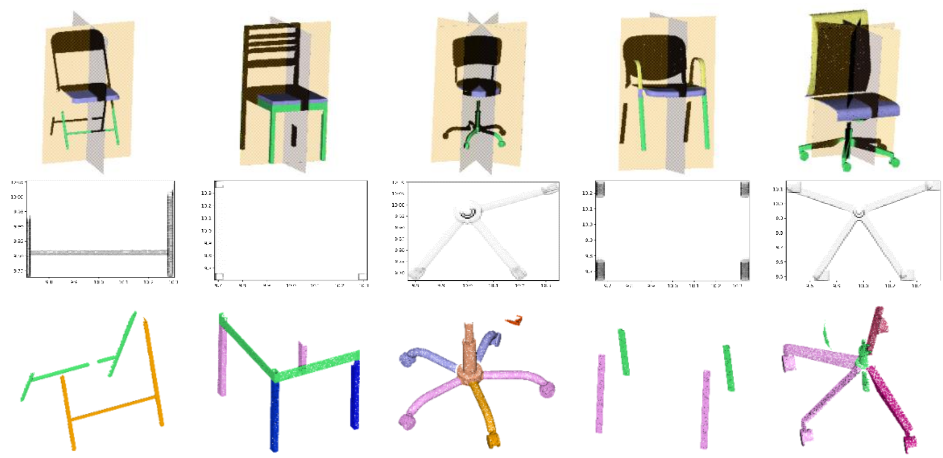
From the extrema, we descend the gradient to find the two cuts candidates: downcut and upcut. This is iteratively refined by studying each extremum, and their relative cuts. When two extrema have a common value for their cuts, they are studied for possible under segmentation, and therefore aggregated into the initial candidate. Cuts candidates are extracted by fitting a linear least-square model to each gradient after extrema's filtering to identify the baseline (Figure 68: line 5). This permits to be robust to varying sampling distance, missing point cloud data and outliers. We then extract the candidate for the Main-Element, and we further process Sub-Elements.



(a) (b) (c) (d) (e)

Figure 68. Each AE from **(a)** to **(e)** is projected in a voxelized space, studied against voxel count per unit over $\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}$ to extract extremums and find patterns that define each Sub-Element as in the ACO. For each object, line 1 is the considered AE, line 2 to 4 illustrates the repartition histogram along $\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}$, line 5 the principal cuts extracted, line 6 the results.

Sub-Step 3. The ontology contains a knowledge-based symmetry indicator which provides insights on possible symmetric properties of each class, expressed regarding the Main-Element (See Supplementary Materials). We first use these and if the test fails we undergo specific symmetry search using analytic knowledge similarly to [312]. In the context of indoor point clouds, we mostly deal with approximate symmetries due to measured data. Therefore, we use a measure of overlap by mapping the pixels as binary grid of the projection on a plane the normal of which is coplanar to the symmetry plane and vice-versa (Figure 69: line 2). The symmetry analysis is conducted over the repartition histogram projected onto the plane $(\overrightarrow{e_2}, \overrightarrow{e_3})$. Then we run mean-shift clustering to detect candidate axis positions among all pairs of neighboring patches similarly to [313].



(a) (b) (c) (d) (e)

Figure 69. Symmetric feature characterization for Sub-Elements of chairs (a) to (e). line 1: symmetric planes; line 2: 2D projection featuring; line 3: similarity's feature tag results to other Sub-Elements.

At this point, we benefit of a better characterization of AE through the ACO and the described threefold mechanism. Each initial element composing the scene (NE, SE and AE) is then processed to extract object relationships. We construct a connected component graph $\mathscr{g}$ in a voxel-space based on the initial bounding-box parameters of $\mathcal{P}$ and the one of the considered element $\mathcal{E}_a$. We then use available topology information $\mathcal{T}_v(v_a, \{v_i\}_{i=1}^{26})$ computed regarding DE-9IM [336] in the voxel space and extended using spatial operators to identify elements relation to $\mathcal{E}_i$ along $\overrightarrow{e_1}$:

$$relation = \{guest, host, twin\}, \tag{4}$$

$$\forall\, i\ \in relation, \exists\ (\mathcal{E}_a, \mathcal{E}_b)\ |\ \mathcal{T}(\mathcal{E}_a, \mathcal{E}_b) = i, \tag{5}$$

$$e(\mathscr{g}) = \ \mathcal{T}(\mathcal{E}_a, \mathcal{E}_b)\ with\ \mathcal{E}_a = \mathscr{v}(\mathscr{g}), \mathcal{E}_b = \mathscr{v}_e(\mathscr{g}) \tag{6}$$

The basic object relationship is defined as $\mathcal{T}(\mathcal{E}_a, \mathcal{E}_b)$ and determined regarding **Algorithm 1.4:**

Algorithm 1.4 Object-relationship definition for indoor elements

Require: A point cloud object $\mathcal{E}_i \in \mathbb{R}^3$ filtered for considering only spatial attributes X, Y, Z along $(\vec{x}, \vec{y}, \vec{z})$ axis, $(\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3})$ the principal directions, X_e, Y_e, Z_e the spatial attributes along $\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}$

1. $\mathcal{T}_v(v_a, \{v_i\}_{i=1}^{26}) \leftarrow$ SPC voxel-based feature for direct voxel topology between $\mathcal{E}$
2. **for each** $\mathcal{E}$ **do**
3. **if** $\mathcal{T}_v(v_a, \{v_i\}_{i=1}^{26}) = true$ **and** $v_i \ \in \mathcal{E}_b$ **then**
4. **if** $\max_{\overrightarrow{e_1}}\left(BBox(\mathcal{E}_b)\right) < \min_{\overrightarrow{e_1}}\left(BBox(\mathcal{E}_a)\right)$ ***and*** $|\min_{\overrightarrow{e_1}}(BBox(\mathcal{E}_a)) - \max_{\overrightarrow{e_1}}(BBox(\mathcal{E}_b))| < KB_{th}$ **then**
5. $\mathcal{T}(\mathcal{E}_a, \mathcal{E}_b) = guest$
6. **elif** $\mathcal{T}(\mathcal{E}_b, \mathcal{E}_a) = guest$ **then** $\mathcal{T}(\mathcal{E}_a, \mathcal{E}_b) = host$
7. **elif** ($\mathcal{T}(\mathcal{E}_a, \mathcal{E}_b) = guest$ **or** $\mathcal{T}(\mathcal{E}_a, \mathcal{E}_b) = host$) **and** $\left|X_{\overrightarrow{e_1}}\left(CenterBBox(\mathcal{E}_a)\right) - \quad X_{\overrightarrow{e_1}}\left(CenterBBox(\mathcal{E}_b)\right)\right| < KB_{th}$ **then** $\mathcal{T}(\mathcal{E}_a, \mathcal{E}_b) = twin$
8. **end if**
9. **end if**
10. **end for**
11. **end**
12. **return** $(\mathcal{T}(\mathcal{E}_a, \mathcal{E}_b))$ for every $\mathcal{E}$ composing $\mathcal{P}$

Regarding AE, Sub-Elements creating new segments are refined similarly using the ACO (e.g. if $\mathcal{E}_i$ is a "Kitchen Chair" and the number of $\mathcal{E}_{sub-i}$ in "someDown" position is inferior to 3, then results are refined as a "Kitchen Chair" is described having at least 3 legs). ACO-guided, we cross-relate $\mathcal{E}_i$ information $\mathcal{F\ell}$, $\mathcal{Wi}$, $\mathcal{Le}$, $\mathcal{F}_{(\overrightarrow{e_2},\overrightarrow{e_3})}$, $\mathcal{F}_{\overrightarrow{e_1}}$, $\mathcal{He}$ and $\mathcal{E\ell}$ with the highest maximas, where:

$$\forall\, X_e < Y_e, \exists\, \mathcal{L}e \in \mathbb{R} \mid \mathcal{L}e = Y_e, \qquad \forall\, X_e > Y_e, \exists\, \mathcal{L}e \in \mathbb{R} \mid \mathcal{L}e = X_e \tag{7}$$

$$\mathcal{Wi} = \min(\overrightarrow{e_1}, \overrightarrow{e_2}), \qquad \mathcal{T}h = Z_e, \qquad \mathcal{F}\ell = \frac{\mathcal{T}h}{\mathcal{Wi}}, \qquad \mathcal{E}\ell = \frac{\mathcal{Wi}}{\mathcal{L}e} \tag{8}$$

These are grouped as a bag of features[24] and we infer modelling rules (e.g. Figure 70) after going through a language processing step to provide three group of features being:

- Bag of features: Flatness $\mathcal{F}\ell$, Width $\mathcal{Wi}$, Length $\mathcal{L}e$, Histogram features ($\mathcal{F}_{(\overrightarrow{e_2},\overrightarrow{e_3})}$, $\mathcal{F}_{\overrightarrow{e_1}}$), Height $\mathcal{H}e$, Elongation $\mathcal{E}\ell$, Thickness $\mathcal{T}h$, main orientation $\overrightarrow{e_1}$.
- Object-relationship information: Topology relation $\mathcal{T}(\mathcal{E}_a, \mathcal{E}_b)$, Relative position to elements in fixed radius $\mathcal{R}(\mathcal{E}_a, \mathcal{E}_b)$, Direct Voxel-based Topology $\mathcal{T}_v(v_a, \{v_i\}_{i=1}^{26})$
- Contextual semantics: Semantic Position $\mathcal{R}_s(\mathcal{E}_a, \mathcal{P})$, Function $\mathcal{F}_i$, Label $\mathcal{L}_i$

Contrary to [252], we do not reason on decision trees that are extracted from an ontology. Indeed, we create a JSON object per element $\mathcal{E}_i$ that holds the Bag of features, Object-relationship information and Contextual semantics. These include concepts of physics and causation such as stability, clearance, proximity and dimensions defined as Knowledge Primitives by Sutton et al. [337]. The reasoning module of the expert system can in turn provide the guiding modelling rules $\mathcal{R}$ for the considered object as developed in Section 6.3.1.

[24] For elements that have rectangular cross-sections, the length and width are defined by the size of the minimum bounding rectangle, whereas for elements that have circular cross-sections, the radius is defined by the size of the minimum bounding circle.

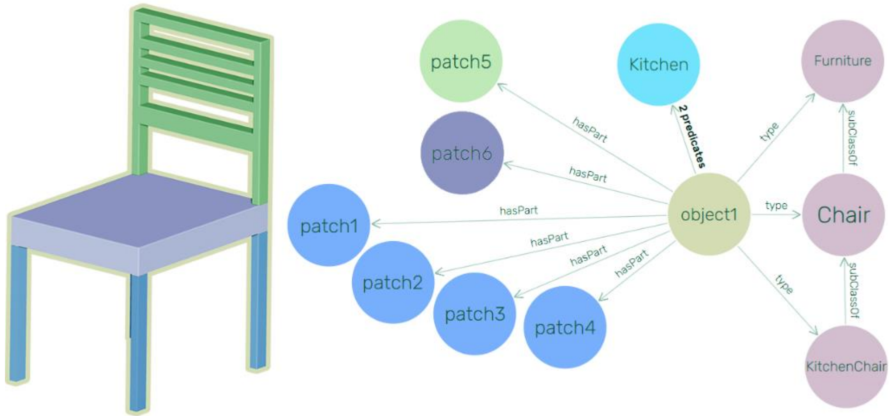


Figure 70. The ACO graph-representation of the chair and the relations within Sub-Elements

In the next sub-section (Figure 62: C) we explain the third step of our global workflow, which provides a modelling approach for obtaining multiple geometries for each $\mathcal{E}_i$ composing $\mathcal{P}$.

6.3.3 Procedural instance 3D modelling

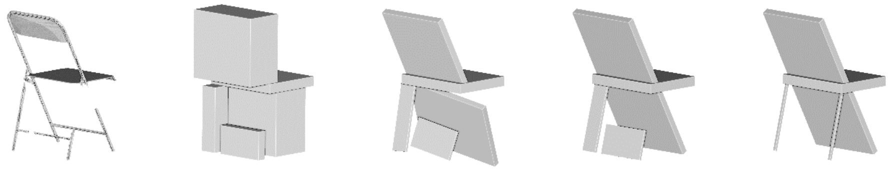
As reviewed in Section 6.2.1, most of indoor scene are primitive-based decompositions. As such we provide a simple yet efficient parametric instance-modelling described in **Algorithm 2** that reconstruct each element $\mathcal{E}_i$ using cuboid and bounded plane representations.

Algorithm 2 Multi-LoD object instance modelling (Figure 62: C)

Require: Algorithm 1 output: $({\mathcal{E}_i}^+, \mathcal{R})$

1. **For each** ${\mathcal{E}_i}^+$ **do**
2. **if** $C_{{\mathcal{E}_i}^+} == 'AE'$ **then**
3. ${\mathcal{E}_i}^+$ goes through part-modelling and part-assemblage
4. **end if**
5. ${\mathcal{E}_i}^+$ parameters refinement through context adaptation following $\mathcal{R}$ specifications
6. $\{\mathcal{M}_{L0}, \mathcal{M}_{L1}, \mathcal{M}_{L2}, \mathcal{M}_{L3}\} \leftarrow$ Primitive fitting for multi-LoD 3D model generation
7. $\mathcal{M}_{3DM} \leftarrow$ Model from 3D Data mining using ModelNet10 based on $\{\mathcal{M}_{L0}, \mathcal{M}_{L1}, \mathcal{M}_{L2}\}$
8. **end for**
9. **end**
10. **return** $(\mathcal{M}_{L0}, \mathcal{M}_{L1}, \mathcal{M}_{L2}, \mathcal{M}_{L3}, \mathcal{M}_{3DM})$

Similarly to Step 2 (Figure 62: B, Section 6.3.2) of the global workflow, the characterization of AE element as ${\mathcal{E}_i}^+$ is accounted for by going through a specific part-modelling and part-assemblage processing (necessitate AE's characterization and part-segmentation as detailed in Section 6.3.2). This is done by considering each Sub-Element of the initial element $\mathcal{E}_i$ an independent element. Then, using intra-$\mathcal{E}_i$ topology (relations between Sub-Elements from part-segmentation) defined within the ACO, geometric parameters are adjusted (Figure 71).



(a) (b) (c) (d) (e)

Figure 71. The different phases of the primitive fitting for AE. **(a)** point cloud; **(b)** Raw parameters and generation of grid-aligned cuboid; **(c)** Refinement by non-constrained PCA-Analysis; **(d)** Refinement by constrained PCA-Analysis; **(e)** parameters refinement through ACO

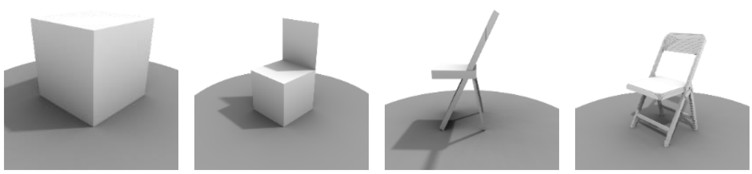
The parametrization of generated models gives users the ability to alter the entire logic of the object construction by adjusting individual parts.

We parametrize a cuboid with three orthogonal directions $\overrightarrow{c_1}, \overrightarrow{c_2}, \overrightarrow{c_3}$ where:

$$\overrightarrow{c_3} = \overrightarrow{c_1} \times \overrightarrow{c_2} \tag{9}$$

The cuboid parameters also include the center coordinates X_c, Y_c, Z_c, as well as the length, width and height respectively along $\overrightarrow{c_1}, \overrightarrow{c_2}, \overrightarrow{c_3}$. Its finite point set representation is obtained by tessellation and generated as an .obj file. A bounded plane is represented by a set of parameters $p = \{p_1, p_2, p_3, p_4\}$ that defines a plane, and a set of edge points e that lies in the plane and describes the vertices of the plane's boundary.

Depending on $\mathcal{E}_i$'s characterization, we obtain a geometric model composed of a bounded plane, a cuboid, a cuboid assembly or a cuboid and bounded plane aggregation. These geometries are then refined to provide multiple LoD. $\mathcal{E}_i$ are found represented as Bounding-box ($\mathcal{M}_{L0}$), Trivial knowledge-based parametric shape ($\mathcal{M}_{L1}$), Parametric assemblage ($\mathcal{M}_{L2}$) and hybrid voxel-based refined model ($\mathcal{M}_{L3}$) as illustrated in Figure 72 and executed in **Algorithm 2**: line 6.



(a) (b) (c) (d)

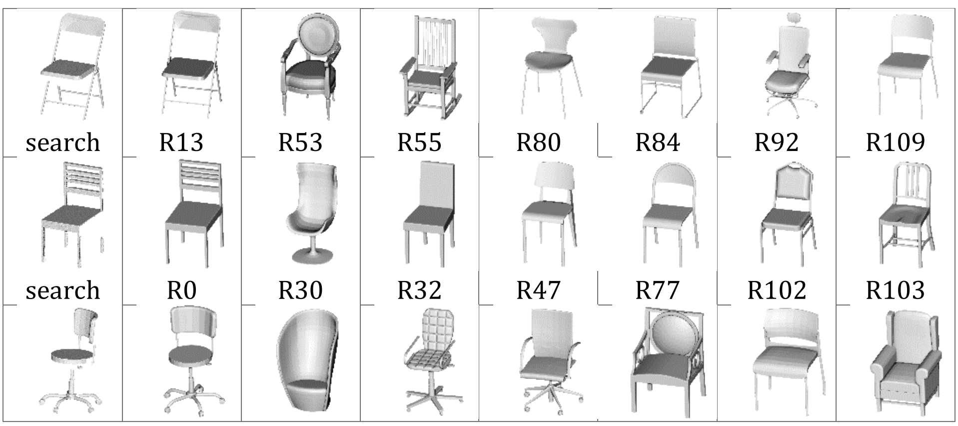
Figure 72. **(a)** Bouding-box $\mathcal{M}_{L0}$; **(b)** KB model $\mathcal{M}_{L1}$; **(c)** Assemblage $\mathcal{M}_{L2}$; **(d)** hybrid model $\mathcal{M}_{L3}$

As an alternative to the 3D models extracted from the procedural engine, we study a 3D database shape matching approach for higher geometric flexibility, but also to provide a way to extract additional information from external database sources through mining. We consider database objects from the ModelNet10 [273] library, specifically the chair, desk and table furniture due to the non-availability of wall, beam, ceiling and floor models (Figure 73). Each candidate is oriented in the same way (Main-Element's $\overrightarrow{e_3}$ is $\vec{z}$-aligned, $\overrightarrow{e_2}$ is $\vec{y}$-aligned) which allows to avoid global alignment search and local refinement via ICP. Only a rigid translation and deformable step is executed. The main challenges include the present noise in scanned data, isotropic shapes, partial views and outliers. As such, global matching methods based on exhaustive search (efficient if we can strongly constrain the space of possible transformations), normalization (not applicable to partial views, or scenes with outliers) and RANSAC (need at least 3 pairs of points) are limited. We investigate an invariance-based method to try and characterize the shapes using properties that are invariant under the desired transformations. We describe each database 3D model by computing a rank $R_{Level-1}$ to define a first set of best fit candidates C_i, where:

$$R_{Level-1}(C_i) = \left|\frac{\mathcal{H}e(\mathcal{M}_{L1})}{\mathcal{L}e(\mathcal{M}_{L1})} - \frac{\mathcal{H}e(C_i)}{\mathcal{L}e(C_i)}\right| + \left|\frac{\mathcal{H}e(\mathcal{M}_{L1})}{\mathcal{W}i(\mathcal{M}_{L1})} - \frac{\mathcal{H}e(C_i)}{\mathcal{W}i(C_i)}\right| \qquad (10)$$

We then compare each rank for each candidate within the database and filter by score. We narrow the set of candidates by comparing its symmetry pointer to $\mathcal{E}_i^{\,+}$ and each Sub-Element (when applicable) that highly constrain the repartition. This is done regarding the symmetry descriptors as defined in Section 6.3.2. The new rank descriptor $R_{Level-2}$ is given by:

$$R_{Level-2}(C_i) = R(similarity_{ratio}) + R_{Level-1}(C_i) \qquad (11)$$



search	R20	R31	R36	R53	R69	R77	R94
(a)	(b)	(c)	(d)	(e)	(f)	(h)	(i)

Figure 73. Results of the shape matching over different datasets **(a)**. The rank RXX represent the score of each candidate. The closer to 0 the best the shape fits the search from **(b)** to **(i)**.

We compute the transformation parameters by matching the centroids of C_i and $\mathcal{M}_{L1}$, refined using $\mathcal{M}_{L2}$ and ACO-inferred ${\mathcal{E}_i}^+$ for its orientation. We finally adjust scale by matching shape parameters. This is done using the symmetry indicators and planes as coplanar constraints for global transformation with independent shape deformation along each principal axis.

Finally, we present in the next section the closing step to aggregate every modelled element to create a 3D model accompanied by object relationships (Figure 62: D, Section 6.3.4).

6.3.4 3D Aggregation for scene modelling

At this stage, every element $\mathcal{E}_i$ composing $\mathcal{P}$ is enhanced to retain contextual information and object-relationship through **Algorithm 1**, becoming ${\mathcal{E}_i}^+$. It is then processed in accordance to $\mathcal{R}$ from the ACO to obtain a set of models $\{\mathcal{M}_{L0}, \mathcal{M}_{L1}, \mathcal{M}_{L2}, \mathcal{M}_{L3}, \mathcal{M}_{3DM}\}$ using **Algorithm 2.** The final step (Figure 62: D) leverage the context with related elements for general modelling, with adjusted parametric reconstruction following notably the topology and symmetric considerations. E.g, the constraints extracted by processing the ACO such as "the feet have the same height" derived from topological reasoning with the ground permit contextual inference. Every element is then aggregated as described in **Algorithm 3**, and object-relationships $\mathcal{R}_{G3D}$ are retained to be usable concurrently to the global 3D indoor model $\mathcal{M}_{G3D}$. This step follows a part-to-whole design which starts with the floors, ceilings, walls, beams/columns and then goes to doors, windows then furniture.

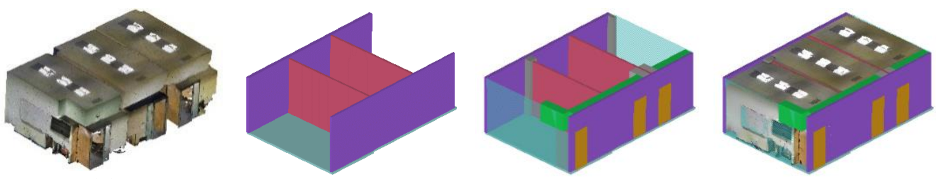
Algorithm 3 Elements aggregation for global geometric and relationship modelling (Figure 62: D)

Require: Algorithm 1 output $({\mathcal{E}_i}^+, \mathcal{R})$ and Algorithm 2 output $(\mathcal{M}_{L0}, \mathcal{M}_{L1}, \mathcal{M}_{L2}, \mathcal{M}_{L3}, \mathcal{M}_{3DM})$

1. $\mathcal{M}_{G3D} \leftarrow \emptyset$
2. **For every** ${\mathcal{E}_i}^+ \in \mathcal{P}$ **do**
3. **if** $\mathcal{L}_i \neq door$ or $\mathcal{L}_i \neq window$ **then**
4. $\mathcal{M}_{G3D} \leftarrow \mathcal{M}_{G3D} \cup \mathcal{M}_{Xi}$: Aggregate desired $\mathcal{E}_i$ geometry ($\mathcal{M}_{Xi}$) from the set of models $\{\mathcal{M}_{L0}, \mathcal{M}_{L1}, \mathcal{M}_{L2}, \mathcal{M}_{L3}, \mathcal{M}_{3DM}\}$ for natural or hybrid models
5. **elif**: $\mathcal{M}_{G3D} \leftarrow (\mathcal{M}_{G3D} \cap \mathcal{M}_{Xi})^C + \mathcal{M}_{Xi} \cap (\mathcal{M}_{G3D} \cap \mathcal{M}_{Xi})^C$
6. **end if**

$\mathcal{R}_{G3D} \leftarrow$ Object Relationship modelling of $\mathcal{M}_{G3D}$
end for
end
return $(\mathcal{M}_{G3D}, \mathcal{R}_{G3D})$

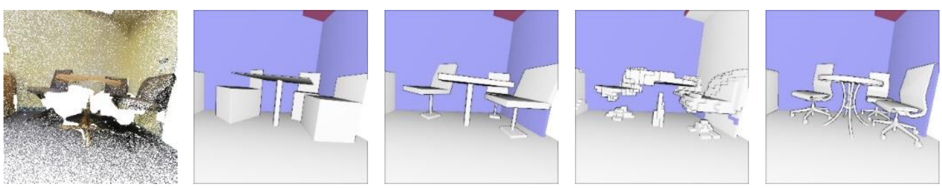
To be topologically consistent in the sense of 3D modelling, we treat overlapping similarly to Fayolle and Pasko [288]. We represent a constructive model by a binary (CSG) construction tree structure with primitive solids at the leaves and operations at the internal nodes of the tree. For any given point in space, an evaluation procedure traverses the tree and evaluates membership predicate at this point. After evaluation, we obtain a consistent 3D model of the entire scene retaining object-relationships. We use the Union (U) and Intersection (∩) operators and their set complements to refine the CSG-tree for modelling the point cloud (Figure 74).



(a) (b) (c) (d)

Figure 74. (a) S3-DIS Point Cloud extract; **(b)** SE modelling (walls, floors, ceillings); **(c)** NE completion and CSG operations; **(d)** 3D global model with point cloud super-imposition

The ability to use several representations from the set of instance models $\{\mathcal{M}_{L0}, \mathcal{M}_{L1}, \mathcal{M}_{L2}, \mathcal{M}_{L3}, \mathcal{M}_{3DM}\}$ permit to obtain hybrid models following the same mechanism as illustrated in Figure 75.



(a) (b) (c) (d) (e)

Figure 75. (a) S3-DIS; **(b)** $(\mathcal{M}_{L0}, \mathcal{M}_{L1})$ model; **(c)** $(\mathcal{M}_{L0}, \mathcal{M}_{L2})$ model; **(d)** $(\mathcal{M}_{L0}, \mathcal{M}_{L3})$ model; **(e)** $(\mathcal{M}_{L1}, \mathcal{M}_{3DM})$ model

In a will to aggregate semantics outside the SPC Infrastructure and for interoperability with existing standards, we obtain a parsing-ready JSON object for IFC file construction. $\mathcal{M}_{L0}, \mathcal{M}_{L1}, \mathcal{M}_{L2}, \mathcal{M}_{L3}$ and $\mathcal{M}_{3DM}$ geometries follow the .obj physical file format to be mapped in the IFC scheme using the EXPRESS data definition language. For beams, floors, and walls, IFC entity types, such as IfcBeam, IfcWall, IfcWallStandardCase, IfcBulidingElementProxy, IfcRelDecomposes and IfcRelConnects are defined with their geometric and connectivity properties following the IFC scheme. In this way, an as-built 3D model of the structural elements compliant with industry standards can be inferred.

6.4 RESULTS

In this section we detail the results of our methodology through several comparisons. We start by describing the underlying datasets, we then present the results of our evaluations to finally provide the details of our implementation and computation time. Then we provide identified limitations and research directions.

6.4.1 Datasets

The methodology was tested over three different datasets. The first dataset (SIM) is simulated data using the ModelNet10 library as scanning environment. The second (DAT) contains real data from actual sites using both the Leica P30 and Trimble TX5 terrestrial laser scanner. The last dataset is the S3-DIS created using the Matterport. The main idea behind using these various datasets is to use the simulated one to test the theoretical base of the proposed approach, while the real datasets cover the difficulties and the efficiency of the method. The two real world scenes (DAT and S3-DIS) represent indoor built environments. For the simulated cases (SIM), the point cloud was generated by 3D mesh tessellation and then adding 2 mm of noise, which is representative of many current laser scanners. Subsampling is not employed at any stage here, both for point clouds and 3D models. The simulated dataset is solely constituted of furniture (chair, desk, table), with around 1 million points per element. The DAT dataset is constituted of 800 million points whereas the S3-DIS dataset contains over 335 million points, with an average of 55 million points per areas (6 Areas). The DAT present many large planar surfaces and a high device accuracy resulting in low noise and a homogeneous point repartition. In contrast, the S3-DIS dataset is very noisy and presents many occluded areas. As these are typical scenes from the built environment, it is worth noting that they all present some significant levels of symmetry and/or self-similarity. Moreover, the DAT and 3D-DIS rely on a scan acquisition methodology that do not cover the full environment presenting many occluded areas.

6.4.2 Comparisons

We tested our approach on both simulated and real-world point clouds of indoor buildings. We first provide in Table 26 results regarding the part-segmentation for AE characterization (Section 6.3.2).

Table 26. Results of the part-segmentation mechanism against manually annotated Sub-Elements. The precision and recall were obtained by studying True Positives, False Positive and False Negatives.

$$precision = \frac{True\ Positive}{True\ Positive + False\ Positive}, recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

Dataset	Element	legs		oneSeat		backRest	
		precision	recall	precision	recall	precision	recall
	chair_0001	99.99%	97.34%	93.52%	99.62%	99.99%	96.87%
	chair_0007	59.18%	99.99%	99.99%	87.22%	99.99%	99.99%
SIM	chair_0008	99.99%	99.99%	98.60%	99.99%	99.99%	98.56%
	chair_0029	99.99%	90.93%	95.82%	99.72%	99.99%	99.99%
	chair_0041	99.99%	99.99%	96.68%	99.99%	99.99%	97.31%
	overall	*91.84%*	*97.65%*	*96.92%*	*97.37%*	*99.99%*	*98.55%*
	F1-score	*94.66%*		*97.14%*		*99.27%*	
		precision	recall	precision	recall	precision	recall
	chair_0001	99.99%	99.90%	98.32%	99.99%	99.99%	99.29%
	chair_0002	99.99%	93.57%	93.59%	83.68%	92.09%	98.97%
	chair_0005	99.99%	99.99%	99.99%	38.36%	29.70%	99.99%
DAT	chair_0006	96.14%	99.99%	87.29%	99.56%	95.02%	96.50%
	chair_0007	96.60%	99.99%	92.35%	88.25%	83.89%	98.79%
	overall	*98.55%*	*98.69%*	*94.31%*	*81.97%*	*80.14%*	*98.71%*
	F1-score	*98.62%*		*87.71%*		*88.46%*	
		precision	recall	precision	recall	precision	recall
	chair_0001	99.99%	99.99%	89.26%	99.99%	99.99%	87.56%
	chair_0002	87.53%	99.99%	80.76%	73.31%	94.98%	91.85%
	chair_0003	99.99%	99.99%	97.51%	99.99%	99.99%	97.54%
S3-DIS	chair_0004	99.99%	99.99%	90.88%	99.99%	99.99%	92.13%
	chair_0005	99.99%	99.99%	90.55%	72.25%	74.94%	92.66%
	overall	97.51%	99.99%	89.79%	89.11%	93.98%	92.35%
	F1-score	*98.74%*		*89.45%*		*93.16%*	

Firstly, we notice an overall precision and recall score above 90% for every Sub-Element of the SIM dataset. Relatively, while the backRest gives the higher F1-score (99.27%), the oneSeat and the legs achieve respectively lower scores of 97.14% and 94.66% . Indeed, these Sub-Elements are more subject to missing data which induce False Negatives. Specifically, the SIM chair_0007 retain a problematic case where the point distribution's feature impact the segmentation through a high number of True Negatives. This can be solved if we include a connectivity step to merge similar Connected Components from looking at their Bag of feature and their voxel topology. As for the recall indices, the problematic zones are often localized at the joints between each Sub-Element, which could be further refined if an additional (time-consuming) nearest neighbour search was implemented.

Logically, the non-simulated datasets DAT and S3-DIS achieve lower scores for the backRest and oneSeat robustness detection. Recall drops by 7.43% in average and precision by 8.90%. This is specifically due to the non-uniform sampling of real world datasets which on top present many occluded areas for these Sub-Element, inducing many False Positive.

Interestingly, we note that both DAT and S3-DIS datasets present an increase of around 4% for the F1-score. Indeed, the low precision and scan angle play in favour of joint identification between the oneSeat and the legs which in turn increase the segmentation accuracy. We could highlight that the quality and robustness of our AE's characterization approach depends on the plane detection quality which is influenced by scanner noise, point density, registration accuracy, and clutter inside of the building.

Secondly, we assess the different LoD $\mathcal{E}_i$ models obtained following Section 6.3.3. As the modelling approach does not aim at a perfect fitting of the underlying point cloud, we used an RMSE indicator for comparison between the model and the different reconstructions (Figure 76). We also compared the different sizes of generated geometries to obtain ratios of precision over complexity.

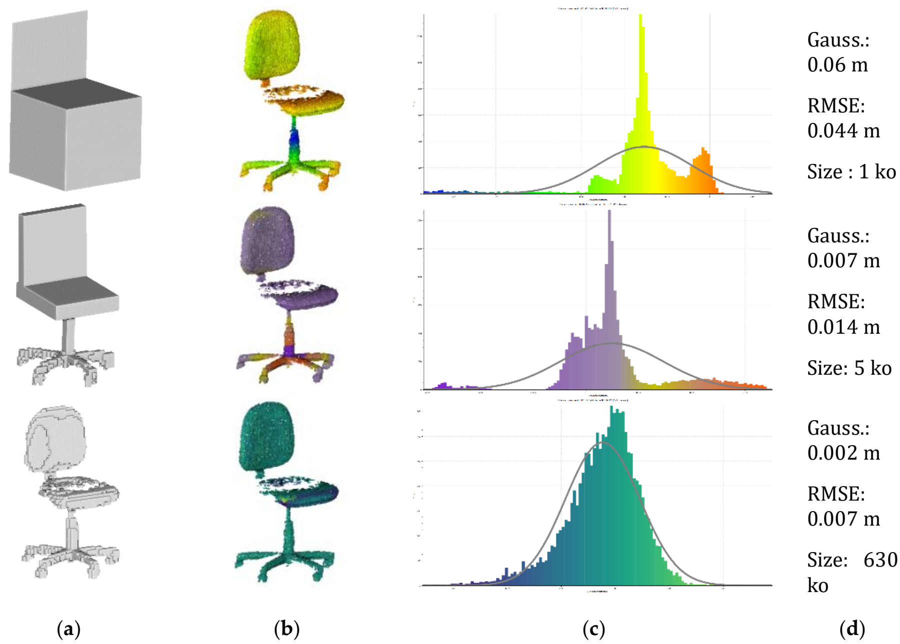


Figure 76. 3D modelling results over $\mathcal{M}_{L1}, \mathcal{M}_{L2}, \mathcal{M}_{L3}$ of the DAT dataset. **(a)** 3D representation; **(b)** colour-coded deviations to **(a)**; studied repartition in **(c)**; main indicators presented in **(d)**.

We notice that the higher the LoD, the better the accuracy but also the higher the data size. Generalized to the elements processed, we extract that RMSE is expected at 5 cm for $\mathcal{M}_{L1}$ (very sensitive to point repartition), whereas $\mathcal{M}_{L2}$ is expected to give a representation with an RMSE of 2 cm, and $\mathcal{M}_{L3}$ is expected to model giving an RMSE of 1 cm. The latest can of course be reduced if the octree level of the voxelisation is lower, causing a higher model size which can become impractical for very large scene. Additionally, $\mathcal{M}_{L3}$ could be refined using the parametric model in zones with a high overlap (e.g. OneSeat area), resulting in a reduced number of vertices. If we look at well-known triangulation modelling methodologies illustrated in Figure 77, while they can provide a higher accuracy, their representation is often incomplete and cannot successfully model occluded areas. On top, their size is in average 6 times bigger than $\mathcal{M}_{L3}$, and the trade-off precision over complexity shows overly complex structures for the precision gains.

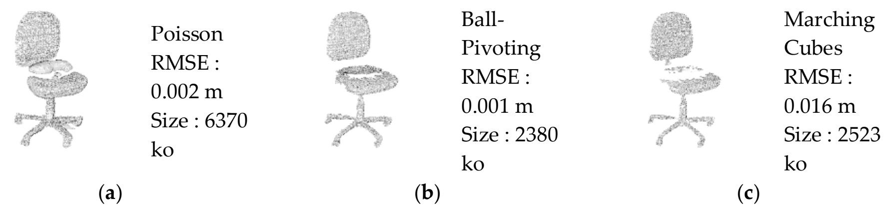
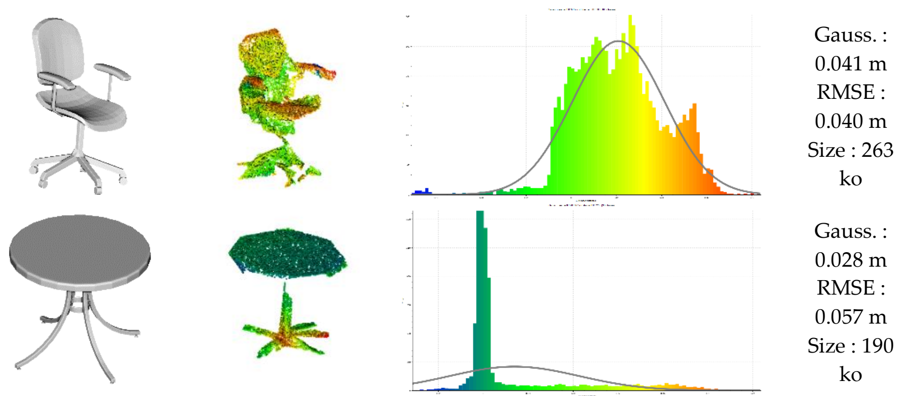
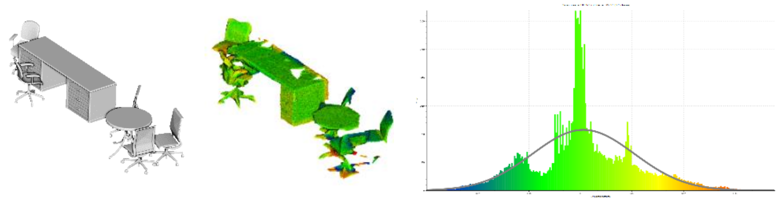


Figure 77. 3D modelling by triangulation. **(a)** Poisson reconstruction [338]; **(b)** Ball-pivoting approach [339]; **(c)** Marching-Cubes approach [340]

As for $\mathcal{M}_{3DM}$, our 3D data mining approach provides interesting results concerning the SIM dataset, but the extension to real world case presents many challenges. Primarily, the fact that the model doesn't exist in the database forces to search for a close candidate and to accommodate intra-$\mathcal{E}_i$ variability. Secondly, the heterogeneity in shapes and forms within the database presents some cases that our algorithm cannot handle, typically when two shapes have a $\mathcal{M}_{L0}$ match, the distinction can provide False Positive. We illustrated the mining results in Figure 78 while assessing the fit's precision.



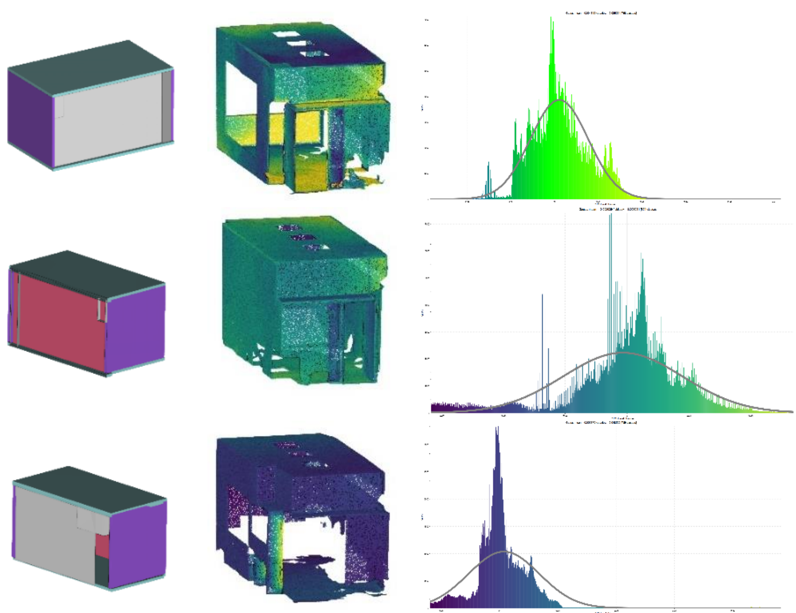


Gauss. : 0.003 m
RMSE : 0.058 m
Size : 3011 ko

(a) (b) (c) (d)

Figure 78. 3D modelling accuracy over the hybrid model. **(a)** constitute the results of the 3D modelling through database mining; **(b)** presents the colour coded deviations to the corresponding model **(a)** and studied by repartition in **(c)**; gaussian, deviation and size are presented in **(d)**.

We notice that using the different LoD for the models within the shape matching approach permit to extract candidates of which the function best fits the indoor scenario. However, the obtained deviations to the S3-DIS point cloud range between 4 and 6 cm which can limit the scenarios of use. Yet, it is important to note that those numbers are heavily influenced by the very high noise of the S3-DIS dataset as well as the large occluded areas. Indeed, one advantage of this 3D mining mechanism is that it provides exhaustive representations from existing models, benefiting asset-management applications. Finally, the proposed methodology permit to reconstruct a global 3D model (Figure 62: C, Section 6.3.4) which was analyzed and illustrated in Figure 79.



Gauss. : 0.004 m
RMSE : 0.026 m
Size : 12.5 ko

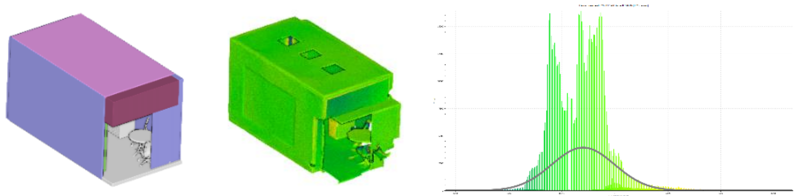
Gauss. : 0.002 m
RMSE : 0.029 m
Size : 18.2 ko

Gauss. : 0.006 m
RMSE : 0.048 m
Size : 14.1 ko

(a) (b) (c) (d)

Figure 79. 3D area-decomposed global model of S3-DIS in $(\mathcal{M}_{L0}, \mathcal{M}_{L1})$. **(a)** constitute the results of the 3D reconstruction modelling; **(b)** presents the colour coded deviations to **(a):** **(c)** represent the deviation analysis; **(d)** regroups main indicators.

We notice that RMSE deviations for $(\mathcal{M}_{L0}, \mathcal{M}_{L1})$ range from 2 cm to 5 cm which are correlated with the scanning method accuracy. On top, the modelling approach which leverages primitives produces an “as-built” reconstruction and therefore doesn’t model small deviations relatively to the global assemblage. If we look closely at the reconstruction of a hybrid $(\mathcal{M}_{L1}, \mathcal{M}_{3DM})$ global model as illustrated in Figure 80, we first notice a very good trade-off between precision reconstruction and size, which is given by its hybrid nature. Moreover, we obtain a coherent watertight CSG assembly usable for simulations as well as 3D printing. On top, the different relations between components allow a selectivity for this printing task.

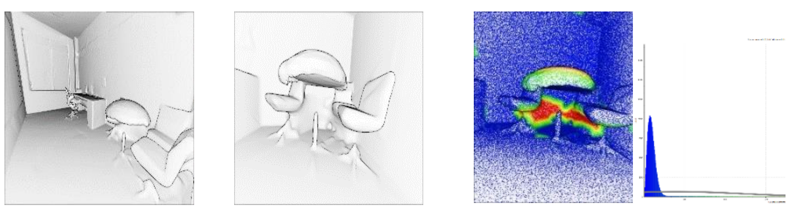


Gauss. : - 0.026 m
RMSE : 0.026 m
Size : 3016 ko

(a) (b) (c) (d)

Figure 80. 3D modelling accuracy over the hybrid model. **(a)** constitute the results of the 3D modelling through database mining; **(b)** presents the colour coded deviations to the corresponding model **(a)** and studied by repartition in **(c)**, and the main indicator are presented in **(d)**.

If we compare to the existing Poisson’s modelling approach (Figure 81), we see that achieved reconstruction’s precisions are often better. Moreover, the size on disk is much larger for the Poisson’s reconstruction. An interesting approach would be to combine a triangulation mechanism such as Poisson to account for small deviations which would extend to “as is” scenarios vs “as-built”.

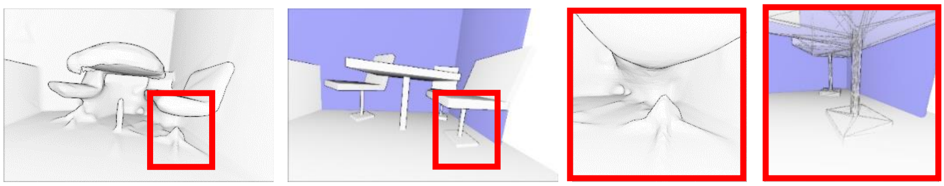


Gauss. : 0.023 m
RMSE : 0.050 m
Size : 25 510 ko

(a) (b) (c) (d)

Figure 81. Poisson reconstruction of the S3-DIS dataset. **(a)** Global view; **(b)** High sensivity to noise and occlusion; **(c)** Poisson’s deviation analysis; **(d)** main indicators

By looking at **(b)** from Figure 81, we also note the high sensitivity to noise and occlusion in the analysed dataset. This is particularly striking for the S3-DIS dataset and strengthen the robustness of our approach to these common artefacts (Figure 82).



(**a**) (**b**) (**c**) (**d**)

Figure 82. Noise and occlusions sensitivity. (a) and (c) shows a Poisson's reconstruction; (b) and (d) shows the $(\mathcal{M}_{L0}, \mathcal{M}_{L2})$ reconstruction.

In the next sub-section, we will investigate the performances and implementation aspects of our approach.

6.4.3 Computation time

We made a prototype implementation of the algorithms described in this paper in different programming languages. All the developments regarding the ACO were made in Java. The different application layers are built on top of RDF and ARQ API of Jena Apache (Java). Jena is an OWL-centric framework that is particularly well suited for our ontology more than OWLAPI, which is RDF-centric. The software Protégé was used as an interface to construct the ACO ontology. The part-segmentation, multi-LoD modelling and database matching were implemented in python using a minimal number of libraries: numpy (for numeric calculations), scikitLearn (for least squares, PCA analysis and signal analysis), matplotlib (for visualization), laspy (for point cloud loading), networkx (for graph and connectivity inference), psycopg2 (for a link to the SPC in-base data, stored in PostgreSQL) and rdfLib (for a connection to RDF triplestore). Visualisations and rendering were made using Three.js or CCLib. All the experiments were conducted on a computer with an Intel Core i7 at 3.30 GHz and 32 GB of RAM. The exchange of information is made through a language processing module which can link SQL statements to JSON, RDF and OWL data, and be manually extended for natural language processing.

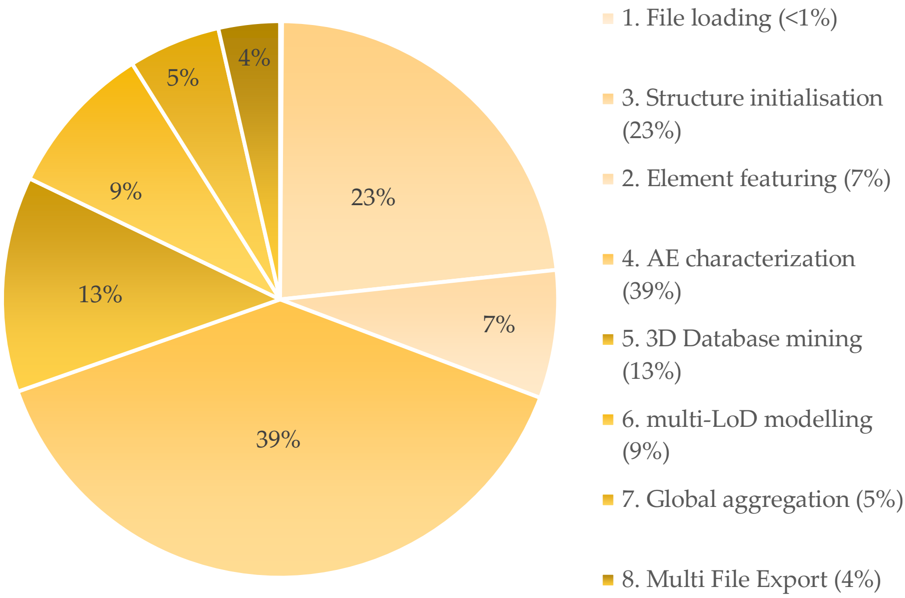


Figure 83. Relative time processing regarding the main elements of the 3D modelling engine

The running times (Figure 83) for the examples presented in this paper as well as some additional experimental datasets range from some seconds for the simpler shapes to several minutes for the more complex shapes. In average the approach takes 85 seconds for the SIM dataset, 32 seconds for the DAT dataset, and 16 seconds for the S3-DIS dataset. Only one thread was used for the computation. The totalling time depends essentially on the size of the point cloud, and therefore the voxelization level retained. We note that there are some threshold and parameters that were determined empirically from our observation, and often their definition has an impact on the runtime. Relatively, the ontology information extraction and inference is quick, followed by the calculation and features in the point cloud (for part-segmentation). The voxelisation is the part that consumes the most memory but can be further optimized by parallelizing its calculation. The structure is already ready for parallel processing. The data mining step can take up to 30 seconds for looking up 900 models in the .off file format and provide the ranking as well as the necessary transformation parameters. Such a search can also be optimized if the models are previously indexed. The CSG integration is quite fast, and usually is done in under 5 seconds. The full workflow from SPC data extraction to multi-LoD modelling and shape matching takes around 5 minutes for a full scene. The IFC file creation is made based on attributes in the JSON file format using the FreeCAD python wrapper.

6.4.4 Limitations

In an attempt to provide a clear list of research directions, we identified ten main points that can be further investigated:

1. In our approach, we consider planar shapes only or manufactured shapes. It would be interesting to extend the method to more complex parametric representations as reviewed in Section 6.2.1.
2. We consider the initial segmentation perfect. While the proposed algorithms are robust to false positives on planar shapes, handling failure cases that can arise when detecting furniture elements would permit to extend the depth of the framework.
3. In our comparison and results analysis, we noticed deviations with elements which present a non-planar morphology. Adding a layer of shape deformation processing to best fit shapes is an opening to provide a compact hybrid model.
4. The ACO was defined using expert's knowledge, shape grammars and standards in use in Europe, thus presents limitations linked to knowledge standardization. Extending the "standards" and features through machine learning could help to better generalize.
5. Our voxel-based clustering approach is dependent on the underlying point data and density, and therefore it can have a high memory footprint thus time execution. We investigate the parallelization of computation to alleviate the processing and extend it to multi-LoD octree-based analysis.
6. The binning and model fitting step (Section 6.3) depend on the initial axis orientation's determination. Extending its sturdiness to highly noisy and non-uniform point sampling would extend the flexibility of the workflow.
7. The considerations in this paper and tests were conducted in indoor built environments only. Research to extend it to other scenes and outdoor scenarios is compelling.
8. Non-standard shapes are difficult to describe through a knowledge-based approach. This limitation comes from the nature of ontologies to be integrated in standardisation and interoperability workflows. A solution would be to compute robust features through a learning network on the existing set of 3D shapes.
9. In our experiments we mainly considered gravity-based scenes with an initial constraint regarding the object orientation. A global registration method would give additional flexibility about the prerequisites for the input dataset.

10. We used the ACO for guiding the modelling process only. Due to its conception, it could be used as an ontology of classification to classify a point cloud in elements described within the OWL.

6.5 PERSPECTIVES

Admittedly, the presented work is merely making one step forward in solving the general problem of 3D Point Cloud modelling. It raises several research directions described in Section 6.4.4, which arise from several identified limitations. It is important to note that our approach is based on a contextual analysis of our environment looking at how elements interact with each other. As such, extending the methodology needs a generalization effort regarding knowledge processing. Indeed, as it is based on an ACO knowledge representation of a specific application, the establishment of the ontology as it stands can in turn limit the interoperability with another domain. However, the approach shows how the context and all its implications regarding object relationships can be used for efficiently modelling point clouds. Going into details, the initial characterization of input shapes needs to be sufficiently meaningful especially the part-segmentation for AE. Furthermore, stitching parts as models together especially for man-made shapes is quite a difficult problem. It often requires resolving topological inconsistencies between parts and the global problem is still an active research area. Our current solution to part assemblage is undeniably simplistic, thus tackling the general problem presents interesting directions for future research. We are seeing a rapid accumulation of 3D models, yet most of these are not semantically described and solely represent geometric shapes. We believe that the analogy to a set of shapes as presented in Section 6.3.4 is a great way for shape retrieval and semantic completion. It can also be used for producing new variations of existing objects as observed by Xu et al. [322]. As shown in this paper, context-based categorization can be an effective mean to this end. Also, Description Logic (DL) complexity of ACO ontology is SHOIQ(D)[25]. OWL2 and its defined relations are the higher level of definition currently defined by OGC specifications. Therefore, in terms of calculation complexity, the proposed ontology is a high-level semantic definition, which requires a heavy calculation process. To reduce this complexity: functional (F), inverse (I), reflexive and disjoint (R) relations will be rethought as much as possible in future work. It is worth mentioning that, while in this work we use the KR for guiding the modelling engine, it can also be used as a classification ontology. Pellet [341] or HermiT [342]

[25] SHOIQ(D) is a naming convention in Description Logic describing the complexity of reasoning in a knowledge base. SHOIQ is a level of complexity between OWL2, which is SROIQ and OWL-DL, which is SHOIN. Each character in the naming convention means that a logic constructor is used.

reasoners are required because of their support of OWL2 and SWRL built-in functions.

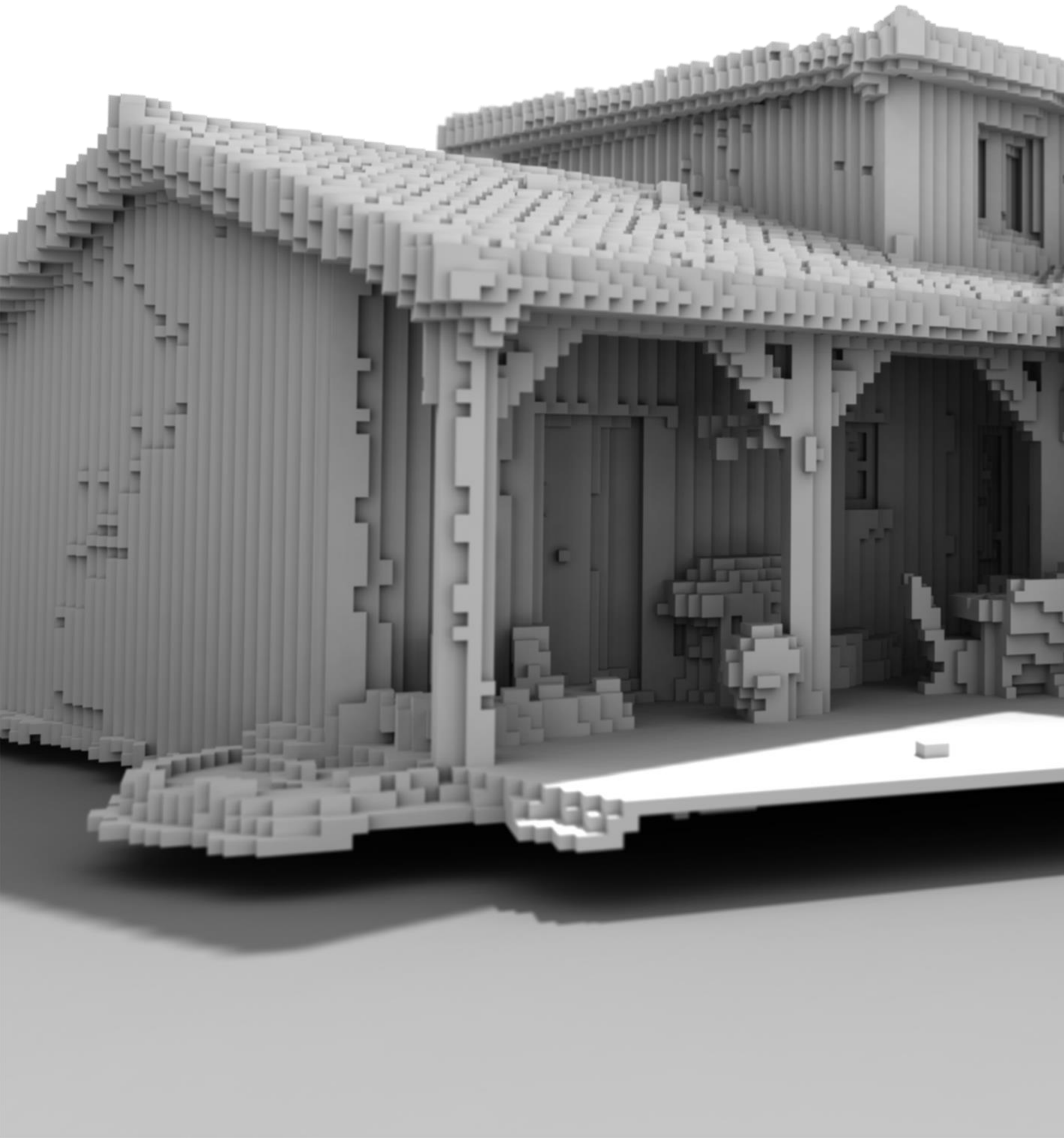
Our part-based segmentation mechanism cannot efficiently handle non-standardized conception which presents problematics for precise identification. On top, complex configurations such as folding chairs are currently not processed by our modelling engine. This could be solved by extending our ACO or through 3D database mining (if models exist in database). Our 3D shape matching procedure is also very interesting for two reasons. Firstly, by using topology, feature similarity and contextual information we can recognize within a given space similar shapes, which provide a new way of modelling incomplete scene, or for variability analysis. Secondly, by looking up a 3D database, we can in turn extract the attached semantics to the fitted candidates and enrich the semantics of the 3D models as well as the underlying point cloud. Finally, it can be used as a mean not only to reconstruct and model an object as a B-Rep or primitive based representation, but to create an open link on the database model and its affiliate information. Indeed, this permits to extract the added information (dimensions, price, availability …) that a hosting database stores for asset management. This provides a great opening to interconnected networks of information that transit and avoid unnecessary multi-existence.

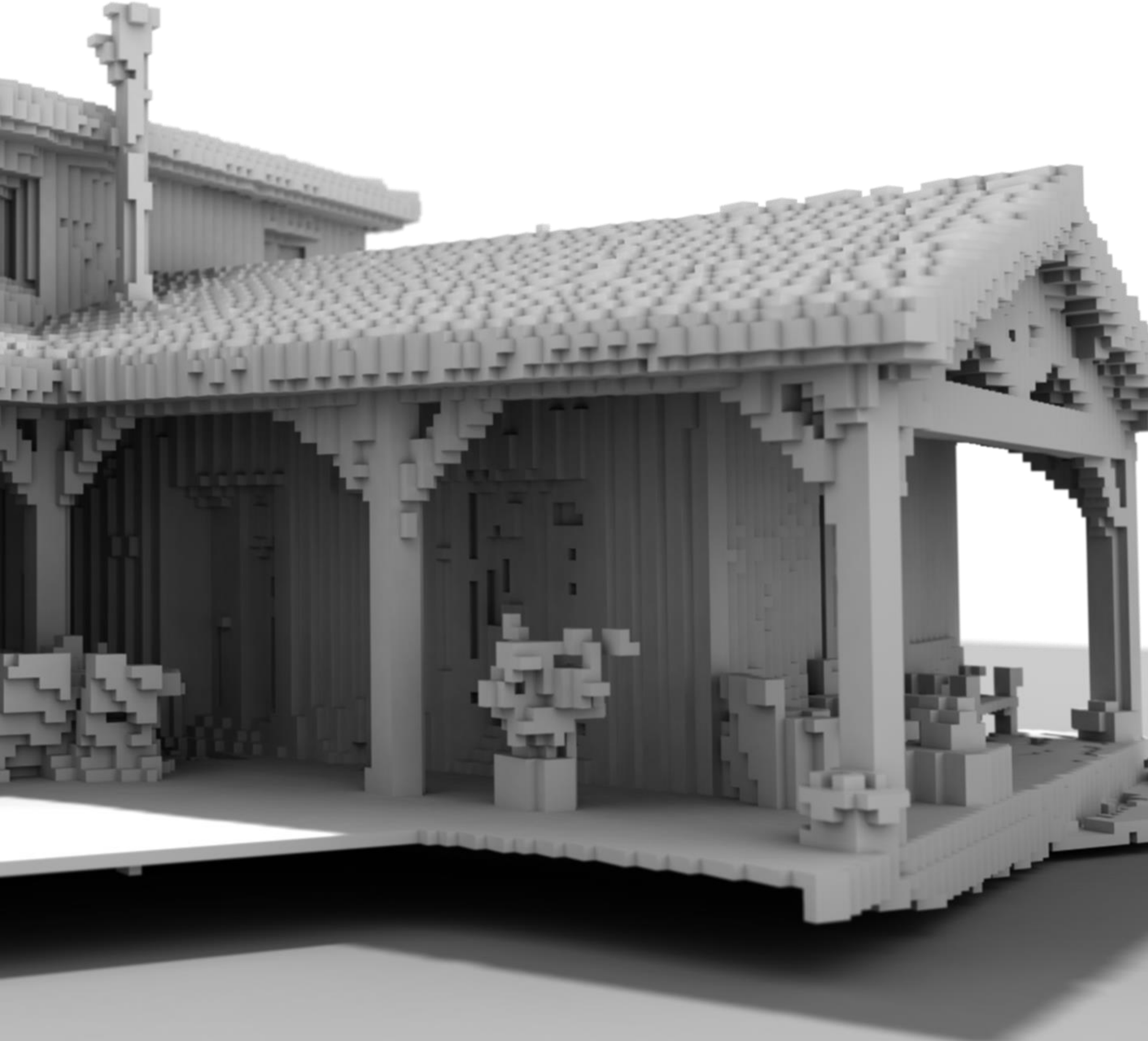
The scalability to bigger building complexes was proven over the real-world datasets, but as denoted, the efficiency can be improved through a better implementation. We focused on a geometry from terrestrial sensors with varying quality, but we intentionally left out color and texture for their high variability in representativity. However, it can be useful in future considerations to better describe shapes or as a mean to extract better feature discrimination. Finally, in our approach we tried to keep in mind the final use of the extracted 3D models similarly to [284]. Whether the goal is the production of indoor CAD models for visualization, more schematic representations that may be enough for navigation or BIM applications, or simply for scene understanding and object localization. In all these scenarios the representation of the final objects differs, but the workflow of our modelling engine to generate several shape representations coupled with object relationship is particularly adapted.

To our eyes, one of the most important perspective concerns the interoperability of the approach within the SPC Infrastructure, acting as a module. Indeed, both are concerned with domain generalization, and the ability to extend workflows to all possible applications. Shape representation in different granularities is a step toward such a flexible use of semantically rich point cloud data. The 3D representation variability given by our multi-LoD approach provides a high flexibility when we look at attaching geometries to subset of points (specifically class instances). This in turn provides queries and filtering capabilities which offer better insight for a new range of scenarios. It also shows how the SPC Infrastructure can be used to provide deliverables for applications such as BIM modelling, virtual inventories or 3D mapping.

6.6 CONCLUSIONS

We presented an automatic method for the global 3D reconstruction of indoor models from segmented point cloud data. Our part-to-whole approach extracts multiple 3D shape representations of the underlying point cloud elements composing the scene before aggregating them with semantics. It provides a full workflow from pre-processing to 3D modelling integrated in a knowledge-based point cloud infrastructure. This permits to leverage domain knowledge through a constructed applicative context ontology for a tailored object characterization at different conceptual considerations. Comprising a 3D modelling step including shape fitting from ModelNet10 for furniture, our approach act as an expert system which output different .obj files as well as a semantic tree. The framework contributes an IFC-inspired as-built reconstruction of the global scene usable by reasoners for automatic decision-making.





Point Cloud voxelisation of my creator's home, Toulouse, France.

CHAPTER 7

Conclusion and research perspectives

Chapter 7 - Conclusion and research perspectives.......213

7.1 Key findings and contribution....... 215
7.1.1 How to structure efficiently point clouds with domain knowledge for interoperable workflows?....... 215
7.1.2 How to leverage a data structure to automate object detection over massive and heterogeneous point clouds?....... 216
7.1.3 How to connect reasoning services for autonomous decision-making scenarios?....... 217
7.1.4 How can open-source database systems integrate point cloud data and semantics? How should one provide domain connectivity? 218
7.1.5 How can a system handle the heterogeneity found within point clouds datasets and semantics for object detection? Can unsupervised frameworks relate to domain concepts?....... 219
7.1.6 How modular is the Smart Point Cloud Infrastructure? How efficient are the proposed point cloud processing modules (segmentation, classification, semantic injection, semantic modelling)? 220
7.2 Research directions....... 221
7.2.1 Better generalization....... 221
7.2.2 Better domain knowledge integration 221
7.2.3 Improve reasoning through AI-based inference....... 222
7.2.4 Dynamic and temporality integration 222
7.2.5 Enhanced domain-based segmentation for archaeological findings 223
7.2.6 Enhanced Classification module....... 223
7.2.7 Integrate a natural language processing module 224
7.2.8 Extending the smart modelling process 224

7.1 KEY FINDINGS AND CONTRIBUTION

The concept of Smart Point Cloud resolves many interrogations that were formulated along this dissertation. Its definition as an integrated knowledge-based point cloud structure extendable through domain formalizations solves several application challenges. The research presents a general automation framework for the development of point cloud semantic representations. The SPC Infrastructure is interoperable and can be used as a starting point for integrating and collaborating around digital reality. Following are short answers regarding the questions initially formulated in Section 1.3.

7.1.1 *How to structure efficiently point clouds with domain knowledge for interoperable workflows?*

Main answers in Chapter 3

The purpose of the SPC is to establish a versatile knowledge-base which can produce domain-dependent 3D semantic representations. It is conceived around an elastic 3-block categorization: device-related knowledge, domain-based knowledge and analytical knowledge. These are co-dependent and can be used in synergy for specific processes as shown in Chapters 3, 4, 5, 6 and summarized in Figure 84.

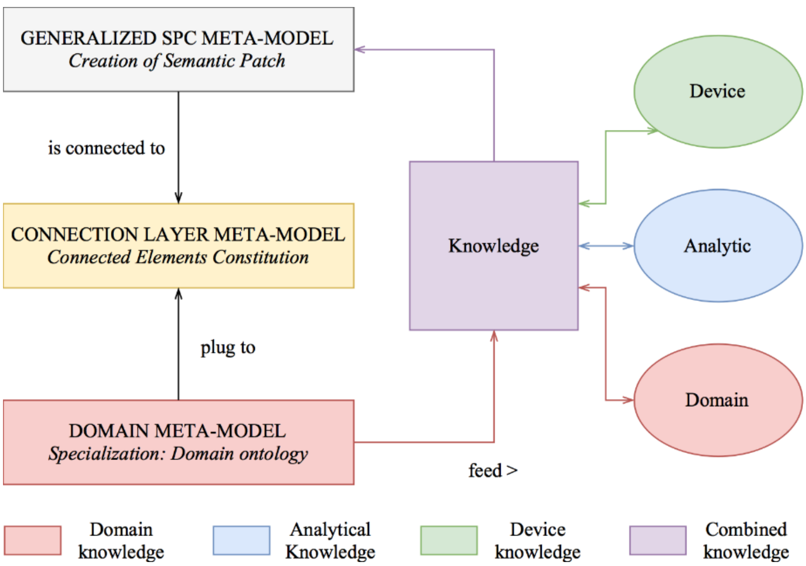


Figure 84. Meta-model articulation for the creation of a SPC

This categorization is found at the structuration level to permit information extraction while retaining meaning and concepts for the creation of semantic depictions. This is made possible by the proposed generic model which defines a conceptualization on which different knowledge formalization can be attached. The approach was thought to allow a maximum flexibility. The data model divides the characterization in different hierarchical levels of abstraction to avoid overlap to existing models and enhance the opening to all possible formalized structure. The core instruction is that the higher levels are closer to a domain representation than lower levels imposing their constraints. The overall structure is a pyramidal assembly, allowing the resolution of thematic problems at higher levels with reference to constraints formally imposed by the lower levels.

7.1.2 How to leverage a data structure to automate object detection over massive and heterogeneous point clouds?

Main answers in Chapter 4

The SPC Infrastructure is initially guided by an unsupervised clustering approach conditioning the data structure (Figure 85). Conceived as a parser module, it enables a spatio-semantic structuration of the point cloud in Connected Elements. By extracting segments without going too deeply in their characterization, the SPC is highly interoperable permitting a compatibility to several application domain. One can also leverage massive point cloud data using the block-storage model constructed using semantic patches and generalized geometries. These entities host compact and representative feature sets such as low-level descriptors and connectivity cues (Chapter 4). They are a flexible and efficient starting point for object detection frameworks such as knowledge-based approaches (4,5) or supervised approaches such as convolutional neural networks (3).

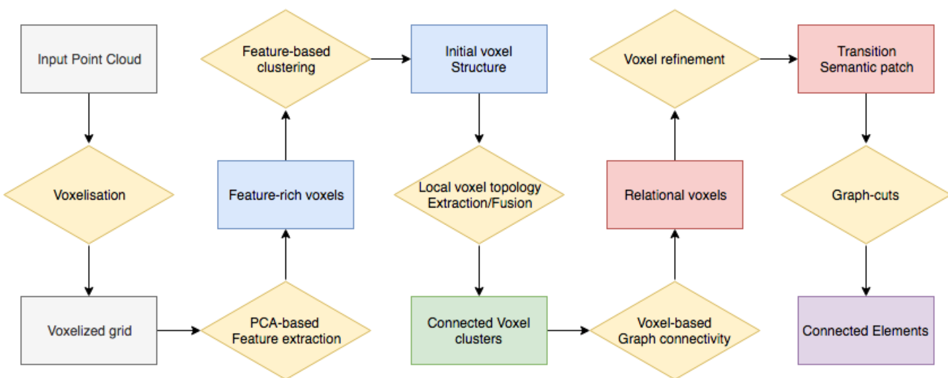


Figure 85. The SPC parser module for a block-model structuration.

7.1.3 *How to connect reasoning services for autonomous decision-making scenarios?*

Main answers in Chapter 6

One of the most important feature of the SPC is its interoperability through a modular conception. While it can be used off-the-shelf as a full-versed solution, its flexibility permits to connect existing decision-making sub-modules. One can link efficiently agent decision support systems for reasoning purposes. While developments were carried in a 3D voxel space with gravity constraints using CEL Bounding-Boxes, shape representation in different granularities is a step toward such a flexible use of semantically rich point cloud data. The 3D representation variability given by our multi-LoD approach provides a high flexibility when we look at attaching geometries to subset of points (specifically class-instances). This in turn provides queries and filtering capabilities which offer better insight for a new range of scenarios. It also shows how the SPC Infrastructure can be used to provide deliverables for applications such as BIM modelling, virtual inventories or 3D mapping.

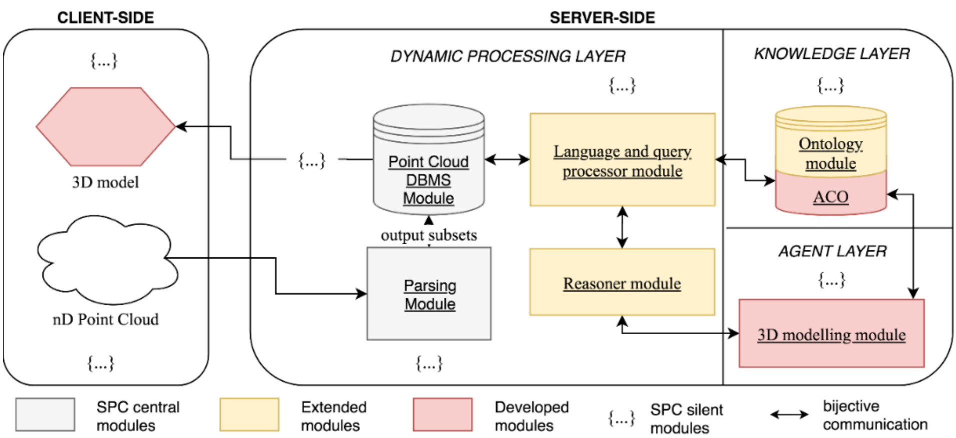


Figure 86. The 3D modelling module part of the agent layer to permit a link to different generalizations and data representations

7.1.4 *How can open-source database systems integrate point cloud data and semantics? How should one provide domain connectivity?*

Main answers in Chapter 3

The Point Cloud DBMS module of the SPCI follows a block-storage model (semanticPatches). The integration in open-source databases such as PostgreSQL is then able to handle massive amounts of point data and semantics.

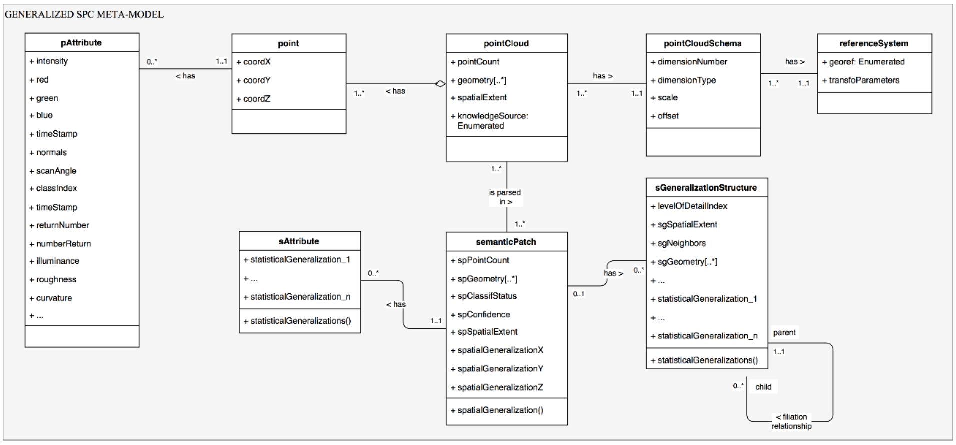


Figure 87. The SPCI conceptual meta-model (level 0) for the structuration of point and semantics.

SemanticPatches are in fact pure voxels or leaf from voxels when features orient a division, then combined to form Connected Elements. To open on different applications, domain knowledge formally expressed as an ontology can directly be connected via two entry-points (Subspaces and Connected Elements) to permit a flexible domain connectivity. For example, we connect the SPC to archaeological ontologies (5) such as CIDOC-CRM, to IFC-inspired ontology (3), to ontology of classification (4, 5, 6) or to objects definition in sub-Elements (6).

7.1.5 How can a system handle the heterogeneity found within point clouds datasets and semantics for object detection? Can unsupervised frameworks relate to domain concepts?

Main answers in Chapter 4 and Chapter 5

Point clouds from reality capture devices suffer from missing data (occlusion, material properties, Improper set-up), erroneous data (noise, outliers, misalignment) and resolution variations which often delivers incomplete datasets. These artefacts augment the heterogeneity in which point clouds are found complicating generalized workflows. The SPC parsing module by only using X,Y and Z attribute is capable of integrating any point cloud, further extracting voxel-based feature sets. This automatic process is very robust to artefacts listed above with an in-depth test on the S3DIS dataset, subject to very high missing and erroneous data. On top, the SPC integrates point cloud from dense-matching sources, TLS, MLS, BMLS, HHLS with no difficulties thanks to a fully unsupervised clustering approach leveraging both low-level features and connectivity cues. It leads to Connected Elements, which are then directly connectable to domain knowledge sources to attach further semantics to these conceptual objects through classification routines. These can be expressed as decision trees, graph-based structures or ontologies.

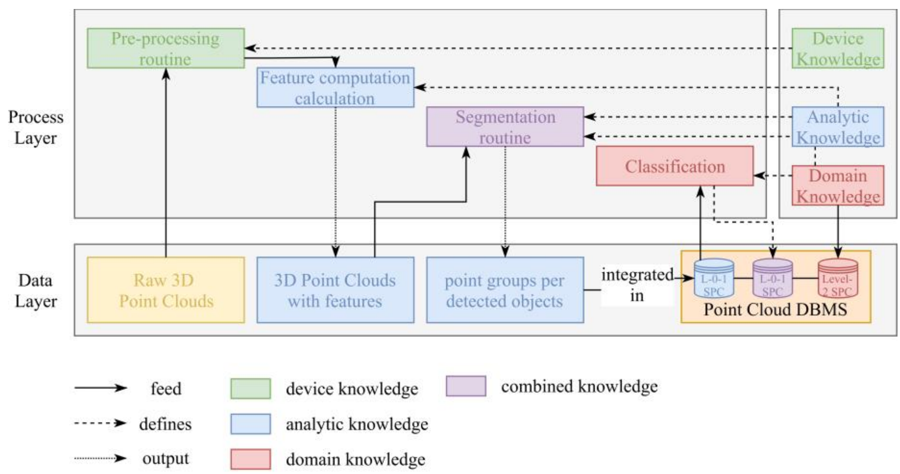


Figure 88. Server-side data management system. Point clouds go through different processing steps, and point groups based on the definition of objects regarding domain knowledge are constituted and populate the SPC database.

7.1.6 *How modular is the Smart Point Cloud Infrastructure? How efficient are the proposed point cloud processing modules (segmentation, classification, semantic injection, semantic modelling)?*

Main answers in Chapter 4, Chapter 5 and Chapter 6

The SPC provides enough elasticity to centralize knowledge for decision-making scenarios. Its conception allows a mapping between domain specializations thanks to the different modules.

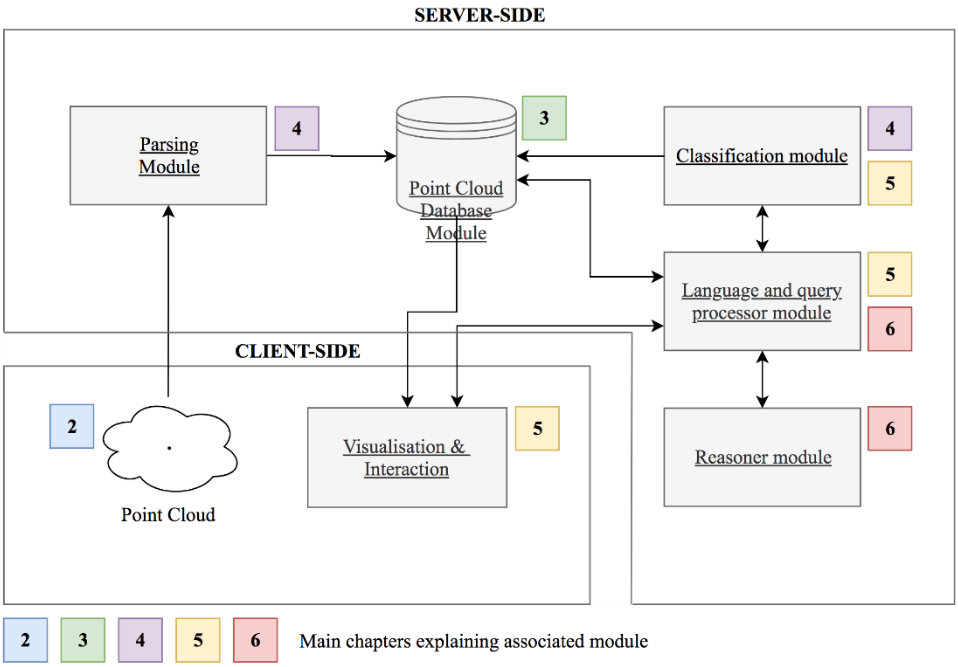


Figure 89. The Smart Point Cloud Infrastructure and its modular architecture associated to the thesis chapters.

The point cloud parsing module is provided as an unsupervised voxel-based clustering approach delivering spatially sound segments. It extracts features at different levels of detail, from the point level to any generalized geometry. It can be extended by adding new feature extraction possibilities or by choosing another neighbouring approach.

The Point Cloud Database module hosts the point cloud in semantic patch with its related semantics. It leverages GIS functionalities and extend 3D analysis by giving multiple geometries of references for detailed and efficient analysis. It can be extended by providing an indexing structure

The Point Cloud Classification Module includes multiple solutions. First, a deep learning method based on PointNet Vanilla to permit supervised learning and the integration of new training data giving more robust results. Then, a decision- tree permitting a well performing classification with a high F_1-score (>85%) for planar-dominant classes comparable to state-of-the-art deep learning. Finally, a domain specialization for tesserae extraction giving an average 95% recognition accuracy for gold tesserae, 97% for faience tesserae, 94% for silver tesserae and 91% for coloured glass.

The reasoner module was populated with a 3D modelling agent that permit initially part-segmentation (F_1-score>95% for chair elements) followed by efficient multi-modal modelling. Comprising a 3D modelling step including shape fitting from ModelNet10 for furniture, the module act as an expert system which output different .obj files as well as a semantic tree.

The visualisation module as a web-based platform is oriented to permit efficient collaboration by different actors and constitute an interactive human-in-the-loop opportunity to establish training datasets.

7.2 RESEARCH DIRECTIONS

7.2.1 Better generalization

Any modelling choice is arbitrary and depends on the conscious or unconscious aspirations of the designer. Although our work responds to a concern for generalization at a spatio-semantic level, it nevertheless remains that it is not totally independent of a certain context. It is for this reason that we wanted to clearly illustrate a privileged domain of application: indoor environments (for BIM, emergency response, inventory management, UAV collision detection …). This choice permits to explore different scales and configurations for deeply and entirely testing our developments. It is also ideal for the definition of new virtual spaces, and the GIS demand associated to such environment is ever increasing. Therefore, as the formalization of domain constantly evolve, research tracks for testing the SPC against other environments and applications is important to improve and tweak its core for a wider deployment.

7.2.2 Better domain knowledge integration

While the SPCI provides direct integration of formalized or inferred domain knowledge, future work including the extensibility of the proposed model to other data types, as well as a better integration of learning routines and knowledge sources is very interesting. As such, some work has been undertaken for backpropagating learning parameters in fuzzy networks to provide a very high flexibility in the definition of ontologies thresholds. Also, the Semantic Web and resources present a variability which is to be thoroughly studied to provide a more flexible way to directly integrate knowledge from trusted sources.

7.2.3 Improve reasoning through AI-based inference

The SPC infrastructure permits to link agent decision support systems and reasoning. However, finding an efficient AI-based agent can become complicated for one application, so generalizing software agents is a massive challenge. To have an intelligent agent that performs reactively and/or pro-actively, interactive tasks need to be tailored to a user's needs without humans or other agents telling it what to do. To accomplish these tasks, it should possess the following general characteristics in regard to [130]: (1) Independence, (2) learning, (3) Cooperation, (4) Reasoning, (5) Intelligence. This relates to the exploratory research field of for Artificial General Intelligence[26] [131] which explicitly justify a need of virtual environments that incentivize the emergence of a cognitive toolkit, such as the SPC infrastructure. As such, SPC-based multi-agent environments provide an opening thanks to its variety (the optimal strategy must be derived optimally) and natural curriculum (the difficulty of the environment is determined by the skill of other agents). This direction while being extremely exciting to avoid purpose-specific algorithm is still a research question that need to be further explored, in which deep Learning may provide a suitable answer.

7.2.4 Dynamic data and temporality integration

While 3-dimensional spaces are strongly inferred in the SPC model, 4-dimension spaces integrating time or by extension n-dimensional spaces are possible characterizations for a greater interoperability. Tests were conducted with static point data only, but varying positions in space and time present additional problematics that could be addressed through new modules or an extended SPC data model. Better integration such as continuous data or the storage and reasoning over datasets covering one location at different time intervals has yet to be further investigated. Indeed, new descriptors emerging from change detection could provide new insights and possibilities for many industries. Also, this opens on continuous

---

[26] Artificial General Intelligence is the intelligence of a machine that could successfully perform any intellectual task that a human can do, including reasoning; judgment calls under uncertainty; KR; planning; learning; natural language communication. Other important capabilities include the ability to sense (e.g. see) and the ability to act (e.g. move and manipulate objects) in the world where intelligent behaviour is to be observed.

Level of Detail possibilities which can provide better ways of handling the amount of data within the Infrastructure.

7.2.5 Enhanced unsupervised feature extraction and segmentation

It is important to retain an unsupervised approach to prepare the data to be usable by many classifiers thus applications. On top, the creation of links between Connected Elements is a novelty which provides interesting perspectives concerning reasoning possibilities that plays on relationships between elements. Additionally, the decomposition in primary, secondary, transition and rest elements is very useful in such contexts as one can specialize or aggregate elements depending on the intended use and application. Indeed, the approach permits to obtain a precise representation of the underlying groups of point contained within Connected Elements and homogenized in Semantic Patches. This can be enhanced by providing a more defined segmentation for non-planar scenes where other features could be extracted permitting an over-segmentation suited for very complex scenarios. Promisingly, the currently extracted relational graph is fully compatible with the Semantic Web and can be used as a base for reasoning services, opening on several possible applications (digital inventories, 3D modelling, BIM, Facility Management …) to be further studied. Of course, it is also easily extensible by improving the feature determination and enhancing the performances.

7.2.6 Enhanced classification approaches

The established ontology-based classification method (indoor assets and archaeological tesserae) is improvable by providing more flexible definition of classes. For example, one can differentiate clutter based on connectivity and proximities to further enhance the classification (E.g., clutter on top of a table may be a computer; clutter linked to the ceiling and in the middle of the room is a light source …). Specifically, it can undergo an ontology refinement to provide a higher characterization and moving thresholds to better adapt the variability in which elements are found in different datasets. Also, supervised learning approaches are promising ways to define a well-performing classifier, but the constitution of several training datasets for each application represent a largely time-consuming task. As such, the ontology approach permits to quickly obtain training data, even in a semi-automatic manner using the SPC visual interface to define the different concepts interactively. As such, the present infrastructure can be used as a training platform for the creation of annotated datasets in a semi-automated manner, with minimal supervision by just giving insight on the domain specialisation. E.g., the ACO (Application Context Ontology) was defined using expert's knowledge, shape grammars and standards in use in Europe, thus presents limitations linked to knowledge standardization. Extending the "standards" and features through machine learning could help to better generalize.

7.2.7 Integrate a natural language processing module

The integration of a natural language module would allow us to extend the possibilities for users to formulate queries that are translated into SQL and SPARQL analogues. Indeed, this would participate into a better automatization and human-machine interface while extending the possible applications to online processing.

7.2.8 Extending the smart modelling process

In the 3D modelling provided method, we consider planar shapes only or manufactured shapes. It would be interesting to extend the method to more complex parametric representations. In the comparison and results analysis, there exist deviations with elements which present a non-planar morphology. Adding a layer of shape deformation processing to best fit shapes is an opening to provide a compact hybrid model. The binning and model fitting step depend on the initial axis orientation's determination. Extending its sturdiness to highly noisy and non-uniform point sampling would extend the flexibility of the workflow. The considerations in this paper and tests were conducted in indoor built environments only. Research to extend it to other scenes and outdoor scenarios is compelling. Also, non-standard shapes are difficult to describe through a knowledge-based approach. This limitation comes from the nature of ontologies to be integrated in standardisation and interoperability workflows. A solution would be to compute robust features through a learning network on the existing set of 3D shapes.

LIST OF TABLES

Table 1. Applications making use of reality-based point cloud data organized by looking at the scale of the studied environment. 30
Table 2. Artefact commonly found in reality-based 3D point clouds. These are obstacles to a broader adoption of point cloud data as the default spatial data type. 33
Table 3. Datasets for benchmarking the SPC Infrastructure 70
Table 4. Example of Basic SQL statements over the SPC infrastructure 73
Table 5. Eigen-based features part of the SF1 feature set 93
Table 6. Geometrical features part of the SF1 feature set 94
Table 7. Relational features of the SF2 feature set for 3D structural connectivity 98
Table 8. The S3DIS dataset and its 6 areas used for testing our methodology 104
Table 9. S3DIS per-area statistics regarding the studied classes. 105
Table 10. Analyses of the impact of feature sets over samples of the S3DIS dataset 106
Table 11. Quantitative CEL segmentation compared to nominal number of elements per class for both a room (Conference room) and area (area 1) 107
Table 12. Global per-class metrics concerning the Area-1 of the S3DIS dataset. SF1 alone and combined SF1SF2 are compared. 107
Table 13. Benchmark results of our semantic segmentation approach against best-performing deep-learning methods 109
Table 14. Benchmark results of our semantic segmentation approach against deep-learning methods without any colour information used. 110
Table 15. Per class metrics for the full S3DIS dataset using our approach 111
Table 16. Overall precision on the full S3DIS dataset against non-machine learning baselines 112
Table 17. Point features computed from the point cloud data after data fusion, before segmentation. 139
Table 18. Segment features computed from the extracted segments 142
Table 19. Example of tessera classification using RDF constraints. 145
Table 20. Segmentation accuracy of tesserae samples. 147
Table 21. Recapitulation of tesserae detection results. 149
Table 22. Segmentation accuracy of the façade of the castle of Jehay over calcareous stones. 149
Table 23. Example of queries over the point cloud. 156
Table 24. Classification of elements based on numerical attributes and topological relations 158
Table 25. Objects to be modelled from the S3-DIS point cloud dataset [133] highlighted in red, identified by a major primitive (Cuboid, Plane, or CSG model assembly) and a category being Normal-Element (NE), Sub-Element (SE) or Aggregated-Element (AE) 185

Table 26. Results of the part-segmentation mechanism against manually annotated Sub-Elements. The precision and recall were obtained by studying True Positives, False Positive and False Negatives. $precision = True\ PositiveTrue\ Positive + False\ Positive, recall = True\ PositiveTrue\ Positive + False\ Negative$ 199
Table 27. Intersection-over-Union on Area 5 of our methodology compared to PointNet [101], SegCloud [160] and SuperPoint Graphs (SPG [161]) v
Table 28. F1-score summary of the 6 areas using our semantic segmentation vii

LIST OF FIGURES

Figure 1. The sensor plays the role of our eyes, the spatial framework becomes a semantic representation, and the scene is tagged familiar using available knowledge 17
Figure 2. 3D point cloud representation vs 3D semantic representation 18
Figure 3. Point cloud datasets using different techniques at different scales and using different methodologies........ 18
Figure 4. The Five Vs of Big Data in the context of point clouds........ 19
Figure 5. The Smart Point Cloud Infrastructure and its modular architecture associated to the thesis chapters. 22
Figure 6. State-of-the art resources given in the thesis. It is gradually organized to first present point clouds specificities and challenges in Chapter 2, then decomposed in Chapter 3 regarding Knowledge Extraction, Integration and Representation further specialized toward automation and structural tasks........ 29
Figure 7. 3D data representation: (a) point cloud; (b) multi-view image; (c) depth map; (d) volumetric; (e) polygonal mesh; (f) parametric model. 36
Figure 8. Classical point cloud workflow from acquisition to delivery of a product for a specific application. 37
Figure 9. Chapter 3: A Smart Point Cloud Data Structure 45
Figure 10. Modular Framework for the creation of an intelligent virtual environment, illustrated over an Indoor Point Cloud. (1) KE; (2) Knowledge Integration; (3) Knowledge Representation; (4) Reasoning from Cognitive Decisions Layer (CDL) 48
Figure 11. Meta-model articulation for the creation of a SPC 57
Figure 12. Level-0 Generalized SPC meta-model (UML). A point cloud constituted of points is block-wise organized through semantic patches. These can be pure spatial conglomerate or retain a coherent semantic relationship between constituting points. Generalizations via different schemes are possible using the generalisation structure to provide additional analysis flexibility........ 58
Figure 13. Example of a basic LOD n-1 Generalization of 3 SemanticPatches from a point cloud with colour attributes only. 59
Figure 14. Level-1 Connection-layer meta-model. It is directly linked to the Level-0: semantic patches constitute ConnectedElements. AggregatedElements and topological notions gives flexibility to the deepness of an element characterization. ConnectedElements can relate to one or multiple spaces defined by their dimensions. These are subsequently divided in subspaces regarding a concept from a domain knowledge characterization, similarly to the world objects (being a specialization of ConnectedElements). 60
Figure 15. Level-2 meta-model example. SeparatorElement and InternalElement are connected to the Level-1 meta-model directly through "WorldObject" and "SubSpace". It is a succession of specialization describing an indoor environment. 62

Figure 16. SPC modular framework. The point cloud is fed to the parsing module (KE) that is directly adjusted regarding the processing module (KE). The different point subsets are extracted and injected in the Point Cloud Database module (KI). This module is central and influenced by the Classification module (KE) and the language and query processing module (KI & KR) which are themselves linked to the Knowledge and agent layers...64
Figure 17. Point cloud parsing methodology: (a) Raw point cloud from TLS; (b) voxelisation at different Octree LoD levels; (c) Segmentation: (d) Voxel-based topology featuring; I Extraction of highly representative points; (f) voxel classification; (g) Connected Element constitution and patch decomposition..........66
Figure 18. Example of Point Cloud in-base topology determinations for relationship extraction (a) CEL block storage; (b) host/guest CEL/ground; (c) Generalization and topology inference: (d) Higher LoD topology inference. 67
Figure 19. Example of voxel space generalization from 3D Point Cloud for A* pathfinding (a) Raw point cloud from TLS; (b) voxelisation at different Octree LoD levels; (c) Segmentation: (d) Voxel-based topology featuring.......................................68
Figure 20. SPC Client/Server Infrastructure implementation...70
Figure 21. Visualisation of four point clouds from the benchmarking datasets and visual impact of the voxel-based featuring and grouping before topology 72
Figure 22. Chapter 4: Knowledge Extraction 81
Figure 23. Voxel-based 3D semantic segmentation. From left to right: Raw point cloud, feature engineering, Connected Elements extraction, Classified point cloud 85
Figure 24. Visual patterns on points from left to right: Not grouped; Proximity criterion; Similarity criterion; Common cluster region; Linear criterion; Parallel criterion: Symmetry criterion..86
Figure 25. Methodological workflow for the constitution of Connected Elements and knowledge-based classification. A point cloud goes through 7 serialized steps (diamonds) to obtain a fully classified dataset (red square)...90
Figure 26. Point Cloud and its extracted voxel structure, where each octree level represents the grid voxels, each subdivided in subsequent 8 voxel children. 91
Figure 27. Feature transfer between octree levels. We note that each non-empty node describes a voxel which can then permit a point-level access for example to compute feature sets (Here, a planar voxel and a corresponding SF1 sample, and a transition voxel and its corresponding SF1 sample)...92
Figure 28. Box plot of primary elements feature variation...95
Figure 29. Direct voxel-to-voxel topology in a 26-connectivity graph. Considered voxel $\mathcal{V}i$ is red, direct connections are either vertex.touch (grey), edge.touch (yellow) or face.touch (orange)..95
Figure 30. Relationship tagging in the voxel-space. (a) represent a mixed relationship $\mathcal{M}r$, (b) a pure vertical relationship $\mathcal{V}r$, (c) a pure horizontal relationship $\mathcal{H}r$...97
Figure 31. S3DIS points within categorized voxels. (a) Full transition voxels, (b) vertical group of points, (c) horizontal group of points, (d) mixed group of points, ...98
Figure 32. Graph representation within a voxel sample of the point cloud. 98

Figure 33. Elements detection and categorization. A point cloud is search for PE, the rest is searched for SE. The remaining from this step is searched for TE, leaving RE. TE permits to extracts graphs through SF2 analysis with PE, SE and RE 99
Figure 34. Edges elements to be decomposed in TE and RE.. 100
Figure 35. different graphs generated on voxel categories. (a) CEL graph, (b) PE graph, (c) SE graph, (d) TE graph, (e) RE graph 101
Figure 36. a) raw point cloud; (b) {PE, SE, TE, RE} groups of voxels; (c) Connected Elements; (d) Classified point cloud.. 103
Figure 37. Normalized score and processing time in function of the defined octree level.. 108
Figure 38. Problematics cases which often include point cloud artefacts such as heavy noise, missing parts, irregular shape geometries, mislabelled data. 109
Figure 39. Confusion matrix of our semantic segmentation approach over the full S3DIS dataset... 112
Figure 40. Results of the semantic segmentation on a room sample. (a) RGB point cloud, (b) Connected Elements, (c) Ground Truth, (d) Results of the semantic segmentation 113
Figure 41. Relative temporal performances of our automatic semantic segmentation workflow.. 114
Figure 42. Chapter 5: Extension of the Smart Point Cloud Infrastructure to the archaeological domain 123
Figure 43. Overview of the methodology developed in this paper. (1) Data pre-processing; (2) Feature computation and segmentation; (3) Object features to ontology; (4) Object classification in regard to ontology 135
Figure 44. The vault of Germigny-des-prés being captured for dense-image matching processing... 137
Figure 45. Point cloud of the church of Germigny-des-Prés. Top View (left) and zone of interest (right). 138
Figure 46. Point cloud from a set of 2D poses reconstructed via dense-image matching using the software ContextCapture v 4.4.6, Bentley Systems, Exton, United States [17]... 138
Figure 47. UML meta-model of the ontology. Tesserae have one or multiple geometries, which are characterized by their regularity (determined by the ontological reasoning framework), and an area. Tesserae also have a temporality (characterized as a time interval, being placed at early Middle Ages or during a restoration at the 19th century) and different materials. These materials retain various properties including light sensitivity... 143
Figure 48. Sub-ontology for the classification of point cloud tessera objects. Blue arrows represent links regarding the tree structure (these are "subClassOf"). The oranges links represent the "hasProperty" relationship that we created to describe the relationship between a Tessera and its properties. It is a simple relationship from domain (Tessera) to Range (Properties)... 144

Figure 49. Detailed ontology for the classification of the mosaic's point cloud. The yellow lines are the links of sub-assumptions, of reasoning. They are in fact links of equivalence between a class and its definition. 144
Figure 50. Connectivity relationship between the CIDOC-CRM and the DSAE ontology to extend interoperability and allow descriptive knowledge for archaeologist to be included. 145
Figure 51. Unique colourisation of a group of golden tesserae (bottom-up view). 146
Figure 52. Zone 1: Classification workflow of tesserae in Zone 1. From left to right: Colour point cloud; abstraction-based segmented point cloud; classified point cloud; 2D geometry over point cloud. 148
Figure 53. DSAE ontology and plugged CIDOC-CRM + CIDOC-CRMba for the detection of objects of interest: calcareous stones, openings and limestone bay frames. 150
Figure 54. Ontology-based classification of the South-South-West façade of the castle of Jehay. From left to right: the façade studied; the result of the segmentation (stones only); the result of the full segmentation; the result of the classification for quasi-planar objects of interest. 151
Figure 55. Server-side data management system. Point clouds go through different processing steps regarding Section 3, and point groups based on the definition of objects regarding domain knowledge are constituted and populate the SPC database. 152
Figure 56. Query result in the WebGL prototype of the optimal viewpoint for faience pieces and gold tesserae in an extracted zone through SQL query. 152
Figure 57. Classification results over the different zones of the mosaics........ 154
Figure 58. Classification and semantization of dark coloured glass. 155
Figure 59. 3D point cloud of the statue of the Egyptian priest Ahmose and his mother, Baket-re. Diorite. C.1490–1400 BC. 18th Dynasty. New Empire. Ny Carlsberg Glyptotek Museum. Copenhagen. Denmark. From left to right: 3D point cloud; feature extraction and segmentation; 3D visualization. 157
Figure 60. CIDOC-CRM ontology for the detection of objects of interest: calcareous stones and openings. Considering the castle of Jehay (B1), it has a building section (BP1) being the studied façade (B2), composed of different elements such as calcareous stones (B3), embrasures (B3) and openings (B4)........ 158
Figure 61. Chapter 6: The Smart Point Cloud reasoning possibilities. 167
Figure 62. Global workflow for modelling indoor point cloud data. Our approach takes as an input semantically-rich point cloud, and process it regarding knowledge-based processes. 179
Figure 63. In red the handled geometry: **(a)** point level; **(b)** patch level; **(c)** object level. 180
Figure 64. The Smart Point Cloud [2] tailored Infrastructure for 3D modelling..... 181
Figure 65. Classes hierarchy in the ACO 182
Figure 66. Framework for the extraction of an object definition from ACO to be readable by other modules of the SPC Infrastructure 183
Figure 67. **(a)** $\mathcal{E}i$ is a Normal-Element; **(b)** $\mathcal{E}i$ is an Aggregated-Element, it goes through part-segmentation; **(c)** $\mathcal{E}i$ is a Sub-Element. 186

Figure 68. Each AE from **(a)** to **(e)** is projected in a voxelized space, studied against voxel count per unit over $e1, e2, e3$ to extract extremums and find patterns that define each Sub-Element as in the ACO. For each object, line 1 is the considered AE, line 2 to 4 illustrates the repartition histogram along $e1, e2, e3$, line 5 the principal cuts extracted, line 6 the results. 190
Figure 69. Symmetric feature characterization for Sub-Elements of chairs (a) to (e). line 1: symmetric planes; line 2: 2D projection featuring; line 3: similarity's feature tag results to other Sub-Elements. 190
Figure 70. The ACO graph-representation of the chair and the relations within Sub-Elements........ 193
Figure 71. The different phases of the primitive fitting for AE. **(a)** point cloud; **(b)** Raw parameters and generation of grid-aligned cuboid; **(c)** Refinement by non-constrained PCA-Analysis; **(d)** Refinement by constrained PCA-Analysis; **(e)** parameters refinement through ACO........ 194
Figure 72. **(a)** Bouding-box $\mathcal{ML}0$; **(b)** KB model $\mathcal{ML}1$; **(c)** Assemblage $\mathcal{ML}2$; **(d)** hybrid model $\mathcal{ML}3$ 195
Figure 73. Results of the shape matching over different datasets **(a)**. The rank RXX represent the score of each candidate. The closer to 0 the best the shape fits the search from **(b)** to **(i)**. 196
Figure 74. **(a)** S3-DIS Point Cloud extract; **(b)** SE modelling (walls, floors, ceillings); **(c)** NE completion and CSG operations; **(d)** 3D global model with point cloud super-imposition........ 197
Figure 75. **(a)** S3-DIS; **(b)** ($\mathcal{ML}0, \mathcal{ML}1$) model; **(c)** ($\mathcal{ML}0, \mathcal{ML}2$) model; **(d)** ($\mathcal{ML}0, \mathcal{ML}3$) model; **(e)** ($\mathcal{ML}1, \mathcal{M}3DM$) model........ 197
Figure 76. 3D modelling results over $\mathcal{ML}1, \mathcal{ML}2, \mathcal{ML}3$ of the DAT dataset. **(a)** 3D representation; **(b)** colour-coded deviations to **(a);** studied repartition in **(c)**; main indicators presented in **(d)**. 200
Figure 77. 3D modelling by triangulation. **(a)** Poisson reconstruction [338]; **(b)** Ball-pivoting approach [339]; **(c)** Marching-Cubes approach [340] 201
Figure 78. 3D modelling accuracy over the hybrid model. **(a)** constitute the results of the 3D modelling through database mining; **(b)** presents the colour coded deviations to the corresponding model **(a)** and studied by repartition in **(c);** gaussian, deviation and size are presented in **(d)**. 202
Figure 79. 3D area-decomposed global model of S3-DIS in ($\mathcal{ML}0, \mathcal{ML}1$). **(a)** constitute the results of the 3D reconstruction modelling; **(b)** presents the colour coded deviations to **(a): (c)** represent the deviation analysis; **(d)** regroups main indicators. 203
Figure 80. 3D modelling accuracy over the hybrid model. **(a)** constitute the results of the 3D modelling through database mining; **(b)** presents the colour coded deviations to the corresponding model **(a)** and studied by repartition in **(c)**, and the main indicator are presented in **(d)**........ 203
Figure 81. Poisson reconstruction of the S3-DIS dataset. **(a)** Global view; **(b)** High sensivity to noise and occlusion; **(c)** Poisson's deviation analysis; **(d)** main indicators 203

Figure 82. Noise and occlusions sensitivity. (a) and (c) shows a Poisson's reconstruction; (b) and (d) shows the $(\mathcal{M}L0, \mathcal{M}L2)$ reconstruction. 204
Figure 83. Relative time processing regarding the main elements of the 3D modelling engine 205
Figure 84. Meta-model articulation for the creation of a SPC 215
Figure 85. The SPC parser module for a block-model structuration........................ 216
Figure 86. The 3D modelling module part of the agent layer to permit a link to different generalizations and data representations 217
Figure 87. The SPCI can manage multiple point clouds and databases. Domain knowledge formally expressed as an ontology can directly b connected via two entry-points(subspaces and Connected Elements) to permit a flexible domain connectivity 218
Figure 88. Server-side data management system. Point clouds go through different processing steps regarding Section 3, and point groups based on the definition of objects regarding domain knowledge are constituted and populate the SPC database. 219
Figure 89. The Smart Point Cloud Infrastructure and its modular architecture associated to the thesis chapters. 220

BIBLIOGRAPHY

1. Treisman, A.M.; Gelade, G. A Feature-Integration of Attention. *Cognitive Psychology* **1980**, *136*, 97–136.
2. Poux, F.; Billen, R. A Smart Point Cloud Infrastructure for intelligent environments. In *International Society for Photogrammetry and Remote Sensing (ISPRS) Book Series*; Lindenbergh, R., Belen, R., Eds.; CRC Press, 2019.
3. Poux, F.; Billen, R. Voxel-based 3D point cloud semantic segmentation: unsupervised geometric and relationship featuring vs deep learning methods. *ISPRS International Journal of Geo-Information* **2019**, 1–34.
4. Poux, F.; Neuville, R.; Van Wersch, L.; Nys, G.-A.; Billen, R. 3D Point Clouds in Archaeology: Advances in Acquisition, Processing and Knowledge Integration Applied to Quasi-Planar Objects. *Geosciences* **2017**, *7*, 96.
5. Poux, F.; Neuville, R.; Nys, G.-A.; Billen, R. 3D Point Cloud Semantic Modelling: Integrated Framework for Indoor Spaces and Furniture. *Remote Sensing* **2018**, *10*, 1412.
6. Hecht, J. Lidar for self-driving cars. *Optics and Photonics News* **2018**, *29*, 26–33.
7. Gerla, M.; Lee, E.-K.; Pau, G.; Lee, U. Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds. In Proceedings of the 2014 IEEE World Forum on Internet of Things (WF-IoT); IEEE, 2014; pp. 241–246.
8. Socher, R. Towards total scene understanding: Classification, annotation and segmentation in an automatic framework. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition; IEEE, 2009; pp. 2036–2043.
9. Pinto, A.M.; Costa, P.; Moreira, A.P.; Rocha, L.F.; Veiga, G.; Moreira, E. Evaluation of Depth Sensors for Robotic Applications. In Proceedings of the 2015 IEEE International Conference on Autonomous Robot Systems and Competitions; IEEE, 2015; pp. 139–143.
10. Nex, F.; Remondino, F. UAV for 3D mapping applications: a review. *Applied Geomatics* **2014**, *6*, 1–15.
11. Slama, C.; Theurer, C.; Henriksen, S. *Manual of photogrammetry*; Slama, C., Theurer, C., Henriksen, S.W., Eds.; 1980; ISBN 0-937294-01-2.
12. Westoby, M.J.; Brasington, J.; Glasser, N.F.; Hambrey, M.J.; Reynolds, J.M. ‘Structure-from-Motion’ photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **2012**, *179*, 300–314.
13. Koenderink, J.J.; van Doorn, A.J. Affine structure from motion. *Journal of the Optical Society of America A* **1991**, *8*, 377.
14. Weinmann, M. *Reconstruction and Analysis of 3D Scenes*; Springer International Publishing: Cham, 2016; ISBN 978-3-319-29244-1.

15. Remondino, F.; Spera, M.G.; Nocerino, E.; Menna, F.; Nex, F. State of the art in high density image matching. *The Photogrammetric Record* **2014**, *29*, 144–166.
16. Pix4D Terrestrial 3D Mapping Using Fisheye and Perspective Sensors – Support Available online: https://support.pix4d.com/hc/en-us/articles/204220385-Scientific-White-Paper-Terrestrial-3D-Mapping-Using-Fisheye-and-Perspective-Sensors#gsc.tab=0 (accessed on Apr 18, 2016).
17. Bentley 3D City Geographic Information System – A Major Step Towards Sustainable Infrastructure Available online: https://www.bentley.com/en/products/product-line/reality-modeling-software/contextcapture.
18. Agisoft LLC Agisoft Photoscan Available online: http://www.agisoft.com/pdf/photoscan_presentation.pdf.
19. nFrames nFrames Available online: http://nframes.com/.
20. Capturing Reality s.r.o RealityCapture Available online: https://www.capturingreality.com.
21. Changchang Wu VisualSFM Available online: http://ccwu.me/vsfm/.
22. Moulon, P. openMVG Available online: http://imagine.enpc.fr/~moulonp/openMVG/.
23. Autodesk Autodesk: 123D catch Available online: http://www.123dapp.com/catch.
24. PROFACTOR GmbH ReconstructMe Available online: http://reconstructme.net/.
25. Eos Systems Inc Photomodeler Available online: http://www.photomodeler.com/index.html.
26. Novel, C.; Keriven, R.; Poux, F.; Graindorge, P. Comparing Aerial Photogrammetry and 3D Laser Scanning Methods for Creating 3D Models of Complex Objects. In Proceedings of the Capturing Reality Forum; Bentley Systems: Salzburg, 2015; p. 15.
27. Sansoni, G.; Carocci, M.; Rodella, R. Calibration and performance evaluation of a 3-D imaging sensor based on the projection of structured light. *IEEE Transactions on Instrumentation and Measurement* **2000**, *49*, 628–636.
28. Galantucci, L.M.; Percoco, G.; Angelelli, G.; Lopez, C.; Introna, F.; Liuzzi, C.; De Donno, A. Reverse engineering techniques applied to a human skull, for CAD 3D reconstruction and physical replication by rapid prototyping. *Journal of medical engineering & technology* **2009**, *30*, 102–11.
29. Wehr, A.; Lohr, U. Airborne laser scanning—an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing* **1999**, *54*, 68–82.
30. Mallet, C.; Bretar, F. Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing* **2009**, *64*, 1–16.
31. Yang, X.; Strahler, A.H.; Schaaf, C.B.; Jupp, D.L.B.; Yao, T.; Zhao, F.; Wang, Z.; Culvenor, D.S.; Newnham, G.J.; Lovell, J.L.; et al. Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-

waveform lidar instrument (Echidna®). *Remote Sensing of Environment* **2013**, *135*, 36–51.
32. Reshetyuk, Y. Self-calibration and direct georeferencing in terrestrial laser scanning, Umea, 2009.
33. Barber, D.; Mills, J.; Smith-Voysey, S. Geometric validation of a ground-based mobile laser scanning system. *ISPRS Journal of Photogrammetry and Remote Sensing* **2008**, *63*, 128–141.
34. James, M.R.; Quinton, J.N. Ultra-rapid topographic surveying for complex environments: the hand-held mobile laser scanner (HMLS). *Earth Surface Processes and Landforms* **2014**, *39*, 138–142.
35. Thomson, C.; Apostolopoulos, G.; Backes, D.; Boehm, J. Mobile Laser Scanning for Indoor Modelling. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **2013**, *II-5/W2*, 289–293.
36. Lauterbach, H.; Borrmann, D.; Heß, R.; Eck, D.; Schilling, K.; Nüchter, A. Evaluation of a Backpack-Mounted 3D Mobile Scanning System. *Remote Sensing* **2015**, *7*, 13753–13781.
37. Berger, M.; Tagliasacchi, A.; Seversky, L.M.; Alliez, P.; Guennebaud, G.; Levine, J.A.; Sharf, A.; Silva, C.T. A Survey of Surface Reconstruction from Point Clouds. *Computer Graphics Forum* **2017**, *36*, 301–329.
38. Dimitrov, A.; Golparvar-Fard, M. Segmentation of building point cloud models including detailed architectural/structural features and MEP systems. *Automation in Construction* **2015**, *51*, 32–45.
39. Thomas, H.; Deschaud, J.-E.; Marcotegui, B.; Goulette, F.; Gall, Y. Le Semantic Classification of 3D Point Clouds with Multiscale Spherical Neighborhoods. In Proceedings of the arXiv (3DV2018); 2018.
40. Ghorpade, V.K.; Checchin, P.; Malaterre, L.; Trassoudaine, L. 3D shape representation with spatial probabilistic distribution of intrinsic shape keypoints. *Eurasip Journal on Advances in Signal Processing* **2017**, *2017*, 52.
41. Shin, D.; Fowlkes, C.C.; Hoiem, D. Pixels, voxels, and views: A study of shape representations for single view 3D object shape prediction. **2018**.
42. Ahmed, E.; Saint, A.; Shabayek, A.E.R.; Cherenkova, K.; Das, R.; Gusev, G.; Aouada, D.; Ottersten, B. Deep Learning Advances on Different 3D Data Representations: A Survey. **2018**.
43. Kazmi, I.K.; You, L.; Zhang, J.J. A survey of 2D and 3D shape descriptors. In Proceedings of the Proceedings - 10th International Conference Computer Graphics, Imaging, and Visualization, CGIV 2013; IEEE, 2013; pp. 1–10.
44. Zhang, L.; João, M.; Ferreira, A. *Survey on 3D shape descriptors*; 2004;
45. Rostami, R.; Bashiri, F.S.; Rostami, B.; Yu, Z. A Survey on Data-Driven 3D Shape Descriptors. In Proceedings of the Computer Graphics Forum; John Wiley & Sons, Ltd (10.1111), 2018.
46. Houshiar, H. Documentation and mapping with 3D point cloud processing, University of Würzburg, 2012.
47. Cao, Z.; Huang, Q.; Ramani, K. 3D Object Classification via Spherical Projections. **2017**.

48. Sinha, A.; Bai, J.; Ramani, K. Deep learning 3D shape surfaces using geometry images. In Proceedings of the ECCV Workshop on Reconstruction and Modeling of Large-Scale 3D Virtual Environments.; 2016; Vol. 9910 LNCS, pp. 223–240.
49. Erdogmus, N.; Marcel, S. Spoofing in 2D face recognition with 3D masks and anti-spoofing with Kinect. In Proceedings of the IEEE 6th International Conference on Biometrics: Theory, Applications and Systems, BTAS 2013; IEEE, 2013; pp. 1–6.
50. Fanelli, G.; Weise, T.; Gall, J.; Gool, L. Van Real Time Head Pose Estimation from Consumer Depth Cameras. **2011**, 101–110.
51. Firman, M. RGBD Datasets: Past, Present and Future. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops; 2016; pp. 661–673.
52. Xiang, Y.; Choi, W.; Lin, Y.; Savarese, S. Data-Driven 3D Voxel Patterns for Object Category Recognition. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* **2015**.
53. Otepka, J.; Ghuffar, S.; Waldhauser, C.; Hochreiter, R.; Pfeifer, N. Georeferenced Point Clouds: A Survey of Features and Point Cloud Management. *ISPRS International Journal of Geo-Information* **2013**, *2*, 1038–1065.
54. van Oosterom, P.; Martinez-Rubi, O.; Ivanova, M.; Horhammer, M.; Geringer, D.; Ravada, S.; Tijssen, T.; Kodde, M.; Gonçalves, R. Massive point cloud data management: Design, implementation and execution of a point cloud benchmark. *Computers & Graphics* **2015**, *49*, 92–125.
55. Cura, R.; Perret, J.; Paparoditis, N. Point cloud server (PCS) : point clouds in-base management and processing. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **2015**, *II-3/W5*, 531–539.
56. Richter, R.; Döllner, J. Concepts and techniques for integration, analysis and visualization of massive 3D point clouds. *Computers, Environment and Urban Systems* **2013**, *45*, 114–124.
57. Yang, J.; Huang, X. A Hybrid Spatial Index for Massive Point Cloud Data Management and Visualization. *Transactions in GIS* **2014**, *18*, 97–108.
58. Tatarchenko, M.; Dosovitskiy, A.; Brox, T. Octree Generating Networks: Efficient Convolutional Architectures for High-resolution 3D Outputs. In Proceedings of the Proceedings of the IEEE International Conference on Computer Vision; 2017; Vol. 2017–Octob, pp. 2107–2115.
59. Zhu, Q.; Gong, J.; Zhang, Y. An efficient 3D R-tree spatial index method for virtual geographic environments. *ISPRS Journal of Photogrammetry and Remote Sensing* **2007**, *62*, 217–224.
60. Gong, J.; Zhu, Q.; Zhong, R.; Zhang, Y.; Xie, X. An Efficient Point Cloud Management Method Based on a 3D R-Tree. *Photogrammetric Engineering & Remote Sensing* **2012**, *78*, 373–381.
61. Scheiblauer, C.; Wimmer, M. Out-of-core selection and editing of huge point clouds. *Computers & Graphics* **2011**, *35*, 342–351.

62. Ben Hmida, H.; Cruz, C.; Boochs, F.; Nicolle, C. From Unstructured 3D Point Clouds to Structured Knowledge - A Semantics Approach. In *Semantics - Advances in Theories and Mathematical Models*; Afzal, M.T., Ed.; InTech: Rijeka, 2012; p. 284 ISBN 978-953-51- 0535-0.
63. Knublauch, H.; Fergerson, R.W.; Noy, N.F.; Musen, M. a. The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications. *International Semantic Web Conference* **2004**, 229–243.
64. Toth, C.; Jóźków, G. Remote sensing platforms and sensors: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing* **2015**, *115*, 22–36.
65. Chen, C.; Zou, X.; Tian, M.; Li, J.; Wu, W.; Song, Y.; Dai, W.; Yang, B. LOW COST MULTI-SENSOR ROBOT LASER SCANNING SYSTEM AND ITS ACCURACY INVESTIGATIONS FOR INDOOR MAPPING APPLICATION. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2017**, *XLII-2/W8*, 83–85.
66. Galindo, C.; Saffiotti, A.; Coradeschi, S.; Buschka, P.; Fernandez-Madrigal, J.A.; Gonzalez, J. Multi-hierarchical semantic maps for mobile robotics. In Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems; IEEE, 2005; pp. 3492–3497.
67. Kaasalainen, S.; Jaakkola, A.; Kaasalainen, M.; Krooks, A.; Kukko, A. Analysis of incidence angle and distance effects on terrestrial laser scanner intensity: Search for correction methods. *Remote Sensing* **2011**, *3*, 2207–2221.
68. Cheng, L.; Chen, S.; Liu, X.; Xu, H.; Wu, Y.; Li, M.; Chen, Y.; Cheng, L.; Chen, S.; Liu, X.; et al. Registration of Laser Scanning Point Clouds: A Review. *Sensors* **2018**, *18*, 1641.
69. Quan, S.; Ma, J.; Hu, F.; Fang, B.; Ma, T. Local voxelized structure for 3D binary feature representation and robust registration of point clouds from low-cost sensors. *Information Sciences* **2018**, *444*, 153–171.
70. Lehtola, V.; Kaartinen, H.; Nüchter, A.; Kaijaluoto, R.; Kukko, A.; Litkey, P.; Honkavaara, E.; Rosnell, T.; Vaaja, M.; Virtanen, J.-P.; et al. Comparison of the Selected State-Of-The-Art 3D Indoor Scanning and Point Cloud Generation Methods. *Remote Sensing* **2017**, *9*, 796.
71. Haala, N. The Landscape of Dense Image Matching Algorithms. **2013**.
72. Pu, S.; Vosselman, G. Knowledge based reconstruction of building models from terrestrial laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing* **2009**, *64*, 575–584.
73. Grilli, E.; Menna, F.; Remondino, F. A review of point clouds segmentation and classification algorithms. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives; 2017; Vol. 42, pp. 339–344.
74. Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Garcia-Rodriguez, J. A Review on Deep Learning Techniques Applied to Semantic Segmentation. **2017**.
75. Arbeiter, G.; Fuchs, S.; Bormann, R.; Fischer, J.; Verl, A. Evaluation of 3D feature descriptors for classification of surface geometries in point clouds. In

Proceedings of the IEEE International Conference on Intelligent Robots and Systems; IEEE, 2012; pp. 1644–1650.
76. Weinmann, M.; Jutzi, B.; Hinz, S.; Mallet, C. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing* **2015**, *105*, 286–304.
77. Safavian, S.R.; Landgrebe, D. *A survey of decision tree classifier methodology*; 1990;
78. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* **2016**, *114*, 24–31.
79. Valentin, J.; Vineet, V.; Cheng, M.-M.; Kim, D.; Shotton, J.; Kohli, P.; Niessner, M.; Criminisi, A.; Izadi, S.; Torr, P. SemanticPaint: Interactive 3D Labeling and Learning at your Fingertips. *ACM Transactions on Graphics* **2015**, *34*.
80. Koppula, H.S.; Anand, A.; Joachims, T.; Saxena, A. Semantic Labeling of 3D Point Clouds for Indoor Scenes. In Proceedings of the Advances in Neural Information Processing Systems; 2011; pp. 244–252.
81. Xiong, X.; Adan, A.; Akinci, B.; Huber, D. Automatic creation of semantically rich 3D building models from laser scanner data. *Automation in Construction* **2013**, *31*, 325–337.
82. Engelmann, F.; Kontogianni, T.; Hermans, A.; Leibe, B. Exploring Spatial Context for 3D Semantic Segmentation of Point Clouds. In Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW); 2018; Vol. 2018–Janua, pp. 716–724.
83. Boulch, A.; Le Saux, B.; Audebert, N. Unstructured point cloud semantic labeling using deep segmentation networks. In Proceedings of the Eurographics Workshop on 3D Object Retrieval; 2017.
84. Yang, B.; Luo, W.; Urtasun, R. PIXOR: Real-time 3D Object Detection from Point Clouds. In Proceedings of the Computer Vision and Pattern Recognition (CVPR 2018); 2018; pp. 7652–7660.
85. Jiang, M.; Wu, Y.; Lu, C. *PointSIFT: A SIFT-like Network Module for 3D Point Cloud Semantic Segmentation*; 2018;
86. Nguyen, C.; Starek, M.J.; Tissot, P.; Gibeaut, J. Unsupervised clustering method for complexity reduction of terrestrial lidar data in marshes. *Remote Sensing* **2018**, *10*, 133.
87. Behl, A.; Paschalidou, D.; Donné, S.; Geiger, A. *PointFlowNet: Learning Representations for 3D Scene Flow Estimation from Point Clouds*; 2018;
88. Engelmann, F.; Kontogianni, T.; Schult, J.; Leibe, B. Know What Your Neighbors Do: 3D Semantic Segmentation of Point Clouds. In Proceedings of the ECCV Workshop 2018; 2018.
89. Li, J.; Chen, B.M.; Lee, G.H. SO-Net: Self-Organizing Network for Point Cloud Analysis. In Proceedings of the Computer Vision and Pattern Recognition (CVPR 2018); 2018.
90. Guerrero, P.; Kleiman, Y.; Ovsjanikov, M.; Mitra, N.J. PCPNet learning local

shape properties from raw point clouds. *Computer Graphics Forum* **2018**, *37*, 75–85.
91. Boulch, A.; Guerry, J.; Le Saux, B.; Audebert, N. SnapNet: 3D point cloud semantic labeling with 2D deep segmentation networks. *Computers & Graphics* **2018**, *71*, 189–198.
92. Su, H.; Jampani, V.; Sun, D.; Maji, S.; Kalogerakis, E.; Yang, M.H.; Kautz, J. SPLATNet: Sparse Lattice Networks for Point Cloud Processing. In Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR 2018); 2018; pp. 2530–2539.
93. Liao, Y.; Donné, S.; Geiger, A. Deep Marching Cubes: Learning Explicit Surface Representations. *Computer Vision and Pattern Recognition (CVPR)* **2018**, 2916–2925.
94. Tangelder, J.W.H.; Veltkamp, R.C. A survey of content based 3D shape retrieval methods. *Multimedia Tools and Applications* **2007**, *39*, 441–471.
95. Nonaka, lkujiro; Takeuchi, H.; Umemoto, K. A theory of organizational knowledge creation. *International Journal of Technology Management* **1996**, *11*.
96. Rusu, R.B.; Marton, Z.C.; Blodow, N.; Dolha, M.; Beetz, M. Towards 3D Point cloud based object maps for household environments. *Robotics and Autonomous Systems* **2008**, *56*, 927–941.
97. Lai, K.; Fox, D. Object Recognition in 3D Point Clouds Using Web Data and Domain Adaptation. *The International Journal of Robotics Research* **2010**, *29*, 1019–1037.
98. Kurgan, L.A.; Musilek, P. A survey of Knowledge Discovery and Data Mining process models. *The Knowledge Engineering Review* **2006**, *21*, 1–24.
99. Fayyad, U.; Wierse, A.; Grinstein, G. *Information visualization in data mining and knowledge discovery*; Morgan Kaufmann, 2002, Ed.; illustrée.; 2002; ISBN 1558606890.
100. Novak, M. *Intelligent Environments: Spatial Aspects of the Information Revolution*; Droege, P., Ed.; P. Droege.; Elsevier, 1997; ISBN 0080534848.
101. Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. PointNet: Deep learning on point sets for 3D classification and segmentation. *30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* **2017**, *2017–Janua*, 77–85.
102. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. *PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space*; 2017;
103. Maturana, D.; Scherer, S. VoxNet: A 3D Convolutional Neural Network for real-time object recognition. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); IEEE, 2015; pp. 922–928.
104. Zhou, Y.; Tuzel, O. VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection. In Proceedings of the arXiv; 2017.
105. Riegler, G.; Ulusoy, A.O.; Geiger, A. OctNet: Learning Deep 3D Representations at High Resolutions. **2016.**
106. Weinmann, M.; Hinz, S.; Weinmann, M. A hybrid semantic point cloud

classification-segmentation framework based on geometric features and semantic rules. *PFG - Journal of Photogrammetry, Remote Sensing and Geoinformation Science* **2017**, *85*, 183–194.
107. Ross, L. Virtual 3D City Models in Urban Land Management - Technologies and Applications, Technischen Universität Berlin, 2010.
108. Poux, F.; Neuville, R.; Hallot, P.; Billen, R. MODEL FOR SEMANTICALLY RICH POINT CLOUD DATA. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2017**, *4*, 107–115.
109. Zlatanova, S.; Rahman, A. Trends in 3D GIS development. *Journal of Geospatial Engineering* **2002**, *4*, 71–80.
110. Liu, X.; Wang, X.; Wright, G.; Cheng, J.; Li, X.; Liu, R. A State-of-the-Art Review on the Integration of Building Information Modeling (BIM) and Geographic Information System (GIS). *ISPRS International Journal of Geo-Information* **2017**, *6*, 53.
111. Kuhn, W. Geospatial Semantics: Why, of What, and How? In *Lecture Notes in Computer Science*; Spaccapietra, S., Zimányi, E., Eds.; Springer: Berlin, Heidelberg, Heidelberg, 2005; Vol. 3534, pp. 1–24.
112. Harvey, F.; Kuhn, W.; Pundt, H.; Bishr, Y.; Riedemann, C. Semantic interoperability: A central issue for sharing geographic information. *The Annals of Regional Science* **1999**, *33*, 213–232.
113. Kolbe, T.H.; Gröger, G.; Plümer, L. CityGML: Interoperable Access to 3D City Models. In *Geo-information for Disaster Management*; Van Oosterom, P., Zlatanova, S., Fendel, E.M., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2005; pp. 883–899 ISBN 3540249885.
114. Biljecki, F.; Stoter, J.; Ledoux, H.; Zlatanova, S.; Çöltekin, A. Applications of 3D City Models: State of the Art Review. *ISPRS International Journal of Geo-Information* **2015**, *4*, 2842–2889.
115. Karim, H.; Abdul Rahman, A.; Boguslawski, P.; Meijers, M.; van Oosterom, P. The Potential of the 3D Dual Half-Edge (DHE) Data Structure for Integrated 2D-Space and Scale Modelling: A Review. In *Advances in 3D Geoinformation*; Abdul-Rahman, A., Ed.; Springer International Publishing, 2017; pp. 477–493 ISBN 978-3-319-25689-4.
116. Dobos, L.; Csabai, I.; Szalai-Gindl, J.M.; Budavári, T.; Szalay, A.S. Point cloud databases. In Proceedings of the Proceedings of the 26th International Conference on Scientific and Statistical Database Management (SSDBM ’14); ACM Press: New York, New York, USA, 2014.
117. Poux, F.; Hallot, P.; Neuville, R.; Billen, R. SMART POINT CLOUD: DEFINITION AND REMAINING CHALLENGES. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2016**, *IV-2*, 119–127.
118. Janowicz, K.; Schade, S.; Bröring, A.; Keßler, C.; Maué, P.; Stasch, C. Semantic Enablement for Spatial Data Infrastructures. *Transactions in GIS* **2010**, *14*, 111–129.
119. Billen, R.; Zaki, C.; Servières, M.; Moreau, G.; Hallot, P. Developing an ontology of space: Application to 3D city modeling. In Proceedings of the Usage,

Usability, and Utility of 3D City Models – European COST Action TU0801; Leduc, T., Moreau, G., Billen, R., Eds.; EDP Sciences: Les Ulis, France, 2012; p. 02007.
120. Gorte, B.; Pfeifer, N. Structuring laser-scanned trees using 3D mathematical morphology. *International Archives of Photogrammetry and Remote* **2004**.
121. Wang, J.; Lindenbergh, R.; Menenti, M. SigVox – A 3D feature matching algorithm for automatic street object recognition in mobile laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* **2017**, *128*, 111–129.
122. Cura, R.; Perret, J.; Paparoditis, N. A scalable and multi-purpose point cloud server (PCS) for easier and faster point cloud data management and processing. *ISPRS Journal of Photogrammetry and Remote Sensing* **2017**, *127*, 39–56.
123. Gong, J.; Ke, S. 3D spatial query implementation method based on R-tree. In Proceedings of the 2011 International Conference on Remote Sensing, Environment and Transportation Engineering; IEEE: Nanjing, 2011; pp. 2828–2831.
124. Cura, R.; Perret, J.; Paparoditis, N. *Implicit LOD for processing, visualisation and classification in Point Cloud Servers*; Saint Mande, 2016;
125. Clementini, E.; Di Felice, P. Approximate topological relations. *International Journal of Approximate Reasoning* **1997**, *16*, 173–204.
126. Zlatanova, S.; Rahman, A.A.; Shi, W. Topological models and frameworks for 3D spatial objects. *Computers & Geosciences* **2004**, *30*, 419–428.
127. Smith, B.; Varzi, A. Fiat and bona fide boundaries. *Philosophical and Phenomenological Research* **2000**, *60*, 401–420.
128. Kakde, H. *Range Searching using Kd Tree*; 2005;
129. Nielsen, J. 10 Usability Heuristics for User Interface Design. *Conference companion on Human factors in computing systems CHI 94* 1995, 152–158.
130. Bui, T.; Lee, J. An agent-based framework for building decision support systems. *Decision Support Systems* **1999**, *25*, 225–237.
131. Goertzel, B.; Pennachin, C. *Artificial General Intelligence*; Goertzel, B., Pennachin, C., Eds.; Springer, 2007; ISBN 9783540237334.
132. Koffka, K. *Principles of Gestalt psychology*; Routledge, 2013;
133. Armeni, I.; Sener, O.; Zamir, A.R.; Jiang, H.; Brilakis, I.; Fischer, M.; Savarese, S. 3D Semantic Parsing of Large-Scale Indoor Spaces. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016; pp. 1534–1543.
134. Ni, H.; Lin, X.; Zhang, J.; Ni, H.; Lin, X.; Zhang, J. Classification of ALS Point Cloud with Improved Point Cloud Segmentation and Random Forests. *Remote Sensing* **2017**, *9*, 288.
135. Bueno, M.; Martínez-Śanchez, J.; Gonźalez-Jorge, H.; Lorenzo, H. Detection of geometric keypoints and its application to point cloud coarse registration. In Proceedings of the IXXIII ISPRS Congress, ISPRS Archives; 2016; Vol. 41, pp. 187–194.

136. Vetrivel, A.; Gerke, M.; Kerle, N.; Nex, F.; Vosselman, G. Disaster damage detection through synergistic use of deep learning and 3D point cloud features derived from very high resolution oblique aerial images, and multiple-kernel-learning. *ISPRS Journal of Photogrammetry and Remote Sensing* **2018**, *140*, 45–59.
137. Blomley, R.; Weinmann, M.; Leitloff, J.; Jutzi, B. Shape distribution features for point cloud analysis - a geometric histogram approach on multiple scales. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **2014**, *II-3*, 9–16.
138. Feng, C.C.; Guo, Z. Automating parameter learning for classifying terrestrial LiDAR point cloud using 2D land cover maps. *Remote Sensing* **2018**, *10*, 1192.
139. Shen, Y.; Feng, C.; Yang, Y.; Tian, D. Mining Point Cloud Local Structures by Kernel Correlation and Graph Pooling. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR); 2018; pp. 4548–4557.
140. Nurunnabi, A.; Belton, D.; West, G. Robust Segmentation for Large Volumes of Laser Scanning Three-Dimensional Point Cloud Data. *IEEE Transactions on Geoscience and Remote Sensing* **2016**, *54*, 4790–4805.
141. Lawin, F.J.; Danelljan, M.; Tosteberg, P.; Bhat, G.; Khan, F.S.; Felsberg, M. Deep projective 3D semantic segmentation. *Computer Analysis of Images and Patterns (CAIP)* **2017**, *10424 LNCS*, 95–107.
142. Mahmoudabadi, H.; Shoaf, T.; Olsen, M.J. Superpixel clustering and planar fit segmentation of 3D LIDAR point clouds. In Proceedings of the 4th International Conference on Computing for Geospatial Research and Application, (COM.Geo); 2013; pp. 1–7.
143. Ioannou, Y.; Taati, B.; Harrap, R.; Greenspan, M. Difference of normals as a multi-scale operator in unorganized point clouds. In Proceedings of the 2012 Second International Conference on 3D Imaging, Modeling,Processing, Visualization and Transmission (3DIMPVT); IEEE: Zurich, 2012; pp. 501–508.
144. Vosselman, G.; Gorte, B.G.H.; Sithole, G.; Rabbani, T. Recognising structure in laser scanner point clouds. In Proceedings of the ISPRS Archives; 2003; Vol. 46, pp. 33–38.
145. Song, T.; Xi, F.; Guo, S.; Ming, Z.; Lin, Y. A comparison study of algorithms for surface normal determination based on point cloud data. *Precision Engineering* **2015**, *39*, 47–55.
146. Weber, C.; Hahmann, S.; Hagen, H. Sharp feature detection in point clouds. In Proceedings of the SMI 2010 - International Conference on Shape Modeling and Applications, Proceedings; IEEE, 2010; pp. 175–186.
147. Ni, H.; Lin, X.; Ning, X.; Zhang, J. Edge Detection and Feature Line Tracing in 3D-Point Clouds by Analyzing Geometric Properties of Neighborhoods. *Remote Sensing* **2016**, *8*, 710.
148. Schnabel, R.; Wahl, R.; Klein, R. Efficient RANSAC for Point Cloud Shape Detection. *Computer Graphics Forum* **2007**, *26*, 214–226.

149. Fischler, M.A.; Bolles, R.C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* **1981**, *24*, 381–395.
150. Hackel, T.; Wegner, J.D.; Schindler, K. Joint classification and contour extraction of large 3D point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* **2017**, *130*, 231–245.
151. Son, H.; Kim, C. Semantic as-built 3D modeling of structural elements of buildings based on local concavity and convexity. *Advanced Engineering Informatics* **2017**, *34*, 114–124.
152. Liu, Y.-S.; Ramani, K. Robust principal axes determination for point-based shapes using least median of squares. *Computer aided design* **2009**, *41*, 293–305.
153. Xu, Y.; Tuttas, S.; Hoegner, L.; Stilla, U. Voxel-based segmentation of 3D point clouds from construction sites using a probabilistic connectivity model. *Pattern Recognition Letters* **2018**, *102*.
154. Xu, Y.; Hoegner, L.; Tuttas, S.; Stilla, U. Voxel- and Graph-Based Point Cloud Segmentation of 3D Scenes Using Perceptual Grouping Laws. In Proceedings of the ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences; 2017; Vol. IV-1/W1, pp. 43–50.
155. Zhu, Q.; Li, Y.; Hu, H.; Wu, B. Robust point cloud classification based on multi-level semantic relationships for urban scenes. *ISPRS Journal of Photogrammetry and Remote Sensing* **2017**, *129*, 86–102.
156. Wang, Y.; Cheng, L.; Chen, Y.; Wu, Y.; Li, M. Building point detection from vehicle-borne LiDAR data based on voxel group and horizontal hollow analysis. *Remote Sensing* **2016**, *8*, 419.
157. Ben-Shabat, Y.; Avraham, T.; Lindenbaum, M.; Fischer, A. Graph based over-segmentation methods for 3D point clouds. *Computer Vision and Image Understanding* **2018**, *174*, 12–23.
158. Ben-Shabat, Y.; Lindenbaum, M.; Fischer, A. *3D Point Cloud Classification and Segmentation using 3D Modified Fisher Vector Representation for Convolutional Neural Networks*; 2017;
159. Rusu, R.B.; Blodow, N.; Beetz, M. Fast Point Feature Histograms (FPFH) for 3D registration. In Proceedings of the 2009 IEEE International Conference on Robotics and Automation; IEEE, 2009; pp. 3212–3217.
160. Tchapmi, L.P.; Choy, C.B.; Armeni, I.; Gwak, J.; Savarese, S. SEGCloud: Semantic Segmentation of 3D Point Clouds. In Proceedings of the International Conference of 3D Vision (3DV 2017); 2017.
161. Landrieu, L.; Simonovsky, M. Large-Scale Point Cloud Semantic Segmentation with Superpoint Graphs. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* **2018**, 4558–4567.
162. Truong-Hong, L.; Laefer, D.F.; Hinks, T.; Carr, H. Flying Voxel Method with Delaunay Triangulation Criterion for Façade/Feature Detection for Computation. *Journal of Computing in Civil Engineering* **2012**, *26*, 691–707.
163. Nourian, P.; Gonçalves, R.; Zlatanova, S.; Ohori, K.A.; Vu Vo, A. Voxelization

algorithms for geospatial applications: Computational methods for voxelating spatial datasets of 3D city models containing 3D surface, curve and point data models. *MethodsX* **2016**, *3*, 69–86.
164. De Lathauwer, L.; De Moor, B.; Vandewalle, J. A Multilinear Singular Value Decomposition. *SIAM Journal on Matrix Analysis and Applications* **2003**, *21*, 1253–1278.
165. He, L.; Ren, X.; Gao, Q.; Zhao, X.; Yao, B.; Chao, Y. The connected-component labeling problem: A review of state-of-the-art algorithms. *Pattern Recognition* **2017**, *70*, 25–43.
166. Krijnen, T.; Beetz, J. An IFC schema extension and binary serialization format to efficiently integrate point cloud data into building models. *Advanced Engineering Informatics* **2017**, *33*, 473–490.
167. Belongie, S.; Malik, J.; Puzicha, J. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2002**, *24*, 509–522.
168. Liu, S.; Xie, S.; Chen, Z.; Tu, Z. Attentional ShapeContextNet for Point Cloud Recognition. *IEEE Conference on Computer Vision and Pattern Recognition* **2018**, 4606–4615.
169. Leute, U. *Archaeometry : an introduction to physical methods in archaeology and the history of art*; VCH Verlagsgesellschaft mbH, 1987; ISBN 3527266313.
170. Koller, D.; Frischer, B.; Humphreys, G. Research challenges for digital archives of 3D cultural heritage models. *Journal on Computing and Cultural Heritage* **2009**, *2*, 1–17.
171. Joukowsky, M. A complete manual of field archaeology : tools and techniques of field work for archaeologists. In *Field archaeology, tools and techniques of field work for archaeologists*; 1980; p. x, 630 p : ISBN 0-13-162164-5.
172. Remondino, F. Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning. *Remote Sensing* **2011**, *3*, 1104–1138.
173. Tucci, G.; Bonora, V. GEOMATICS & RESTORATION – Conservation of Cultural Heritage in the Digital Era. In Proceedings of the ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences; 2017; Vol. XLII-5/W1, pp. 1–1.
174. Patias, P. Cultural heritage documentation. In Proceedings of the International summer school "Digital recording and 3D Modeling"; 2006; pp. 24–29.
175. Binding, C.; May, K.; Tudhope, D. Semantic interoperability in archaeological datasets: Data mapping and extraction via the CIDOC CRM. In Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer Berlin Heidelberg: Berlin, Heidelberg, 2008; Vol. 5173 LNCS, pp. 280–290.
176. Lüscher, P.; Weibel, R.; Burghardt, D. Integrating ontological modelling and Bayesian inference for pattern classification in topographic vector data. *Computers, Environment and Urban Systems* **2009**, *33*, 363–374.

177. Poux, F.; Neuville, R.; Hallot, P.; Van Wersch, L.; Jancsó, A.L.; Billen, R.; Luczfalvy Jancsó, A.; Billen, R. DIGITAL INVESTIGATIONS OF AN ARCHAEOLOGICAL SMART POINT CLOUD: A REAL TIME WEB-BASED PLATFORM TO MANAGE THE VISUALISATION OF SEMANTICAL QUERIES. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2017**, *XLII-5/W1*, 581–588.
178. Pieraccini, M.; Guidi, G.; Atzeni, C. 3D digitizing of cultural heritage. *Journal of Cultural Heritage* **2001**, *2*, 63–70.
179. Remondino, F.; Rizzi, A. Reality-based 3D documentation of natural and cultural heritage sites—techniques, problems, and examples. *Applied Geomatics* **2010**, *2*, 85–100.
180. Hartmann-Virnich, A. Transcrire l'analyse fine du bâti: un plaidoyer pour le relevé manuel dans l'archéologie monumentale. In *la pierre et sa mise en oeuvre dans l'art médiéval : mélanges d'Histoire de l'art offerts à Eliane Vergnolle*; Gallet, Y., Ed.; Brepols Publishers: Turnhout, 2011; pp. 191–202 ISBN 978-2-503-53563-0.
181. Lambers, K.; Remondino, F. Optical 3D Measurement Techniques in Archaeology: Recent Developments and Applications. In Proceedings of the Layers of perception. Proceedings of the 35th International conference on Computer Applications and Quantitative Methods in Archeology (CAA); Posluschny, A., Ed.; Berlin, 2007; pp. 27–35.
182. Boto-varela, G.; Hartmann-virnich, A.; Nussbaum, N.; Reveyron, N.; Boto-varela, P.D.D.G.; Hartmann-virnich, A.; Nussbaum, N. Archéologie du bâti : du mètre au laser. *Perspectives* **2012**, 329–346.
183. Forte, M.; Dell'Unto, N.; Di Giuseppantonio Di Franco, P.; Galeazzi, F.; Liuzza, C.; Pescarin, S. The virtual museum of the Western Han Dynasty: 3D documentation and interpretation. In Proceedings of the Space, Time, Place, Third International Conference on Remote Sensing in Archaeology; 2010; Vol. 2118, pp. 195–199.
184. Pavlidis, G.; Koutsoudis, A.; Arnaoutoglou, F.; Tsioukas, V.; Chamzas, C. Methods for 3D digitization of Cultural Heritage. *Journal of Cultural Heritage* **2007**, *8*, 93–98.
185. Chase, A.F.; Chase, D.Z.; Weishampel, J.F.; Drake, J.B.; Shrestha, R.L.; Slatton, K.C.; Awe, J.J.; Carter, W.E. Airborne LiDAR, archaeology, and the ancient Maya landscape at Caracol, Belize. *Journal of Archaeological Science* **2011**, *38*, 387–398.
186. Chase, A.F.; Chase, D.Z.; Fisher, C.T.; Leisz, S.J.; Weishampel, J.F. Geospatial revolution and remote sensing LiDAR in Mesoamerican archaeology. In Proceedings of the Proceedings of the National Academy of Sciences; National Academy of Sciences, 2012; Vol. 109, pp. 12916–12921.
187. Evans, D.H.; Fletcher, R.J.; Pottier, C.; Chevance, J.-B.; Soutif, D.; Tan, B.S.; Im, S.; Ea, D.; Tin, T.; Kim, S.; et al. Uncovering archaeological landscapes at Angkor using lidar. In Proceedings of the Proceedings of the National Academy of Sciences; National Academy of Sciences, 2013; Vol. 110, pp.

12595–12600.
188. Coluzzi, R.; Lanorte, A.; Lasaponara, R. On the LiDAR contribution for landscape archaeology and palaeoenvironmental studies: The case study of Bosco dell'Incoronata (Southern Italy). *Advances in Geosciences* **2010**, *24*, 125–132.
189. Lerma, J.L.; Navarro, S.; Cabrelles, M.; Villaverde, V. Terrestrial laser scanning and close range photogrammetry for 3D archaeological documentation: the Upper Palaeolithic Cave of Parpallo as a case study. *Journal of Archaeological Science* **2010**, *37*, 499–507.
190. Armesto-González, J.; Riveiro-Rodríguez, B.; González-Aguilera, D.; Rivas-Brea, M.T. Terrestrial laser scanning intensity data applied to damage detection for historical buildings. *Journal of Archaeological Science* **2010**, *37*, 3037–3047.
191. Entwistle, J.A.; McCaffrey, K.J.W.; Abrahams, P.W. Three-dimensional (3D) visualisation: the application of terrestrial laser scanning in the investigation of historical Scottish farming townships. *Journal of Archaeological Science* **2009**, *36*, 860–866.
192. Grussenmeyer, P.; Landes, T.; Voegtle, T.; Ringle, K. Comparison methods of terrestrial laser scanning, photogrammetry and tacheometry data for recording of cultural heritage buildings. In Proceedings of the The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.; Beijing, 2008; Vol. Vol. XXXVI, pp. 213–218.
193. Gonzalez-Aguilera, D.; Muñoz-Nieto, A.; Rodriguez-Gonzalvez, P.; Menéndez, M. New tools for rock art modelling: Automated sensor integration in Pindal Cave. *Journal of Archaeological Science* **2011**, *38*, 120–128.
194. Galindo Domínguez, R.E.; Bandy, W.L.; Mortera Gutiérrez, C.A.; Ortega Ramírez, J. Geophysical-Archaeological Survey in Lake Tequesquitengo, Morelos, Mexico. *Geofísica Internacional* **2013**, *52*, 261–275.
195. Wenzel, K.; Rothermel, M.; Haala, N.; Fritsch, D. SURE – The ifp Software for Dense Image Matching. In Proceedings of the Photogrammetric Week 2013; 2013; pp. 59–70.
196. Guarnieri, A.; Remondino, F.; Vettore, A. Digital photogrammetry and TLS data fusion applied to cultural heritage 3D modeling. In Proceedings of the International Archives of Photogrammetry and Remote Sensing; Maas, H.-G., Schneider, D., Eds.; ISPRS: Dresden, Germany, 2006; Vol. XXXVI.
197. Nunez, M.A.; Buill, F.; Edo, M. 3D model of the Can Sadurni cave. *Journal of Archaeological Science* **2013**, *40*, 4420–4428.
198. Leberl, F.; Irschara, A.; Pock, T.; Meixner, P.; Gruber, M.; Scholz, S.; Wiechert, A. Point Clouds: Lidar versus 3D Vision. *Photogrammetric Engineering and Remote Sensing* **2010**, *76*, 1123–1134.
199. Al-kheder, S.; Al-shawabkeh, Y.; Haala, N. Developing a documentation system for desert palaces in Jordan using 3D laser scanning and digital photogrammetry. *Journal of Archaeological Science* **2009**, *36*, 537–546.
200. Poux, F.; Neuville, R.; Hallot, P.; Billen, R. Point clouds as an efficient

multiscale layered spatial representation. In Proceedings of the Eurographics Workshop on Urban Data Modelling and Visualisation; Vincent, T., Biljecki, F., Eds.; The Eurographics Association, 2016.
201. Koch, M.; Kaehler, M. Combining 3D laser-Scanning and close-range Photogrammetry - An approach to Exploit the Strength of Both methods. In Proceedings of the Computer Applications to Archaeology; Williamsburg, 2009; pp. 1–7.
202. Stanco, F.; Battiato, S.; Gallo, G. *Digital Imaging for Cultural Heritage Preservation: Analysis, Restoration, and Reconstruction of Ancient Artworks*; 2011; ISBN 1439821739.
203. Billen, R. Nouvelle perception de la spatialité des objets et de leurs relations. Développement d'une modélisation tridimensionnelle de l'information spatiale, University of Liège, 2002.
204. Agapiou, A.; Lysandrou, V. Remote sensing archaeology: Tracking and mapping evolution in European scientific literature from 1999 to 2015. *Journal of Archaeological Science: Reports* **2015**, *4*, 192–200.
205. Rollier-Hanselmann, J.; Petty, Z.; Mazuir, A.; Faucher, S.; Coulais, J.-F. Développement d'un SIG pour la ville médiévale de Cluny. *Archeologia e Calcolatori* **2014**, *Supplement*, 164–179.
206. Katsianis, M.; Tsipidis, S.; Kotsakis, K.; Kousoulakou, A. A 3D digital workflow for archaeological intra-site research using GIS. *Journal of Archaeological Science* **2008**, *35*, 655–667.
207. Apollonio, F.I.; Gaiani, M.; Benedetti, B. 3D reality-based artefact models for the management of archaeological sites using 3D Gis: A framework starting from the case study of the Pompeii Archaeological area. *Journal of Archaeological Science* **2012**, *39*, 1271–1287.
208. Galeazzi, F.; Callieri, M.; Dellepiane, M.; Charno, M.; Richards, J.; Scopigno, R. Web-based visualization for 3D data in archaeology: The ADS 3D viewer. *Journal of Archaeological Science: Reports* **2016**, *9*, 1–11.
209. Stefani, C.; Brunetaud, X.; Janvier-Badosa, S.; Beck, K.; De Luca, L.; Al-Mukhtar, M. Developing a toolkit for mapping and displaying stone alteration on a web-based documentation platform. *Journal of Cultural Heritage* **2014**, *15*, 1–9.
210. De Reu, J.; Plets, G.; Verhoeven, G.; De Smedt, P.; Bats, M.; Cherretté, B.; De Maeyer, W.; Deconynck, J.; Herremans, D.; Laloo, P.; et al. Towards a three-dimensional cost-effective registration of the archaeological heritage. *Journal of Archaeological Science* **2013**, *40*, 1108–1121.
211. Scianna, A.; Gristina, S.; Paliaga, S. Experimental BIM applications in archaeology: A work-flow. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; 2014; Vol. 8740, pp. 490–498 ISBN 978-3-319-13695-0; 978-3-319-13694-3.
212. Campanaro, D.M.; Landeschi, G.; Dell'Unto, N.; Leander Touati, A.M. 3D GIS for cultural heritage restoration: A "white box" workflow. *Journal of Cultural Heritage* **2016**, *18*, 321–332.

213. Soler, F.; Melero, F.J.; Luzon, M.V. A complete 3D information system for cultural heritage documentation. *Journal of Cultural Heritage* **2017**, *23*, 49–57.
214. Lieberwirth, U. 3D GIS Voxel-Based Model Building in Archaeology. In Proceedings of the Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA); 2007; Vol. 2, pp. 1–8.
215. Orengo, H.A. Combining terrestrial stereophotogrammetry, DGPS and GIS-based 3D voxel modelling in the volumetric recording of archaeological features. *ISPRS Journal of Photogrammetry and Remote Sensing* **2013**, *76*, 49–55.
216. Moscati, A.; Lombardo, J.; Losciale, L.V.; De Luca, L. Visual browsing of semantic descriptions of heritage buildings morphology. In Proceedings of the Digital Media and its Applications in Cultural Heritage (DMACH) 2011; Amman, 2011; pp. 1–16.
217. Janowicz, K. Observation-Driven Geo-Ontology Engineering. *Transactions in GIS* **2012**, *16*, 351–374.
218. Klein, L.A. *Sensor and Data Fusion: A Tool for Information Assessment and Decision Making*; SPIE Press, Ed.; illustrée.; 2004; ISBN 0819454354.
219. Petrie, G. Systematic oblique ae using multiple digitarial photography i frame cameras. *Photogrammetric Engineering and Remote Sensing* **2009**, *75*, 102–107.
220. Van Wersch, L.; Kronz, A.; Simon, K.; Hocquet, F.-P.; Strivay, D. *Matériaux des mosaïques de Germigny-des-Prés.*; Germigny-des-prés;
221. Poux, F.; Neuville, R.; Billen, R. POINT CLOUD CLASSIFICATION OF TESSERAE FROM TERRESTRIAL LASER DATA COMBINED WITH DENSE IMAGE MATCHING FOR ARCHAEOLOGICAL INFORMATION EXTRACTION. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **2017**, *IV-2/W2*, 203–211.
222. Ben Hmida, H.; Boochs, F.; Cruz, C.; Nicolle, C. Knowledge Base Approach for 3D Objects Detection in Point Clouds Using 3D Processing and Specialists Knowledge. *International Journal on Advances in Intelligent Systems* **2012**, *5*, 1–14.
223. Lin, H.; Gao, J.; Zhou, Y.; Lu, G.; Ye, M.; Zhang, C.; Liu, L.; Yang, R. Semantic decomposition and reconstruction of residential scenes from LiDAR data. *ACM Transactions on Graphics (TOG)* **2013**, *32*, 1.
224. Ochmann, S.; Vock, R.; Wessel, R.; Klein, R. Automatic reconstruction of parametric building models from indoor point clouds. *Computers & Graphics* **2016**, *54*, 94–103.
225. Chaperon, T.; Goulette, F. Extracting cylinders in full 3D data using a random sampling method and the Gaussian image. In Proceedings of the Computer Vision and Image Understanding; 2001; pp. 35–42.
226. Nurunnabi, A.; Belton, D.; West, G. Robust Segmentation in Laser Scanning 3D Point Cloud Data. In Proceedings of the 2012 International Conference on

Digital Image Computing Techniques and Applications; IEEE: Fremantle, WA, 2012; pp. 1–8.

227. Rusu, R.B.; Blodow, N. Close-range scene segmentation and reconstruction of 3D point cloud maps for mobile manipulation in domestic environments. *Intelligent Robots and Systems IROS IEEE/RSJ International Conference on* **2009**, 1–6.

228. Samet, H.; Tamminen, M. Efficient component labeling of images of arbitrary dimension represented by linear bintrees. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **1988**, *10*, 579–586.

229. Girardeau-Montaut, D.; Roux, M.; Thibault, G. Change Detection on points cloud data acquired with a ground laser scanner. In Proceedings of the ISPRS Workshop Laser Scanning; ISPRS: Enschede, 2005.

230. Girardeau-Montaut, D. Détection de changement sur des données géométriques tridimensionnelles, Télécom ParisTech, 2006.

231. Douillard, B.; Underwood, J. On the segmentation of 3D LIDAR point clouds. In Proceedings of the International Conference on Robotics and Automation; IEEE: Shangai, 2011; pp. 2798–2805.

232. Aijazi, A.; Checchin, P.; Trassoudaine, L. Segmentation Based Classification of 3D Urban Point Clouds: A Super-Voxel Based Approach with Evaluation. *Remote Sensing* **2013**, *5*, 1624–1650.

233. Sapkota, P. Segmentation of coloured point cloud data, International Institute for Geo-Information science and earth observation, 2008.

234. Poux, F.; Hallot, P.; Jonlet, B.; Carre, C.; Billen, R. Segmentation semi-automatique pour le traitement de données 3D denses: application au patrimoine architectural. *XYZ: la Revue de l' association française de topographie* **2014**, *141*, 69–75.

235. Library, P.C. Statistical Outlier Fliter Available online: http://pointclouds.org/documentation/tutorials/statistical_outlier.php.

236. Kaufman, A.; Yagel, R.; Cohen, D. Modeling in Volume Graphics. In *Modeling in Computer Graphics*; Springer Berlin Heidelberg: Berlin, Heidelberg, 1993; pp. 441–454.

237. Brinkhoff, T.; Kriegel, H.; Schneider, R.; Braun, A. Measuring the Complexity of Polygonal Objects. In Proceedings of the Proceedings of the 3rd ACM International Workshop on Advances in Geographic Information Systems; Baltimore, Maryland, Maryland, 1995; p. 109.

238. Arp, R.; Smith, B.; Spear, A.D. *Building Ontologies with Basic Formal Ontology*; The MIT Press, 2015; ISBN 9788578110796.

239. Neuville, R.; Poux, F.; Hallot, P.; Billen, R. Towards a normalized 3D geovizualisation: The viewpoint management. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **2016**, *IV-2/W1*, 179–186.

240. Ronzino, P.; Niccolucci, F.; Felicetti, A.; Doerr, M. CRMba a CRM extension for the documentation of standing buildings. *International Journal on Digital Libraries* **2016**, *17*, 71–78.

241. Garstka, J.; Peters, G. Learning Strategies to Select Point Cloud Descriptors for 3D Object Classification: A Proposal. In Proceedings of the Eurographics 2014; Paulin, M., Dachsbacher, C., Eds.; The Eurographics Association: Strasbourg, France, 2014.
242. Franken, M.; van Gemert, J.C. Automatic Egyptian hieroglyph recognition by retrieving images as texts. In Proceedings of the Proceedings of the 21st ACM international conference on Multimedia - MM '13; ACM Press: New York, New York, USA, 2013; pp. 765–768.
243. Hmida, H.B.. Ben; Cruz, C..; Boochs, F..; Nicolle, C.. From 9-IM topological operators to qualitative spatial relations using 3D selective Nef complexes and logic rules for bodies. In Proceedings of the KEOD 2012 - Proceedings of the International Conference on Knowledge Engineering and Ontology Development; 2012; pp. 208–213.
244. Forte, M. Cyber-archaeology: an eco-approach to the virtual reconstruction of the past. In Proceedings of the International Symposium on "Information and Communication Technologies in Cultural Heritage"; 2008; pp. 91–106.
245. Remondino, F. From point cloud to surface. In Proceedings of the International Workshop on Visualization and Animation of Reality-based 3D Models; ETH, Swiss Federal Institute of Technology Zurich, I. of G. and P., Ed.; ETH, Swiss Federal Institute of Technology Zurich, Institute of Geodesy and Photogrammetry: Tarasp-Vulpera, 2003; Vol. XXXIV.
246. Hussain, M. Efficient simplification methods for generating high quality LODs of 3D meshes. *Journal of Computer Science and Technology* **2009**, *24*, 604–608.
247. Gordon, S.J.; Lichti, D.D.; Stewart, M.P.; Franke, J. Modelling Point Clouds for Precise Structural Deformation Measurement. *International Archives of Photogrammetry and Remote Sensing* **2004**, *XXXV*.
248. Bosche, F.; Haas, C.T.; Akinci, B. Automated Recognition of 3D CAD Objects in Site Laser Scans for Project 3D Status Visualization and Performance Control. *Journal of Computing in Civil Engineering* **2009**, *23*, 311–318.
249. Bueno, M.; Bosché, F.; González-Jorge, H.; Martínez-Sánchez, J.; Arias, P. 4-Plane congruent sets for automatic registration of as-is 3D point clouds with 3D BIM models. *Automation in Construction* **2018**, *89*, 120–134.
250. Labsch, M.; Kr, K.; Grash, F.; Kohl, S. Content Creation for a 3D Game with Maya and Unity 3D. In Proceedings of the The 15th Central European Seminar on Computer Graphics; Viničn, Slovakia, 2011.
251. Braun, E. System and method for cross-platform and cross-game virtual asset creation and management 2012.
252. Dietenbeck, T.; Torkhani, F.; Othmani, A.; Attene, M.; Favreau, J.-M.M. Multi-layer ontologies for integrated 3D shape segmentation and annotation. In *Studies in Computational Intelligence*; Springer, Cham, 2017; Vol. 665, pp. 181–206.
253. Cao, Y.; Lu, F. Three-Dimensional Modeling for Buildings Evacuation Management, 2012.

254. Hinks, T.; Laefer, D.F.; Carr, H. Viability Considerations for the Application of Aerial LiDAR for Infrastructure Documentation and Monitoring. In Proceedings of the Proceedings of the Transportation Research Board 87th Annual Meeting; Washington, DC, USA, 2008.
255. Adjiski, V.; Mirakovski, D.; Despodov, Z.; Mijalkovski, S. Simulation and optimization of evacuation routes in case of fire in underground mines. *Journal of Sustainable Mining* **2015**, *14*, 133–143.
256. Lee, J.; Zlatanova, S. A 3D data model and topological analyses for emergency response in urban areas. *Geospatial Information Technology for Emergency Response* **2008**, 143–167.
257. Pu, S.; Zlatanova, S. Evacuation route calculation of inner buildings. In *Geo-information for Disaster Management*; 2005; pp. 1143–1161 ISBN 3540249885.
258. Daniel, S.; Harrap, R.; Power, M. Getting Into Position: Serious Gaming in Geomatics. In Proceedings of the World Conference on E-Learning in …; Association for the Advancement of Computing in Education, 2009; Vol. 2009, pp. 213–219.
259. Conconi, A.; Ganchev, T.; Kocsis, O.; Papadopoulos, G.; Fernández-Aranda, F.; Jiménez-Murcia, S. PlayMancer: A serious gaming 3D environment. In Proceedings of the Proceedings - 4th International Conference on Automated Solutions for Cross Media Content and Multi-Channel Distribution, Axmedis 2008; IEEE, 2008; pp. 111–117.
260. Escarieu, F.; Pousset, Y.; Aveneau, L.; Vauzelle, R. Outdoor and indoor channel characterization by a 3D simulation software. *12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications. PIMRC 2001. Proceedings (Cat. No.01TH8598)* **2001**, *1*, 105–111.
261. Cichon, D.J.; Zwick, T.; Lahteenmaki, J. Ray optical indoor modeling in multi-floored buildings: simulations and measurements. In Proceedings of the Antennas and Propagation Society International Symposium, 1995. AP-S. Digest; IEEE, 1995; Vol. 1, pp. 522–525 vol.1.
262. Korkalainen, M.; Sallinen, M.; Kärkkäinen, N.; Tukeva, P. Survey of wireless sensor networks simulation tools for demanding applications. In Proceedings of the Proceedings of the 5th International Conference on Networking and Services, ICNS 2009; IEEE, 2009; pp. 102–106.
263. Lauer, A.; Bahr, A.; Wolff, I. FDTD simulations of indoor propagation. In Proceedings of the Proceedings of IEEE Vehicular Technology Conference (VTC); IEEE, 1994; pp. 883–886.
264. Xiao, J.; Furukawa, Y. Reconstructing the World's Museums. *International Journal of Computer Vision* **2014**, *110*, 243–258.
265. Zhou, C.; Liu, X.; Xue, F.; Bo, H.; Li, K. Research on static service BOM transformation for complex products. *Advanced Engineering Informatics* **2018**, *36*, 146–162.
266. Liu, R.; Asari, V.K. 3D indoor scene reconstruction and change detection for robotic sensing and navigation. In Proceedings of the Proceedings of SPIE -

The International Society for Optical Engineering; Agaian, S.S., Jassim, S.A., Eds.; International Society for Optics and Photonics, 2017; Vol. 10221, p. 102210D.
267. Prieto, S.A.; Quintana, B.; Adán, A.; Vázquez, A.S. As-is building-structure reconstruction from a probabilistic next best scan approach. *Robotics and Autonomous Systems* **2017**, *94*, 186–207.
268. Shaojie Shen, N.M.V.K. Autonomous Indoor 3D Exploration with a Micro-Aerial Vehicle. *2012 IEEE International Conference on Robotics and Automation* **2012**, 1–7.
269. Zhu, Y.; Mottaghi, R.; Kolve, E.; Lim, J.J.; Gupta, A.; Fei-Fei, L.; Farhadi, A. Target-driven Visual Navigation in Indoor Scenes using Deep Reinforcement Learning. *2017 IEEE International Conference on Robotics and Automation (ICRA)* **2016**, 3357–3364.
270. Chow, J.C.K.; Peter, M.; Scaioni, M.; Al-durgham, M. Indoor Tracking , Mapping , and Navigation : Algorithms , Technologies , and Applications. *Journal of Sensors* **2018**, *2018*, 3–6.
271. Balado, J.; Díaz-Vilariño, L.; Arias, P.; Garrido, I. Point clouds to indoor/outdoor accessibility diagnosis. In Proceedings of the ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences; ISPRS, 2017; Vol. 4, pp. 287–293.
272. Hayes-Roth, F.; Waterman Donald, A.; Lenat Douglas, B. *Building Expert System*; 1983;
273. Zhirong Wu; Song, S.; Khosla, A.; Fisher Yu; Linguang Zhang; Xiaoou Tang; Xiao, J. 3D ShapeNets: A deep representation for volumetric shapes. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); IEEE, 2015; pp. 1912–1920.
274. Fathi, H.; Dai, F.; Lourakis, M. Automated as-built 3D reconstruction of civil infrastructure using computer vision: Achievements, opportunities, and challenges. *Advanced Engineering Informatics* **2015**, *29*, 149–161.
275. Zhang, D.; Lu, G. Review of shape representation and description techniques. *Pattern Recognition* **2004**, *37*, 1–19.
276. Tang, P.; Huber, D.; Akinci, B.; Lipman, R.; Lytle, A. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in Construction* **2010**, *19*, 829–843.
277. Remondino, F. From point cloud to surface: the modeling and visualization problem. In Proceedings of the International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences; 2003; Vol. XXXIV, pp. 24–28.
278. Cazals, F.; Giesen, J. Delaunay Triangulation Based Surface Reconstruction : Ideas and Algorithms. *INRIA Rapport de recherche* **2004**, 1–45.
279. Bommes, D.; Bruno, L.; Pietroni, N.; Levy, B.; Pietroni, N.; Puppo, E.; A, C.S.; Tarini, M.; Zorin, D. State of the Art in Quad Meshing. In Proceedings of the Eurographics STARS; 2012; Vol. xx, pp. 1–24.
280. Valero, E.; Adán, A.; Cerrada, C. Automatic Method for Building Indoor

Boundary Models from Dense Point Clouds Collected by Laser Scanners. *Sensors* **2012**, *12*, 16099–16115.
281. Berger, M.; Tagliasacchi, A.; Seversky, L.; Alliez, P.; Levine, J.; Sharf, A.; Silva, C. *State of the Art in Surface Reconstruction from Point Clouds*; 2014; Vol. 1;.
282. Canny, J. *The complexity of robot motion planning*; MIT Press, 1987; ISBN 9780262031363.
283. Campbell, R.J.; Flynn, P.J. A Survey Of Free-Form Object Representation and Recognition Techniques. *Computer Vision and Image Understanding* **2001**, *81*, 166–210.
284. Kim, Y.M.; Mitra, N.; Yang, D.-M.; Guibas, L. Acquiring 3D indoor environments with variability and repetition. *ACM Transactions on Graphics* **2012**, *31*.
285. Budroni, A.; Boehm, J. Toward automatic reconstruction of interiors from laser data. In Proceedings of the Proceedings of the Virtual Reconstruction and Visualization of Complex Architectures (3D-ARCH); Trento, Italy, 2009.
286. Pătrăucean, V.; Armeni, I.; Nahangi, M.; Yeung, J.; Haas, C. State of research in automatic as-built modelling. *Advanced Engineering Informatics* **2015**, *29*, 162–171.
287. Lee, J.; Son, H.; Kim, C.; Kim, C. Skeleton-based 3D reconstruction of as-built pipelines from laser-scan data. *Automation in Construction* **2013**, *35*, 199–207.
288. Fayolle, P.A.; Pasko, A. An evolutionary approach to the extraction of object construction trees from 3D point clouds. *CAD Computer Aided Design* **2016**, *74*, 1–17.
289. Adan, A.; Huber, D. 3D reconstruction of interior wall surfaces under occlusion and clutter. In Proceedings of the Proceedings - 2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, 3DIMPVT 2011; IEEE, 2011; pp. 275–281.
290. Stamos, I.; Yu, G.; Wolberg, G.; Zokai, S. 3D modeling using planar segments and mesh elements. In Proceedings of the Proceedings - Third International Symposium on 3D Data Processing, Visualization, and Transmission, 3DPVT 2006; 2007; pp. 599–606.
291. Lafarge, F.; Keriven, R.; Brédif, M. Insertion of 3-D-primitives in mesh-based representations: Towards compact models preserving the details. *IEEE Transactions on Image Processing* **2010**, *19*, 1683–1694.
292. Lafarge, F.; Mallet, C. Creating large-scale city models from 3D-point clouds: A robust approach with hybrid representation. *International Journal of Computer Vision* **2012**, *99*, 69–85.
293. De Luca, L.; Veron, P.; Florenzano, M. Reverse engineering of architectural buildings based on a hybrid modeling approach. *Computers & Graphics* **2006**, *30*, 160–176.
294. Schnabel, R.; Wessel, R.; Wahl, R.; Klein, R. Shape recognition in 3d point-clouds. In Proceedings of the 16th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision; 2008; Vol. 8.
295. Sanchez, V.; Zakhor, A. Planar 3D modeling of building interiors from point

cloud data. In Proceedings of the Proceedings - International Conference on Image Processing, ICIP; IEEE, 2012; pp. 1777–1780.
296. Bosché, F. Plane-based registration of construction laser scans with 3D/4D building models. *Advanced Engineering Informatics* **2012**, *26*, 90–102.
297. Budroni, A.; Boehm, J. Automated 3D Reconstruction of Interiors from Point Clouds. *International Journal of Architectural Computing* **2010**, *08*, 55–74.
298. Macher, H.; Landes, T.; Grussenmeyer, P. From Point Clouds to Building Information Models: 3D Semi-Automatic Reconstruction of Indoors of Existing Buildings. *Applied Sciences* **2017**, *7*, 1030.
299. Jung, J.; Hong, S.; Jeong, S.; Kim, S.; Cho, H.; Hong, S.; Heo, J. Productive modeling for development of as-built BIM of existing indoor structures. *Automation in Construction* **2014**, *42*, 68–77.
300. Okorn, B.; Xiong, X.; Akinci, B.; Huber, D. Toward Automated Modeling of Floor Plans. In Proceedings of the Symposium on 3D Data Processing, Visualization and Transmission (3DPVT 10); Paris, 2010; Vol. 2.
301. Brédif, M.; Vallet, B.; Ferrand, B. Distributed Dimensonality-Based Rendering of LIDAR Point Clouds. In Proceedings of the The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences; Copernicus GmbH: La Grande Motte, 2015; Vol. 40, p. 559.
302. Nurunnabi, A.; Sadahiro, Y.; Lindenbergh, R. Robust cylinder fitting in three-dimensional point cloud data. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives; 2017; Vol. 42, pp. 63–70.
303. Kim, C.; Son, H.; Kim, C. Fully automated registration of 3D data to a 3D CAD model for project progress monitoring. *Automation in Construction* **2013**, *35*, 587–594.
304. Ortiz, P.; Sánchez, H.; Pires, H.; Pérez, J.A. Experiences about fusioning 3D digitalization techniques for cultural heritage documentation. In Proceedings of the International Archives of Photogrammetry and Remote Sensing; Maas, H.-G., Schneider, D., Eds.; 2006; Vol. XXXVI.
305. Nurunnabi, A.; Belton, D.; West, G. Diagnostic-robust statistical analysis for local surface fitting in 3D point cloud data. In Proceedings of the ISPRS Annals of the Photogrammetry Remote Sensing and Spatial Information Sciences; 2012; Vol. I-3, pp. 269–274.
306. Mura, C.; Mattausch, O.; Jaspe Villanueva, A.; Gobbetti, E.; Pajarola, R. Automatic room detection and reconstruction in cluttered indoor environments with complex room layouts. *Computers and Graphics* **2014**, *44*, 20–32.
307. Xu, Y.; Tuttas, S.; Hoegner, L.; Stilla, U. Reconstruction of scaffolds from a photogrammetric point cloud of construction sites using a novel 3D local feature descriptor. *Automation in Construction* **2018**, *85*, 76–95.
308. Osada, R.; Funkhouser, T.; Chazelle, B.; Dobkin, D. Matching 3D models with shape distributions. In Proceedings of the Proceedings - International Conference on Shape Modeling and Applications, SMI 2001; IEEE Computer

Society, 2001; pp. 154–166.
309. Mitra, N.J.; Pauly, M.; Wand, M.; Ceylan, D. Symmetry in 3D geometry: Extraction and applications. In Proceedings of the Computer Graphics Forum; Wiley/Blackwell (10.1111), 2013; Vol. 32, pp. 1–23.
310. Ikehata, S.; Yang, H.; Furukawa, Y.S. Structured Indoor Modeling Supplementary Material. In Proceedings of the Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV); 2015; pp. 1323–1331.
311. Martinet, A.; Soler, C.; Holzschuch, N.; Sillion, F.X. Accurate detection of symmetries in 3D shapes. *ACM Transactions on Graphics* **2006**, *25*, 439–464.
312. Kovács, I.; Várady, T.; Salvi, P. Applying geometric constraints for perfecting CAD models in reverse engineering. *Graphical Models* **2015**, *82*, 44–57.
313. Kazhdan, M.; Funkhouser, T.; Rusinkiewicz, S. Symmetry descriptors and 3D shape matching. In Proceedings of the Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing - SGP '04; 2004; p. 115.
314. Lafarge, F.; Descombes, X.; Zerubia, J.; Pierrot-Deseilligny, M. Structural approach for building reconstruction from a single DSM. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2010**, *32*, 135–147.
315. Nan, L.; Xie, K.; Sharf, A. A search-classify approach for cluttered indoor scene understanding. *ACM Transactions on Graphics* **2012**, *31*, 1.
316. Bosché, F. Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction. *Advanced Engineering Informatics* **2010**, *24*, 107–118.
317. Besl, P.J.; McKay, N.D. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **1992**, *14*, 239–256.
318. Sharp, G.C.; Lee, S.W.; Wehe, D.K. ICP registration using invariant features. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2002**, *24*, 90–102.
319. Gruen, A.; Akca, D. Least squares 3D surface and curve matching. *ISPRS Journal of Photogrammetry and Remote Sensing* **2005**, *59*, 151–174.
320. Rabbani, T.; Dijkman, S.; van den Heuvel, F.; Vosselman, G. An integrated approach for modelling and global registration of point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* **2007**, *61*, 355–370.
321. Boykov, Y.; Kolmogorov, V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2004**, *26*, 1124–1137.
322. Xu, K.; Li, H.; Zhang, H.; Cohen-Or, D.; Xiong, Y.; Cheng, Z.-Q. Style-content separation by anisotropic part scales. *ACM Transactions on Graphics* **2010**, *29*, 1.
323. Xu, K.; Zheng, H.; Zhang, H.; Cohen-Or, D.; Liu, L.; Xiong, Y. Photo-inspired model-driven 3D object modeling. In Proceedings of the ACM SIGGRAPH 2011 papers on - SIGGRAPH '11; 2011; p. 1.
324. Zlatanova, S.; Van Oosterom, P.J.M.; Lee, J.; Li, K.-J.; Lemmen, C.H.J. LADM AND

INDOORGML FOR SUPPORT OF INDOOR SPACE IDENTIFICATION. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **2016**, *IV-2/W1*, 257–263.
325. Lee, J.; Li, K.-J.; Zlatanova, S.; Kolbe, T.H.; Nagel, C.; Becker, T. *OGC®indoorgml*; 2014;
326. Vanlande, R.; Nicolle, C.; Cruz, C. IFC and building lifecycle management. *Automation in Construction* **2008**, *18*, 70–78.
327. Bhojraj, S.; Lee, C.M.C.; Oler, D.K. What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research. *Journal of Accounting Research* **2003**, *41*, 745–774.
328. Quintana, B.; Prieto, S.A.; Adán, A.; Vázquez, A.S. Semantic scan planning for indoor structural elements of buildings. *Advanced Engineering Informatics* **2016**, *30*, 643–659.
329. Xiong, X.; Huber, D. Using Context to Create Semantic 3D Models of Indoor Environments. In Proceedings of the British Machine Vision Conference; BMVA Press, 2010; pp. 1–11.
330. Fisher, R.B. Applying knowledge to reverse engineering problems. *Computer-Aided Design* **2004**, *36*, 501–510.
331. Xu, M.; Petrou, M. 3D Scene interpretation by combining probability theory and logic: The tower of knowledge. *Computer Vision and Image Understanding* **2011**, *115*, 1581–1596.
332. Sainz, M.; Pajarola, R. Point-based rendering techniques. *Computers & Graphics* **2004**, *28*, 869–879.
333. Bui, G.; Le, T.; Morago, B.; Duan, Y. Point-based rendering enhancement via deep learning. *The Visual Computer* **2018**, *34*, 829–841.
334. Kobbelt, L.; Botsch, M. A survey of point-based techniques in computer graphics. *Computers & Graphics* **2004**, *28*, 801–814.
335. Rusu, R.B.; Cousins, S. 3d is here: Point cloud library (pcl). In Proceedings of the 2011 IEEE International Conference on Robotic and Automation; IEEE: Shangai, 2011; pp. 1–4.
336. Clementini, E.; Di Felice, P.; van Oosterom, P. A small set of formal topological relationships suitable for end-user interaction. In *Advances in Spatial Databases*; Abel, D., Chin Ooi, B., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1993; pp. 277–295 ISBN 978-3-540-47765-5.
337. Sutton, M.; Stark, L.; Bowyer, K. Function from visual analysis and physical interaction: a methodology for recognition of generic classes of objects. *Image and Vision Computing* **1998**, *16*, 745–763.
338. Kazhdan, M.; Bolitho, M.; Hoppe, H. Poisson surface reconstruction. *Eurographics symposium on Geometry processing* **2006**, 1–10.
339. Bernardini, F.; Mittleman, J.; Rushmeier, H.; Silva, C.; Taubin, G. The ball-pivoting algorithm for surface reconstruction. *Transactions on Visualization and Computer Graphics* **1999**, *5*, 349–359.
340. Lorensen, W.E.; Cline, H.E.; Lorensen, W.E.; Cline, H.E. Marching cubes: A high resolution 3D surface construction algorithm. In Proceedings of the

Proceedings of the 14th annual conference on Computer graphics and interactive techniques - SIGGRAPH '87; ACM Press: New York, New York, USA, 1987; Vol. 21, pp. 163–169.
341. Sirin, E.; Parsia, B.; Grau, B.C.; Kalyanpur, A.; Katz, Y. Pellet: A practical OWL-DL reasoner. *Web Semantics: Science, Services and Agents on the World Wide Web* **2007**, *5*, 51–53.
342. Glimm, B.; Horrocks, I.; Motik, B.; Stoilos, G.; Wang, Z. HermiT: An OWL 2 Reasoner. *Journal of Automated Reasoning* **2014**, *53*, 245–269.
343. Ogden, C.; Richards, I.; Ranulf, S.; Cassirer, E. *The Meaning of Meaning. A Study of the Influence of Language upon Thought and of the Science of Symbolism*; A Harvest.; Harcourt, Brace & World Inc.: New York, 1923; ISBN 978-0156584463.

CURRICULUM VITAE

Florent Poux (1990) was born in Toulouse (France) were he obtained a DEUG of Science at the University of Toulouse III - Paul Sabatier. He graduated in 2013 as a Geomatics Engineer from the renowned CNAM “Ecole Supérieure des Géomètres et Topographes” (Le Mans, France), with a speciality in spatial data processing accomplished within the Schulich School of Engineering of the University of Calgary (Canada).

Florent then gained valuable industry knowledge by specializing within the French Order of Surveyors for two years, making him eligible to the title of “Géomètre-Expert”. Within this time, he worked as the head of the 3D reality capture department of the OxyGéo company. He participated in many projects involving precise topography, UAVs and 3D laser scanners applied to buildings monitoring, military demining, cultural heritage reconstruction and transportation systems. He worked all over Europe and developed new international partnerships with key actors.

In 2014, Florent co-founded the company Geovast 3D which offers innovative point cloud solutions to speed-up workflows and increase the collaborative process. These are deployed across the world and used by major companies solving data management, automation and visualisation problematics. Currently, he holds the position of director and oversees the development of the structure and gives the principal vision.

In parallel, he decided to join in October 2015 the Geomatics Unit of the University of Liège and was nominated as a teaching and research assistant. Florent gives practical and courses in topography, 3D reality capture and automation to geographical science students from License to Master. He has also been involved in supervising BSc and MSc thesis, and participated in many projects such as Interreg, VR-labs and international 3D reconstruction of cultural heritage. As part of his research duties, he was selected as a PhD student under the supervision of Prof. Roland Billen, head of the Geomatics group and vice-dean of the Faculty of Sciences. The research resulted in software and multiple papers in books, international journals and conferences (pp. ii), with an outreach spanning several countries and domains. He carried out a research visit at the Visual Computing Institute of the renowned RWTH Aachen University (Germany), focusing on human-in-the-loop semantic segmentation problematics using deep learning approaches.

During his young academic and scientific career, Florent has served on the scientific & program committee of several international conferences and assisted as a reviewer for leading international GIS and Automation journals such as Remote Sensing, IJGI, Land and Sensors.

LIST OF PUBLICATIONS

INTERNATIONAL JOURNAL PAPERS (ISI – OPEN ACCESS)

1. **Poux, F.**, & Billen, R. (2019, in press). Voxel-based 3D point cloud semantic segmentation: unsupervised geometric and relationship featuring vs deep learning methods. *ISPRS International Journal of Geo-Information*.

2. Neuville, R., Pouliot, J., **Poux, F.**, & Billen, R. (2019). 3D Viewpoint Management and Navigation in Urban Planning: Application to the Exploratory Phase. *Remote Sensing*, *11*(3), 236. https://doi.org/10.3390/rs11030236

3. **Poux, F.**, Neuville, R., Nys, G.-A., & Billen, R. (2018). 3D Point Cloud Semantic Modelling: Integrated Framework for Indoor Spaces and Furniture. *Remote Sensing*, *10*(9), 1412. https://doi.org/10.3390/rs10091412

4. Neuville, R., Pouliot, J., **Poux, F.**, de Rudder, L., & Billen, R. (2018). A Formalized 3D Geovisualization Illustrated to Selectivity Purpose of Virtual 3D City Model. *ISPRS International Journal of Geo-Information*, *7*(5), 194. https://doi.org/10.3390/ijgi7050194

5. **Poux, F.**, Neuville, R., Van Wersch, L., Nys, G.-A., & Billen, R. (2017). 3D Point Clouds in Archaeology: Advances in Acquisition, Processing and Knowledge Integration Applied to Quasi-Planar Objects. *Geosciences*, *7*(4), 96. https://doi.org/10.3390/GEOSCIENCES7040096

6. **Poux, F.**, Hallot, P., Jonlet, B., Carre, C., & Billen, R. (2014). Segmentation semi-automatique pour le traitement de données 3D denses: application au patrimoine architectural. *XYZ: La Revue de l' Association Française de Topographie*, *141*(4), 69–75. Retrieved from http://orbi.ulg.be/handle/2268/173626

BOOKS AND BOOKS CHAPTERS

7. **Poux, F.**, & Billen, R. (2019, in press). A Smart Point Cloud Infrastructure for intelligent environments. In *ISPRS Book Series,* R. Lindenbergh & R. Belen (Eds.). CRC Press.

8. **Poux Florent**. (2013). *Vers de nouvelles perspectives lasergrammétriques : Optimisation et automatisation de la chaîne de production de modèles 3D*. (EUE, Ed.) (Univ Europ).

INTERNATIONAL REFEREED CONFERENCES

9. Neuville, R., Pouliot, J., **Poux, F.**, Hallot, P., De Rudder, L., & Billen, R. (2017). Towards A Decision Support Tool For 3d Visualisation: Application To Selectivity Purpose Of Single Object In A 3d City Scene. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Vol. 4). https://doi.org/10.5194/isprs-annals-IV-4-W5-91-2017

10. **Poux, F.**, Neuville, R., & Billen, R. (2017). Point Cloud Classification Of Tesserae From Terrestrial Laser Data Combined With Dense Image Matching For Archaeological Information Extraction. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *IV-2/W2*, 203–211. https://doi.org/10.5194/isprs-annals-IV-2-W2-203-2017

11. **Poux, F.**, Neuville, R., Hallot, P., Van Wersch, L., Jancsó, A. L., Billen, R., Luczfalvy Jancsó, A., & Billen, R. (2017). Digital Investigations Of An Archaeological Smart Point Cloud: A Real Time Web-Based Platform To Manage The Visualisation Of Semantical Queries. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-5/W1*(5W1), 581–588. https://doi.org/10.5194/isprs-archives-XLII-5-W1-581-2017

12. **Poux, F.**, Neuville, R., Hallot, P., & Billen, R. (2017). Model For Reasoning From Semantically Rich Point Cloud Data. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *IV-4/W5*, 107–115. https://doi.org/https://doi.org/10.5194/isprs-annals-IV-4-W5-107-2017

13. Luczfalvy Jancsó, A., Jonlet, B., Hallot, P., **Poux, F.**, Hoffsummer, P., & Billen, R. (2016). CASTLE4D: An Archaeological Information System Based on 3D Point Clouds. In J. L. Lerma & M. Cabrelles (Eds.), *Proceedings of the 8th International Congress on Archaeology, Computer Graphics, Cultural Heritage and Innovation* Valencia. https://doi.org/http://dx.doi.org/10.4995/arqueologica8.2016.4210

14. **Poux, F.**, Neuville, R., Hallot, P., & Billen, R. (2016). Point clouds as an efficient multiscale layered spatial representation. In T. Vincent & F. Biljecki (Eds.), *Eurographics Workshop on Urban Data Modelling and Visualisation*. The Eurographics Association. https://doi.org/10.2312/udmv.20161417

15. **Poux, F.**, Hallot, P., Neuville, R., & Billen, R. (2016). Smart Point Cloud: Definition And Remaining Challenges. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *IV-2/W1*, 119–127. https://doi.org/10.5194/isprs-annals-IV-2-W1-119-2016

16. Neuville, R., **Poux, F.**, Hallot, P., & Billen, R. (2016). Towards a normalized 3D geovizualisation: The viewpoint management. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *IV-2/W1*, 179–186. https://doi.org/10.5194/isprs-annals-IV-2-W1-179-2016

17. Novel, C., Keriven, R., **Poux, F.**, & Graindorge, P. (2015). Comparing Aerial Photogrammetry and 3D Laser Scanning Methods for Creating 3D Models of Complex Objects. In *Capturing Reality Forum* (p. 15). Salzburg: Bentley Systems. Retrieved from http://orbi.ulg.ac.be/handle/2268/190320

18. **Poux, F.**, & Billen, R. (2015). Smart point cloud: Toward an intelligent documentation of our world. In *PCON* (p. 11). Liège. Retrieved from http://orbi.ulg.ac.be/handle/2268/190318

.

APPENDIX A

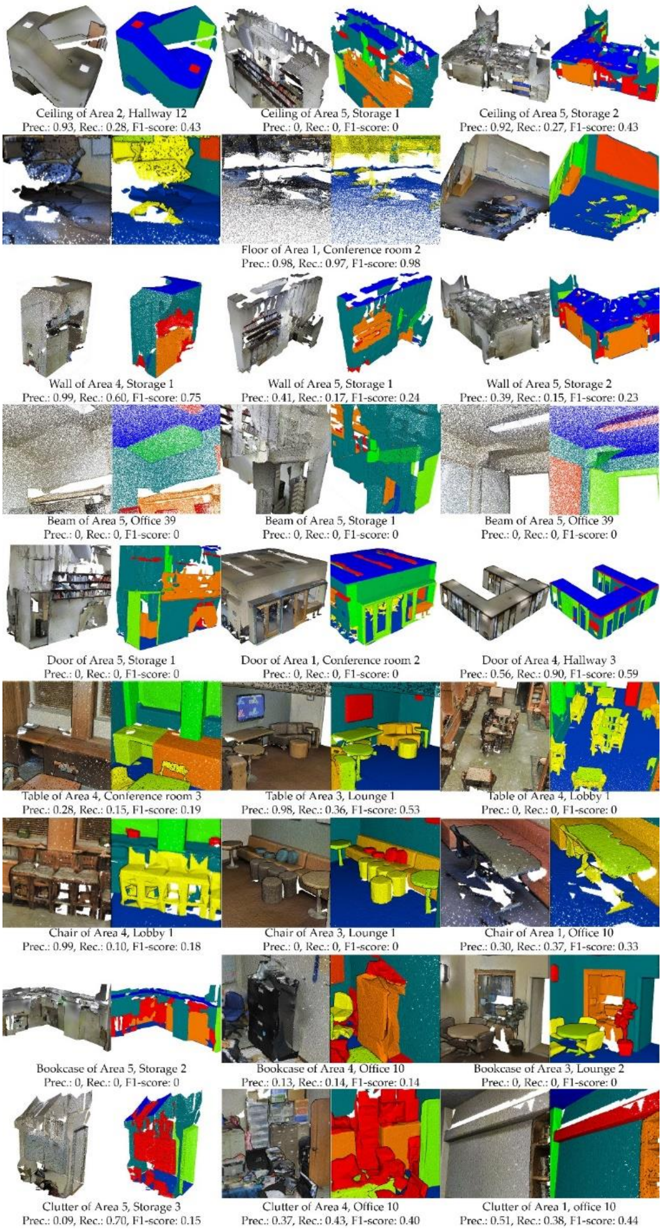
We decided to hold $\overline{IoU}$ metrics to get an idea of the worst possible scores and compare them with the three methods listed in Table 27.

Table 27. Intersection-over-Union on Area 5 of our methodology compared to PointNet [101], SegCloud [160] and SuperPoint Graphs (SPG [161])

IoU for Area-5	Ceiling 0	Floor 1	Wall 2	Beam 3	Door 6	Table 7	Chair 8	Bookcase 10	Clutter 12
PointNet [101]	88.8	97.33	69.8	0.05	10.76	58.93	52.61	40.28	33.22
SegCloud [160]	90.06	96.05	69.86	0	23.12	70.4	75.89	58.42	41.6
SPG [161]	91.49	97.89	75.89	0	52.29	77.4	86.35	65.49	50.67
Ours	85.78	92.91	71.32	0	7.54	31.15	29.02	23.48	21.91

We see that scores obtained for the floor and the ceilings are comparable to the ones obtained by the three deep learning approaches. However, the wall detection ratio outperforms both PointNet and SegCloud, but SPG are still showing better performances. This is explained by the high level of noise and irregular structure. The beam presents a null score (as benchmarked methods) due to the very little number of points and specificity of the 3 beams in the ground truth dataset labelled containing 22 424 points.

APPENDIX B


Ceiling of Area 2, Hallway 12
Prec.: 0.93, Rec.: 0.28, F1-score: 0.43
Ceiling of Area 5, Storage 1
Prec.: 0, Rec.: 0, F1-score: 0
Ceiling of Area 5, Storage 2
Prec.: 0.92, Rec.: 0.27, F1-score: 0.43
Floor of Area 1, Conference room 2
Prec.: 0.98, Rec.: 0.97, F1-score: 0.98
Wall of Area 4, Storage 1
Prec.: 0.99, Rec.: 0.60, F1-score: 0.75
Wall of Area 5, Storage 1
Prec.: 0.41, Rec.: 0.17, F1-score: 0.24
Wall of Area 5, Storage 2
Prec.: 0.39, Rec.: 0.15, F1-score: 0.23
Beam of Area 5, Office 39
Prec.: 0, Rec.: 0, F1-score: 0
Beam of Area 5, Storage 1
Prec.: 0, Rec.: 0, F1-score: 0
Beam of Area 5, Office 39
Prec.: 0, Rec.: 0, F1-score: 0
Door of Area 5, Storage 1
Prec.: 0, Rec.: 0, F1-score: 0
Door of Area 1, Conference room 2
Prec.: 0, Rec.: 0, F1-score: 0
Door of Area 4, Hallway 3
Prec.: 0.56, Rec.: 0.90, F1-score: 0.59
Table of Area 4, Conference room 3
Prec.: 0.28, Rec.: 0.15, F1-score: 0.19
Table of Area 3, Lounge 1
Prec.: 0.98, Rec.: 0.36, F1-score: 0.53
Table of Area 4, Lobby 1
Prec.: 0, Rec.: 0, F1-score: 0
Chair of Area 4, Lobby 1
Prec.: 0.99, Rec.: 0.10, F1-score: 0.18
Chair of Area 3, Lounge 1
Prec.: 0, Rec.: 0, F1-score: 0
Chair of Area 1, Office 10
Prec.: 0.30, Rec.: 0.37, F1-score: 0.33
Bookcase of Area 5, Storage 2
Prec.: 0, Rec.: 0, F1-score: 0
Bookcase of Area 4, Office 10
Prec.: 0.13, Rec.: 0.14, F1-score: 0.14
Bookcase of Area 3, Lounge 2
Prec.: 0, Rec.: 0, F1-score: 0
Clutter of Area 5, Storage 3
Prec.: 0.09, Rec.: 0.70, F1-score: 0.15
Clutter of Area 4, Office 10
Prec.: 0.37, Rec.: 0.43, F1-score: 0.40
Clutter of Area 1, office 10
Prec.: 0.51, Rec.: 0.38, F1-score: 0.44

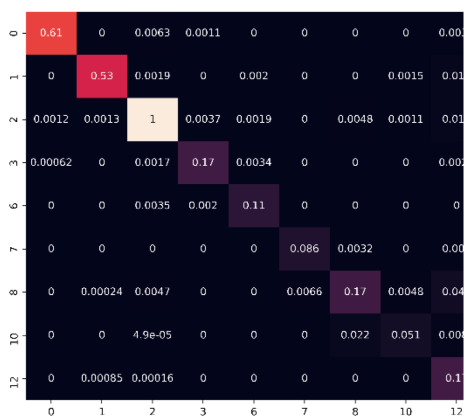
APPENDIX C

We provide the summary of our analysis conducted per area in Table 28.

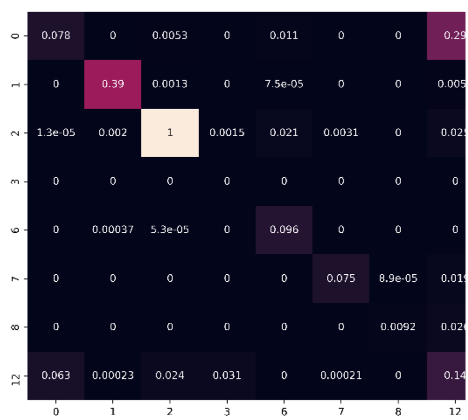
Table 28. F1-score summary of the 6 areas using our semantic segmentation

F1-score	Ceiling 0	Floor 1	Wall 2	Beam 3	Door 6	Table 7	Chair 8	Bookcase 10	Clutter 12
Area-1	0.97	0.96	0.80	0.66	0.24	0.48	0.48	0.26	0.47
Area-2	0.85	0.94	0.70	0.15	0.22	0.11	0.12	0.26	0.32
Area-3	0.98	0.98	0.78	0.61	0.21	0.41	0.61	0.38	0.50
Area-4	0.90	0.97	0.78	0.00	0.12	0.25	0.40	0.24	0.35
Area-5	0.92	0.96	0.83	0.00	0.14	0.48	0.45	0.38	0.36
Area-6	0.95	0.97	0.78	0.58	0.24	0.54	0.53	0.28	0.43

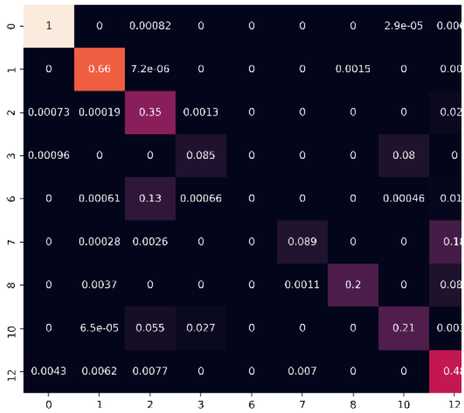
We note that Area-2 is responsible for a drop of performance in ceiling and floor detection, as well as Area-4, which is explained by the very irregular structures of the ceiling and the presence of multilevel stairs. Wall detection is constant among areas whereas beams are very irregular and explain the drop of performances in non-weighted. The classes in Areas 2, 4 and 5 are very specific and in a very low number of occurrences. Table and chair detection rates are very constant and give place for future improvements. Bookcase and clutter also show very similar detection rates per area and demand a global classification optimization for higher performances. As seen above, 'table' presents an unsatisfying detection rate. This is due to the very low recall score, as our classifier only tagged points which were surely a table.



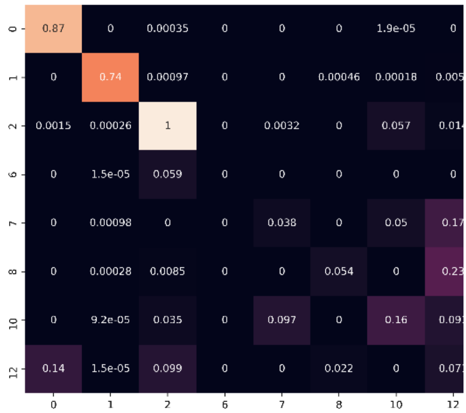
Area 1 – Conference room 1



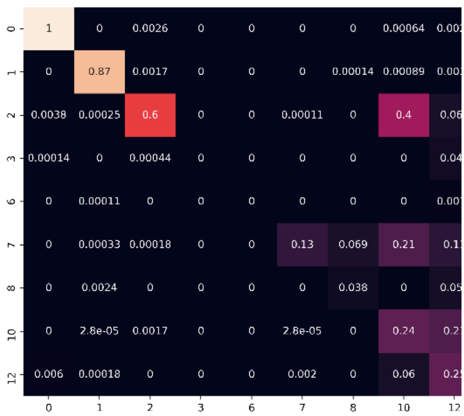
Area 2 – Office 10



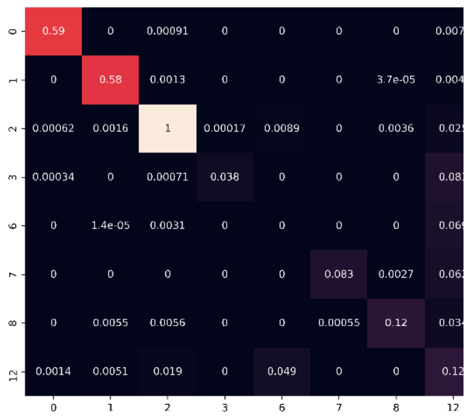
Area 3 – Office 2



Area 4 – Conference room 3



Area 5 – Office 39



Area 6 – Conference room 1