2021 European Conference on Computing in Construction
Online eConference
July 26-28, 2021

"BIM-TO-SCAN" FOR SCAN-TO-BIM: GENERATING REALISTIC SYNTHETIC GROUND TRUTH POINT CLOUDS BASED ON INDUSTRIAL 3D MODELS

Florian Noichl, Alexander Braun, André Borrmann
Technical University of Munich, Germany

ABSTRACT

In the field of Scan-to-BIM, recent developments achieve promising results in accuracy and flexibility, leveraging tools from the field of deep learning for semantic segmentation of raw point cloud data. Those methods demand large-scale, domain-specific datasets for training. Promising ideas to fulfill this need use primitive synthetic point cloud data, which predominantly lack distinct point cloud properties, such as missing patches due to occlusions in the scene. To solve this issue, we use a specialized laser scan simulation tool from the domain of Geosciences in a toolchain that allows generating realistic ground truth data based on 3D models. In this context, we introduce a comprehensive taxonomy for the industrial point cloud context. Furthermore, we provide the missing link for a comprehensive, open-source toolchain that is flexible towards any use case in the field.

INTRODUCTION

Motivation

With the introduction of the Building Information Modeling (BIM) method, the sector of Architecture, Engineering and Construction (AEC) is trying to overcome conventional planning processes that are document-based and inherently describe the built environment using 2D plans (Borrmann et al. 2015). While the adoption of BIM has been largely focused on the disciplines of planning and design, recent efforts target the method's potential in the operation phase. In this, relevant up-to-date information is fed into the digital representation (as-built model) and then leveraged to enable improved decision-making using the so-called Digital Twin. While this concept originated from the field of mechanical engineering (Kritzinger et al. 2018), it is expected to have a significant impact in the AEC sector (Sacks, Brilakis, Pikas, Xie & Girolami 2020).

In the early stages of a project, the primary potential of BIM lies in the model-based cooperation between stakeholders, planners, and contractors. During the operations phase, the model itself plays a major part by providing all necessary data for operations to be optimized, both on efficiency and sustainability. However, due to slow adaption in the industry, good semantic 3D models rarely exist for currently used facilities, especially when the existing facilities are ten or more years old. The conventional way to solve this problem is to capture the facility's as-is condition, and then work with the resulting data to remodel the facility manually. This task is very time-consuming, expensive and error-prone. For owners and operators who would like to benefit from an actual as-built model (also referred to as ”as-is model”, especially if data is collected independently of the construction process (Anil et al. 2013)), but have no useful legacy data, this is very challenging.

Therefore, numerous research projects are conducted in this direction that the community has coined with the term “Scan-to-BIM”, which describes procedures that aim to facilitate remodeling of existing facilities after capturing the current state on-site. More generally, related approaches are gathered within the so-called ”field-to-BIM” domain (Sacks, Girolami & Brilakis 2020). For the sake of completeness, we mention the closely related yet distinct field of ”Scan-vs-BIM”, where the as-is condition of the asset is aligned with an existing model for further processing.

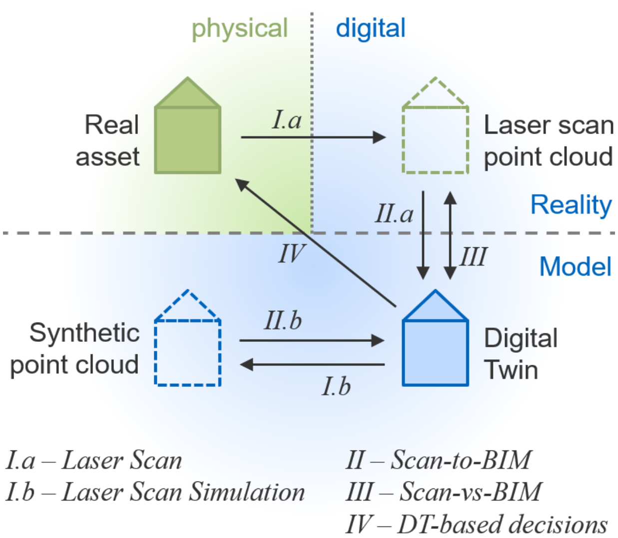


Figure 1: Digital Twin framework to illustrate the domains of Scan-to-BIM and Scan-vs-BIM, extended to incorporate our proposed synthetic data approach in its environment

To capture the actual as-is condition of a facility, recent approaches predominantly use 3D point cloud data captured by LiDAR sensors or photogrammetry. The raw data from these sources are point clouds representing the surfaces of visible objects during capture. This data's main shortcoming is that it does not inherently carry any information regarding the captured objects' semantics, such as information regarding object type or material. However, this information is currently the most accurate and suitable in the generation of a semantic digital building or asset twin or a BIM model.

The point cloud is usually classified and segmented according to the underlying semantics in the first processing step. Recent artificial intelligence (AI) methods have shown great potential in flexibility and precision in this regard.

For these approaches, large-scale, diverse, ground truth datasets are required for two main reasons: Firstly, the precision of AI-based approaches inherently improves with the size and quality of available training datasets. Secondly, both for AI-based and conventional approaches (such as RANSAC and Hough Transform), datasets with ground truth labels are required to provide quantitative results to precisely evaluate those approaches' actual performance. However, the improvement of well-performing solutions for well-known benchmark datasets outpaces the development of diverse, use-case specific and publicly available ground truth datasets. To counteract these developments, we propose a flexible and scalable strategy that provides realistic, synthetic ground truth data.

Related work

A comprehensive overview of solutions using deep learning on point clouds, in general, can be found in (Guo et al. 2020). Patraucean et al. (2015) collect approaches specifically for the field of Scan-to-BIM. In related research areas, publishing labeled datasets (cf. Table 1) to foster the development of approaches that allow automated reasoning about a point cloud's content has already become an essential part of the scientific practice.

Still, the research community occupied with 3D semantic segmentation is over-all lacking labeled datasets compared to the advances made in areas like computer vision, as comprehensivley summarized in Gao et al. (2020). Popular datasets include KITTI (Geiger et al. 2013), Semantic3D (Hackel et al. 2017), Vaihingen (Rottensteiner et al. 2013) and Paris-Lille (Roynard et al. 2018) for outdoor scenes; datasets for indoor scenes have been published less often, popular examples are the S3DIS (Armeni et al. 2016) and ScanNet (Dai et al. 2017), the latter however stored in their voxel representation instead of point clouds.

For our use case in the industrial construction sector, latest research includes a publication on a manually labeled collection of industrial point cloud data (Agapaki et al. 2019). Unfortunately, to this date, the data presented in this work is not accessible to the authors, and therefore the domain remains without a publicly accessible ground truth dataset. Agapaki (2020) provides a comprehensive overview of available annotated point cloud datasets, which we summarize and extend with their acclaimed CLOI dataset in Table 1.

Table 1: Annotated point cloud datasets

Name	Context	No. of points
Semantic3D	urban	400×10^7
KITTI	urban	180×10^7
S3DIS	indoor, office	27×10^7
CLOI	indoor, industrial	14×10^7
Paris-Lille	urban	4.3×10^7

While recent research includes various important approaches for the detection of elements for the industrial use case (Maalek et al. 2019, Son et al. 2015), there remains a lack of data for those use cases. The reason for this are diverse. For one, there is a lack of intrinsic motivation by facility operators to share captured as-is data, as this is not within their traditional scope of business. Furthermore, in places where data is captured, evaluated, and processed, warranted doubts towards open publishing arise for multiple reasons, regarding employee privacy, datasets possibly containing trade secrets, or other justifiably confidential information.

This lack of accessibility to real world annotated data is a bottleneck for the field of Scan-to-BIM. As point clouds from laser scans and photogrammetry consist of points captured from the surface of 3D objects, another way to tackle the issue is to generate synthetic data. This data has been generated using various techniques, including randomly distributing points on 3D surfaces (Ma et al. 2020), adding random noise to each point (Schnabel et al. 2007), or even capturing scenes out of video games (Yue et al. 2018). Another approach to generate large-scale synthetic point cloud data was used in the work of Shen (2020). Ma et al. (2020) have used BIM models for the generation of point clouds by randomly placing points on the object surfaces. Using this synthetic data as an addition to their real world point clouds in the training phase resulted in an increase of IoU (intersection over union) of 7.1%.

However, the data used in their approach lacks realistic properties of point clouds acquired by a laser scanner, especially regarding occlusions, which frequently occur in industrial scenes. Helios (Bechtold & Höfle 2016) and Blensor (Gschwandtner et al. 2011) are tools to generate more realistic laser scan point clouds based on 3D surfaces by simulating laser beams and their reflected signal measurements. The latter has been used for an approach similar to ours to create a ground truth point cloud for a large-scale urban scene (Griffiths & Boehm 2019).

Research gap

Access to large-scale ground truth datasets is crucial to develop methods for semantic segmentation and other ways to automatically process and enrich point cloud data to advance the field of Scan-to-BIM further. Unfortunately, currently no such data is publicly available for the use case of industrial assets.

Industrial companies operating complex facilities have been working with 3D models for many years. For example, in 2009, this was documented in an official report of the German Association of the Automotive Industry (VDA 2009) that contains an effort to communicate industry-wide standardization. Scan data captured in active industrial facilities is delicate regarding issues of employee privacy and nondisclosure. Therefore, this data is restricted, and usually, it is impossible to publish for academic usage and benchmarking of state-of-the-art solutions. The opportunity of working with models is that their content can be checked with ease, elements that are critical can be replaced or removed. The synthetic data generated upon those models will not contain any information that is not present in the model. This leads to the limitation that the resulting point clouds only contain objects included in the model, which can not be circumvented. Our goal is therefore to exploit detailed 3D model data to satisfy the need for 3D scan data. As related work shows the lack of data and the potential of synthetic data, we aim to advance our specific use case in this regard by providing a method to effortlessly generate large amounts of realistic ground truth point cloud data.

METHOD

Overview

In our approach, we utilize the laser scan simulation tool Helios (Bechtold & Höfle 2016), since it is an open source tool that allows to set up the desired simulation in a modular and flexible way. Figure 2 shows an overview of our workflow. The following subsections introduce the steps of this workflow, along with a description of the utilized tools and file formats.

Model import

Starting from the authoring tool, we export the model using preferably vendor-neutral formats that include object classes. The usage of the Industry Foundation Class (IFC) format (ISO 2018) is therefore the best solution. In our testing, we tried the IFC export of Autodesk Revit with good results. If no such information is included in the original data and the model is stored in its plain geometry, we make use of the FBX format. Both tested authoring applications (Autodesk Revit and Bentley Microstation v8i) have

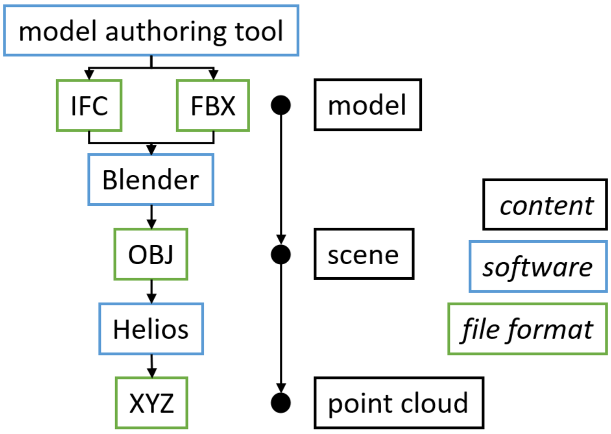


Figure 2: The workflow for "BIM-To-Scan", distinguishing between content, software application used and file format of (intermediate) results

shown a robust export to FBX with reasonable file sizes.

This intermediate file is then imported into the Blender application[1] for further processing. While FBX import is a standard feature with Blender, we use the BlenderBIM plugin[2] for IFC import.

Model preparation

To provide a suitable set of classes for the point cloud labels, we have developed a taxonomy for industrial point cloud data, as depicted in Figure 3. We base this structure on the approach for a point cloud taxonomy presented by Kim et al. (2016) for construction point clouds, and adapt the content to fulfill the requirements of our project, which is the scope of structural and MEP (Mechanical, electrical, and plumbing) elements. To do this, we further base our proposed taxonomy on the findings of Agapaki et al. (2018), who have investigated important objects and shapes in industrial asset models.

Our taxonomy is distributed over three levels. The first level covers a very basic separation into crafts. Level 2 allows to obtain more detailed insight for each level 1 class. The competing goals in this are to keep the classes distinct in regard to their geometry and context, while providing sufficient insight to provide valuable information for the engineering perspective. Grouping all classes into the aforementioned levels allows us to consider classes in varying granularity, which can also be used to adapt classes according to results per applied method in semantic segmentation.

Furthermore, top-down observation of the levels allows for a simple separation of the labeled point cloud into more compact sub-clouds without losing class-relevant data. In comparison to the categories introduced with the CLOI dataset (Agapaki et al. 2019), our approach adds the craft-wise separation as a novel perspective. Additionally, we omit the sepa-

[1] https://blender.org/
[2] https://blenderbim.org/

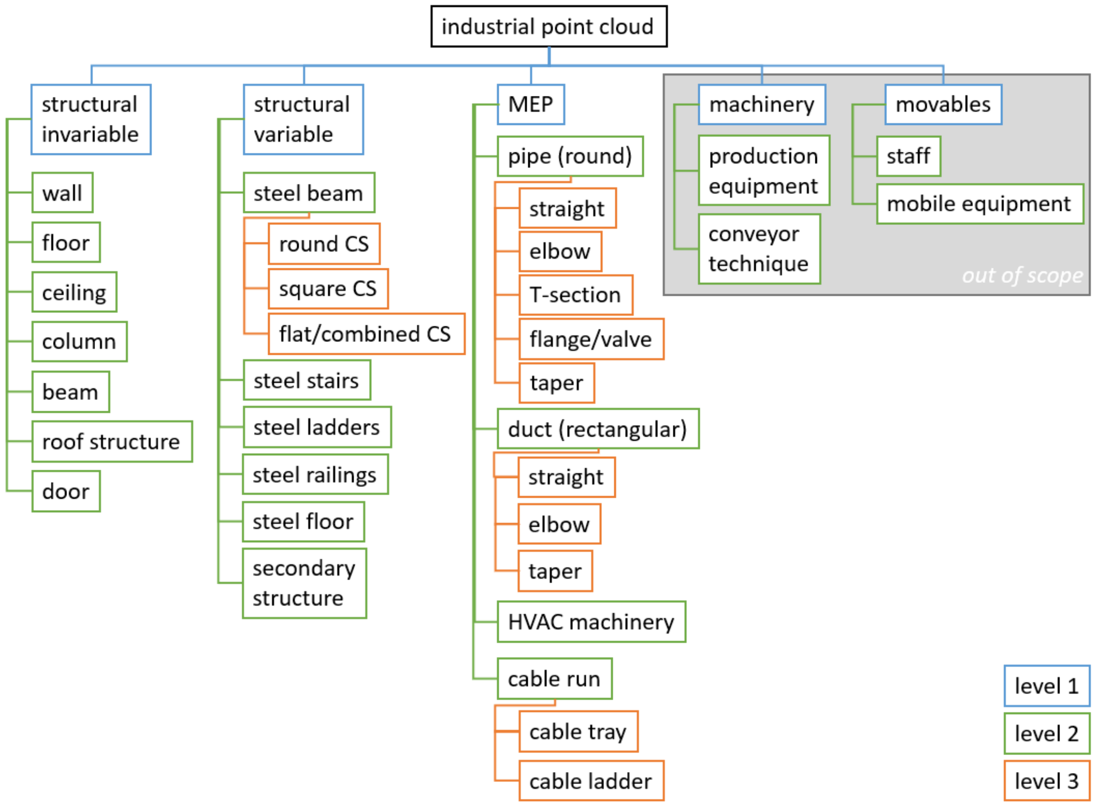


Figure 3: Taxonomy for the industrial use case: The complete point cloud is separated into object classes, that are divided into sub-classes over three levels of granularity

ration of steel cross sections that result in a combination of flat sections (I, L, etc.). From our perspective, this differentiation is overly precise at this point and can, if necessary, better be introduced downstream by analyzing geometric features of previously correctly identified segments, as presented by Kim et al. (2020).

Within Blender, the assignment to categories is executed by grouping geometries in distinct collections. In the case of imported CAD geometries without semantics, this step is fully manual in our approach. Depending on the information content of the original model, it is possible to parse the information from the imported model data to allocate objects to their respective categories automatically. As a proof of concept, this has been tested for an IFC file (IFC2x3) exported from Autodesk Revit.

Scene preparation

To prepare the scene for simulation in Helios, the objects must be stored in single OBJ files, with information on their location and orientation in the scene stored in a XML file. The OBJ (Wavefront OBJ) file format is chosen because it is the standard input requirement for Helios. In this part of our toolchain, we make use of the Blender plugin provided by Neumann (2020). To pass on the information of the previously designated object classes, a separate material dictionary is is used, to map label integers to object class strings.

Survey preparation

In advance of a stationary laser scan in any environment, scan planning is required regarding scan positions. Recent research includes approaches to optimize this using building models in 2D (Díaz-Vilariño et al. 2018) and 3D (Kabir Biswas et al. 2015), a module to optimize planning for simple geometries is included in Helios (Bechtold & Höfle 2016).

In practice, the surveyor usually selects locations by his expert opinion and his assessment on-site, to maximize coverage with a preferably low number of total scans. Additionally, he needs to assure sufficient overlap to guarantee precise registration. In our workflow, the single resulting point clouds do not require registration since the simulation for each scan is performed in the same coordinate system, as previously defined in the model. Scan planning is straightforward in Helios, as it allows the user to place single scans in the scene using the user interface directly.

Laser scan simulation

Helios' code and comprehensive documentation can be found in the project's Github repository[3]. In its core, the simulation is performed by casting single rays as per the defined equipment's functionality. For our use case, this usually means the rotating mirror of a terrestrial laser scanner. In brief, if the simulated ray intersects the surface of the model, the measured location of the intersection is returned along with the measured waveform and material identifier. By including equipment-specific parameters regarding measurement precision (cf. Table 2), the simulated measurements include measurement errors accordingly. In our case, we store the object classes, as defined in our taxonomy in the material definition (as a workaround), such that each point of the synthetic point cloud obtains its distinct class affiliation in the simulation itself. This is necessary to generate fully labeled data for the training of a deep learning network for semantic segmentation.

As an output, we receive a single point cloud per individual scan, just like one would performing a real laser scan. Those can subsequently be combined to a full point cloud by merging them directly. As each point cloud is delivered in the same global coordinate system, registration is not necessary. Each resulting point cloud is stored by default in a XYZ file, which is an open ASCII file format to store point cloud data.

[3] https://github.com/GIScience/helios

CASE STUDY AND RESULTS

In order to compare the quality of the synthetically created point clouds to real ones, we performed both a real laser scan and a laser scan simulation in the frame of a case study. To do so, we chose a facility for which we have access to a detailed as-planned 3D CAD model, which allows us to perform the laser scan simulation for direct comparison.

As scanning equipment, we use a terrestrial laser scanner model FARO FOCUS S 150, with a chosen resolution of 8192×3413 points resulting in 27,959,296 points per scan. This choice is made to reflect industry standard settings for this kind of facility, in accordance with the experts we worked with. The settings are translated to simulation parameters, collected in Table 2, hardware specifications were adopted from the manufacturer[4].

Table 2: Case study simulation parameters

parameter	value
scan frequency	16 Hz
pulse frequency	122,000 Hz
ranging error	± 1 mm
beam divergence	19 arcsec
vertical field of view	300°
horizontal field of view	360°
head rotation	$1.29\ \frac{\circ}{\mathrm{s}}$

A total of 28 scanning locations was chosen in the asset by the surveyor, to allow for solid surface coverage and cloud-to-cloud registration. Registering the single point clouds of this scan amounts to a total of 7.1×10^8 points, which is a size comparable those of the datasets introduced in Table 1. For the laser scan simulation, we used the 3D CAD model and the full workflow introduced in this paper. The 3D CAD model resulted in a triangulated surface model with a total of 4,291,794 vertices and 1,430,466 faces.

To evaluate the computational effort, we ran a single leg of the scan on a notebook with an Intel(R) Core(TM) i7-8665U CPU and 16GB RAM, using Microsoft Windows 10 and Helios in headless mode (no graphical output). Computation time depends on the used hardware, but also heavily on scanning parameters and complexity of the scene. On the described setup, initializing the simulation by loading the scene takes 3 min 38 sec, the simulation of a single scan finishes after 9 min 38 sec if we disable full waveform (FWF) computation. In our current testing, we do not include material properties, therefore we omit FWF computation. We performed the simulation for each of the 28 scan locations of the real laser scan.

To help the general understanding of the process, we provide overview snapshots of the used model, along with the synthetic and real laser scan point

[4] https://faro.com/

clouds in figures 7, 8 and 9. Figure 4 depicts the point cloud category split per level 1 of our taxonomy in the point cloud of a selected single scan. Selected properties of this point cloud are summarized in Table 3.

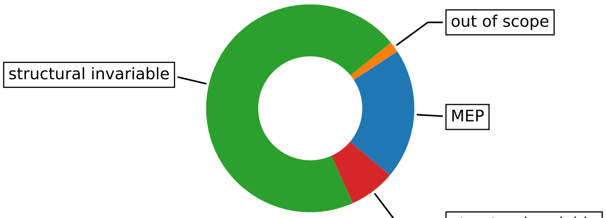


Figure 4: Point cloud category split, per level 1 in our taxonomy

Table 3: Properties of a selected single scan, surface density is calculated for a local neighborhood with r =0.05 m, densities and distances are described with their median values

property	laser scan	synthetic data
points $[pts.]$	25.3×10^6	27.7×10^6
surface density $\left[\frac{pts.}{\pi r^2}\right]$	5602	4584
distance to model [m]	0.01	1×10^{-6}

The variation in these properties regarding density and distance is considerable due to the deviation between model and the actual existing facility. To investigate this, we select the content of a small cuboid volume from the model, the synthetic and real point cloud. In Figure 5 we compare the points contained in this clipping box in two rows that illustrate different aspects of the real scan (upper row) and the synthetic point cloud (lower row).

The difference between Figures 5b and 5e is due to inaccuracies in the laser scan simulation itself. As introduced, we used the same scanning parameters in our real laser scan and in our simulation. However, there remains a divergence between the resulting point densities, which is due to differences between real and modeled surfaces. This is the main shortcoming of synthetic data, and can best be recognized in the distances between point cloud and underlying model surface: While Figure 5f shows very small values because the simulated laser scan stays true to the underlying surface with some divergence, the distances between the real laser scan and the model depicted in Figure 5c are in some areas significant. This mainly due to the fact that our model is not an accurate as-built representation but an as-planned status. In this, object types such as secondary steel structures might be missing completely (as can be seen in the lower parts of Figure 5), others might be less detailed (such as valves, see Figure 5, center parts and 6a) or misplaced (such as the valves, and second pipe run in the lower parts of Figure 5).

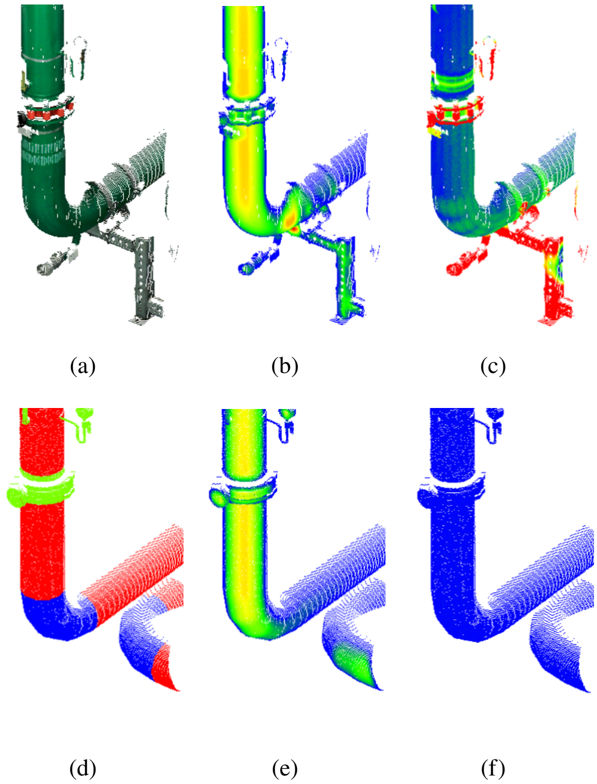
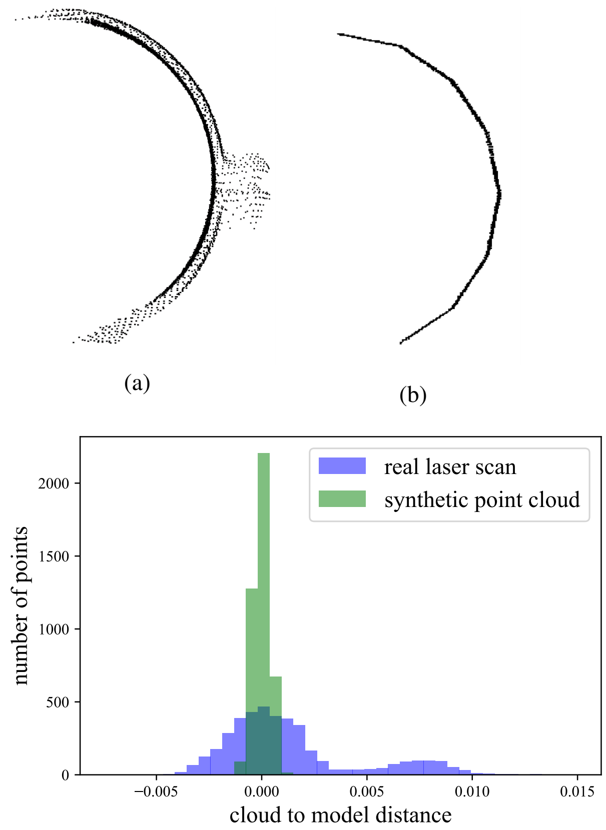


Figure 5: Comparison between real laser scan (upper row, 5a-5c) and synthetic laser scan point clouds (lower row, 5d-5f): RGB values (5a), object classes (5d), surface density with $r = 0.05$ m (5b, 5e), cloud to model distances (5c, 5f)

The deviation of the model from the reality is not only a question of all elements being present in the model, but also of how the geometries are modeled. In our case, this can be observed in the surface of round pipes. The distributions of points over the horizontal pipe section depicted in the lower parts of Figure 5 are shown in Figure 6. This illustrates that the general distribution is qualitatively accurate regarding noise and the occluded rear surface of the pipe, but the underlying surface geometry shows significant differences. While the real pipes exhibit a circular cross section, the pipes in our model are represented by a polygon as the model is handled in the form of a triangulated mesh. In Figure 6c the distances of the points depicted in Figures 6a and 6b to the underlying model surface are collected for comparison. In this, the real laser scan point cloud differs in two main aspects: Firstly, the pipe section includes pipe brackets, which leads to the tail on the right side of the distribution. Secondly, this distribution has a larger standard deviation, which can be explained by deviations between the circular shape of the pipe and the polygon. Furthermore, the synthetic data is generated using the manufacturer's technical specifications, which usually slightly overestimate the precision of the equipment in situ.



(a) (b)

(c)

Figure 6: Distribution of points over the cross sections of a round pipe with according histogram for point to model distances: Real laser scan (6a, 6c blue histogram) and synthetic point cloud (6b, 6c green histogram)

CONCLUSIONS

Opportunities and pitfalls

The main motivation for choosing this approach is that it allows to generate realistic and fully labeled ground truth point cloud data as a basis for large-scale training of deep learning networks for semantic point cloud segmentation. The approach is intrinsically flexible towards various use cases, the only fixed requirement is the availability of a detailed 3D model-ideally but not necessarily as a product model with individual semantic information per object. The manual effort and possible errors in point cloud labeling are completely avoided. The modular concept of the Helios simulation tool allows for changes in choice and placement of scanning equipment. At the same time the approach is freely scalable, and easily repeatable for changing scanning strategies or even a changed taxonomy. The objects will be assigned to their respective classes in the model preparation step, the rest of the task will remain unchanged, point cloud quality unimpaired. Two of the software formats in use (FBX and OBJ) are proprietary, but well established and popular solutions in the field. All software tools used in this are open-source, the toolchain we use in this contribution is accessible in an open Github repository[5].

As we have found in the presented case study, the used parameters overestimate the equipment's precision on site, which means the data is less noisy than the real laser scan point cloud and therefore less realistic in that aspect. This might lead to the case where approaches trained on our synthetic data will rely too much on the higher precision and perform significantly worse on real data. To finally verify the added value of the synthetic data created in our toolchain with regard to training deep learning approaches for semantic segmentation, testing needs to be done along with real data as a next step.

Another limitation in our approach is the remaining manual effort for model preparation. In comparison to completely manually labeling the point clouds, however, this can be done in a fraction of time.

Outlook

Synthetic data is useful as a support for the development of AI solutions in data-weak domains. Whereas straightforward approaches to create such data are known and in use, our workflow allows us to achieve a higher level of realism in the resulting data while keeping the process manageable. Our toolchain allows us to process large-scale models, seamlessly fulfilling all requirements that the utilized simulation tool introduces. We introduce a comprehensive draft for a point cloud taxonomy in the industrial environment for our ground truth labels along with the workflow.

In a next step, we investigate the applicability of approaches that have been trained on our synthetic data to real-world datasets, to evaluate the potential that lies in the method presented in this contribution regarding its intended purpose. We outline our next steps as follows: First we will train and test current state of the art deep learning architectures for semantic segmentation on synthetic point cloud data created as presented in this contribution. Then, we will test the networks' performance on real scan data. To be able to do that, we are simultaneously working on the manual labeling of our real scan data.

ACKNOWLEDGEMENTS

This research was conducted within the scope of a project funded by Audi AG, Ingolstadt. The funding is highly appreciated.

[5] https://github.com/fnoi/B2H

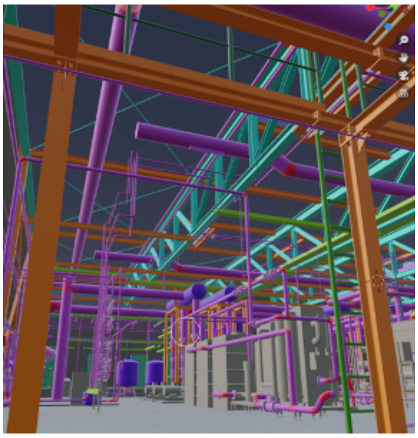


Figure 7: 3D model, objects color-coded according to our taxonomy

Figure 8: Synthetic point cloud, based on the model depicted in figure 7

Figure 9: Real laser scan point cloud of the facility, with actual RGB values

References

Agapaki, E. (2020), Automated Object Segmentation in Existing Industrial Facilities, PhD thesis.

Agapaki, E., Glyn-Davies, A., Mandoki, S. & Brilakis, I. (2019), 'CLOI: A Shape Classification Benchmark Dataset for Industrial Facilities', Computing in Civil Engineering pp. 66–73.

Agapaki, E., Miatt, G. & Brilakis, I. (2018), 'Prioritizing object types for modelling existing industrial facilities', Automation in Construction **96**, 211–223.

Anil, E. B., Tang, P., Akinci, B. & Huber, D. (2013), 'Deviation analysis method for the assessment of the quality of the as-is Building Information Models generated from point cloud data', Automation in Construction **35**, 507–516.
URL: http://dx.doi.org/10.1016/j.autcon.2013.06.003

Armeni, I., Sener, O., Zamir, A. R., Jiang, H., Brilakis, I., Fischer, M. & Savarese, S. (2016), '3D semantic parsing of large-scale indoor spaces', Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition **2016-Decem**, 1534–1543.

Bechtold, S. & Höfle, B. (2016), '{HELIOS}: A Multi-Purpose LiDAR Simulation Framework for Research, Planning and Training of Laser Scanning Operations with Airborne, Ground-Based Mobile and Stationary Platforms', ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences **III-3**, 161–168.

Borrmann, A., König, M., Koch, C. & Beetz, J. (2015), Building Information Modelling, Springer.

Dai, A., Chang, A. X., Savva, M., Halber, M., Funkhouser, T. & Nießner, M. (2017), 'ScanNet: Richly-annotated 3D reconstructions of indoor scenes', Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 **2017-Janua**, 2432–2443.

Díaz-Vilariño, L., Frías, E., Balado, J. & González-Jorge, H. (2018), 'Scan planning and route optimization for control of execution of as-designed bim', International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives **42**(4), 217–224.

Gao, B., Pan, Y., Li, C., Geng, S. & Zhao, H. (2020), 'Are We Hungry for 3D LiDAR Data for Semantic Segmentation?', pp. 1–26.
URL: http://arxiv.org/abs/2006.04307

Geiger, A., Lenz, P., Stiller, C. & Urtasun, R. (2013), 'Vision meets robotics: The KITTI dataset. The International Journal of Robotics Research', The International Journal of Robotics Research (October), 1–6.

Griffiths, D. & Boehm, J. (2019), 'SynthCity: A large scale synthetic point cloud', pp. 1–6.
URL: http://arxiv.org/abs/1907.04758

Gschwandtner, M., Kwitt, R., Uhl, A. & Pree, W. (2011), 'BlenSor: Blender sensor simulation toolbox', Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **6939 LNCS**(PART 2), 199–208.

Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L. & Bennamoun, M. (2020), 'Deep Learning for 3D Point Clouds: A Survey', IEEE Transactions on Pattern Analysis and Machine Intelligence pp. 1–1.

Hackel, T., Savinov, N., Ladicky, L., Wegner, J. D., Schindler, K. & Pollefeys, M. (2017), 'Semantic3D.Net: a New Large-Scale Point Cloud Classification Benchmark', ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **4**(1W1), 91–98.

ISO (2018), Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries — Part 1: Data schema, Standard, International Organization for Standardization, Geneva, CH.

Kabir Biswas, H., Bosché, F. & Sun, M. (2015), Planning for Scanning Using Building Information Models: A Novel Approach with Occlusion Handling, in 'ISARC. Proc. Int. Symp. Autom. Robot. Constr.'.

Kim, H., Kim, K. & Kim, H. (2016), 'Data-driven scene parsing method for recognizing construction site objects in the whole image', Automation in Construction **71**(Part 2), 271–282.
URL: http://dx.doi.org/10.1016/j.autcon.2016.08.018

Kim, Y., Hong, C., Nguyen, P. & Choi, Y. (2020), 'Automatic pipe and elbow recognition from three-dimensional point cloud model of industrial plant piping system using convolutional neural network-based primitive classification', Automation in Construction **116**(December 2019), 103236.
URL: https://doi.org/10.1016/j.autcon.2020.103236

Kritzinger, W., Karner, M., Traar, G., Henjes, J. & Sihn, W. (2018), 'Digital Twin in manufacturing: A categorical literature review and classification', IFAC-PapersOnLine **51**(11), 1016–1022.

Ma, J. W., Czerniawski, T. & Leite, F. (2020), 'Semantic segmentation of point clouds of building interiors with deep learning: Augmenting training datasets with synthetic BIM-based point clouds', Automation in Construction **113**(March), 103144.
URL: https://doi.org/10.1016/j.autcon.2020.103144

Maalek, R., Lichti, D. D., Walker, R., Bhavnani, A. & Ruwanpura, J. Y. (2019), 'Extraction of pipes and flanges from point clouds for automated verification of pre-fabricated modules in oil and gas refinery projects', Automation in Construction **103**(March), 150–167.

Neumann, M. (2020), '{Blender2Helios} - Github Repository'.
URL: https://github.com/neumicha/Blender2Helios

Patraucean, V., Armeni, I., Nahangi, M., Yeung, J., Brilakis, I. & Haas, C. (2015), 'State of research in automatic as-built modelling', 32nd International Symposium on Automation and Robotics in Construction and Mining: Connected to the Future, Proceedings (January 2015).

Rottensteiner, F., Sohn, G., Gerke, M. & Wegner, J. D. (2013), 'ISPRS Test Project on Urban Classification and 3D Building Reconstruction', ISPRS - Commission III - Photogrammetric Computer Vision and Image Analysis Working Group III / 4 - 3D Scene Analysis pp. 1–16.

Roynard, X., Deschaud, J. E. & Goulette, F. (2018), 'Paris-Lille-3D: A large and high-quality ground-truth urban point cloud dataset for automatic segmentation and classification', International Journal of Robotics Research **37**(6), 545–557.

Sacks, R., Brilakis, I., Pikas, E., Xie, H. S. & Girolami, M. (2020), 'Construction with digital twin information systems', Data-centric Engineering .

Sacks, R., Girolami, M. & Brilakis, I. (2020), 'Building Information Modelling, Artificial Intelligence and Construction Tech', Developments in the Built Environment p. 100011.

Schnabel, R., Wahl, R. & Klein, R. (2007), 'Efficient RANSAC for point-cloud shape detection', Computer Graphics Forum **26**(2), 214–226.

Shen, J. (2020), A Simulated Point Cloud Implementation of a Machine Learning Segmentation and Classification Algorithm, Master thesis, Purdue University.

Son, H., Kim, C. & Kim, C. (2015), '3D reconstruction of as-built industrial instrumentation models from laser-scan data and a 3D CAD database based on prior knowledge', Automation in Construction **49**, 193–200.
URL: http://dx.doi.org/10.1016/j.autcon.2014.08.007

VDA (2009), 3D Datenaustausch in der Fabrikplanung, Technical report.

Yue, X., Wu, B., Seshia, S. A., Keutzer, K. & Sangiovanni-Vincentelli, A. L. (2018), 'A LiDAR point cloud generator: From a virtual world to autonomous driving', arXiv pp. 458–464.