Scan2BIM – A Review on the Automated Creation of Semantic-Aware Geometric as-is Models of Bridges

Scan2BIM – Eine Übersicht zur automatisierten Erstellung von geometrisch-semantischen Bestandsmodellen von Brücken

Maximilian Kellner, Hristo Vassilev, Antonia Busch, Robert Blaskow, Mariana Ferrandon Cervantes, Kwasi Nyarko Poku-Agyemang, Annette Schmitt, Sven Weisbrich, Hans-Gerd Maas, Frank Neitzel, Alexander Reiterer, Jörg Blankenbach

Bridges are one of the most important elements in the worldwide traffic infrastructure. Therefore, they require regular inspection and maintenance. In Germany, for instance, many bridges are older than 50 years, which means they need special treatment to keep them in service. One very promising concept for optimizing the maintenance of bridges is BIM or a Digital Twin.
An essential basis of BIM as well as Digital Twins are geometric-semantic models. However, due to the age of many bridges, there are often no digital models available. That's why there is an increasing demand for workflows for the automated creation of geometric-semantic models of existing bridges. This paper presents the general workflow for automated generation for Digital Twins. This literature review starts with an overview of different sensors and data acquisition techniques. Afterwards, different methods for data pre-processing like denoising and voxelization will be discussed. The main part of the literature review shows different AI methods for enriching point clouds with semantic information and extracting the contained geometry. Additionally, the current challenges within this workflow are shown and an outlook for future research is suggested.

Keywords: Bridge infrastructure, Scan2BIM, reality capture, data fusion, semantic segmentation, automated modelling, challenges, Digital Twin, BIM

Brücken sind wesentliche Verkehrsinfrastrukturbauwerke, die regelmäßig inspiziert und gewartet werden müssen. In Deutschland sind viele Brücken älter als 50 Jahre, was zur Folge hat, dass sie einer besonderen Behandlung bedürfen, um sie in Betrieb zu halten. Als vielversprechendes Konzept zur Optimierung der Erhaltung der Verkehrsinfrastruktur werden BIM und Digitale Zwillinge gesehen. Eine wesentliche Grundlage von BIM wie auch von Digitalen Zwillingen sind geometrisch-semantische Bauwerksmodelle. Aufgrund des Alters vieler Brücken sind jedoch oft keine digitalen Modelle vorhanden. Deshalb besteht ein zunehmender Bedarf an Workflows zur automatisierten Erstellung von geometrisch-semantischen As-is-Modellen von Brücken. In diesem Beitrag wird ein Überblick zur automatisierten Generierung von Bestandsmodellen vorgestellt. Nach einer Einführung in geeignete Erfassungssensorik und -techniken werden zunächst unterschiedliche Methoden zur Datenvorverarbeitung diskutiert. Im Hauptteil werden verschiedene KI-Methoden zur Anreicherung von Punktwolken mit semantischen Informationen und zur

DOI: 10.14627/avn.2024.3.4

Extraktion der enthaltenen Geometrie behandelt. Zusätzlich werden die aktuellen Herausforderungen innerhalb dieses Workflows aufgezeigt und ein Ausblick auf zukünftige Forschung gegeben.

Schlüsselwörter: Brückeninfrastruktur, Scan2BIM, Reality Capture, Datenfusion, Semantische Segmentierung, Automatisierte Modellierung, Digitale Zwillinge, BIM

1 INTRODUCTION

Bridges are key elements of efficient transport networks. To keep bridge structures in service, systematic bridge maintenance is carried out as a process. This process starts with the collection of data about the condition state of the bridges all the way through the execution of well-informed maintenance decisions about the bridge inventory /Haardt & Holst 2008/. The German bridge inventory of federal highways is composed of more than 40 000 bridges, 4.6 % of which are categorized as having an inadequate or worse condition according to the March 2023 report from the Federal Highway Research Institute (BASt) /BASt 2023/. The categorization of the condition of the bridges is the result of observations from site inspections, where findings about the bridge state are documented with sketches and photographs /BVBS 2013/. This process may be influenced by the subjective perception of the inspector /Moore et al. 2001/, /Moufti et al. 2014/. Moreover, although a significant proportion of bridges remains in satisfactory or still sufficient condition, it is still crucial to consistently monitor and inspect these structures to detect potential issues at an early stage /Office of Transportation Maintenance 2008/. In this context, the Scan-to-BIM or Scan2BIM process can assist in the timely generation of high-quality data to support maintenance decision making. Scan2BIM is the process of creating semantically rich three-dimensional (3D) digital models, so called BIM models, from reality capture data /Bosché et al. 2015/.

Remote sensing techniques based on laser scanning or photogrammetry are used to map the current outer surface geometry of built structures in a digital format. One sought-after attribute in Scan2BIM for bridges applications is the ability to swiftly acquire data with ease and flexibility. This is particularly valuable because bridge structures are typically large and can include hard-to-reach areas. Moreover, data capture may be affected by traffic. The collected digital data is further processed to produce a set of 3D coordinate points. This set of points is usually referred to as a 3D point cloud, and it will often consist of many millions of points. In addition to point coordinates, data acquisition techniques may also provide attributes such as RGB color or brightness values, both of which may help in the condition assessment of bridges in the later stages of the Scan2BIM process. While remote sensing techniques only capture visible surfaces further details about non-visible building elements can be inferred by automatically integrating "as-planned" data from available legacy construction drawings into the BIM model or by acquiring current "as-is" data through non-destructive testing methods.

Building Information Modelling (BIM) evolved from a tool to design to a collaborative workflow integrating geometrical, topological and functional information stored in a consistent database /Weygant 2011/, /NIBS 2014/, /Azhar et al. 2012/. This rich information repository is accessible and manageable by all stakeholders throughout the whole life cycle of an asset. An essential output of the BIM method is a semantically-rich 3D and object-oriented digital model, so called BIM model /Sacks et al. 2008/. Depending on the life cycle phase this model can be an as-planned/as-designed, as-built or as-is model. Especially, the use of BIM in the operational phase requires as-is models, just like Digital Twins (DTs) that are currently being intensively discussed. DTs are proposed as a coupling of a real asset and its digital representation, where changes in state of one of the representations are communicated and trigger responses into the other /Oliver et al. 2018/. For DTs, BIM models that reflect the actual geometric-semantic state are an important basis which are combined with other models (e. g. numerical models, sensor models) within the Digital Twin.

This article describes a methodical approach to the organisation of the digital twin creation process which consists of several steps. These steps are distributed among the present sections. We first describe the available data sources (Section 2) and how they can be pre-processed (Section 3). To enrich the data, we review several techniques for acquiring semantic understanding (Section 4). We will present unsupervised and supervised learning methods, for completeness. Knowing point relations allows the extraction of geometric properties (Section 5) and the creation of digital models (Section 6). Finally, we present state-of-the-art implementations utilizing the steps described (Section 7). *Fig. 1* provides an overview of the sequential steps and sections.

2 DATA SOURCES

Optical 3D measurement techniques are a powerful tool for the acquisition of 3D object geometries with high spatial resolution. Among manifold techniques, the term comprises photogrammetric multi-view stereo techniques as well as terrestrial laser scanning. These

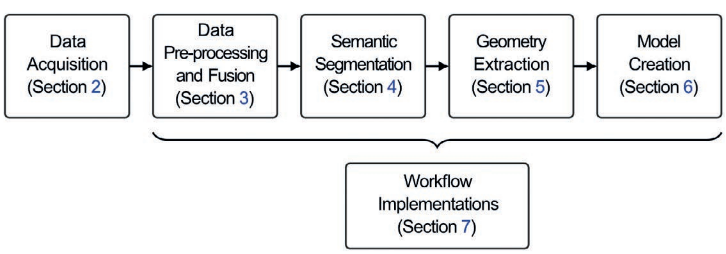

Fig. 1 | Flowchart representation of the Scan2BIM workflow reviewed in this study. Numbers in parenthesis refer to the corresponding sections within this study where further details are found.

two techniques are predestined for complete and accurate data capture of complex objects such as bridges. They may be applied in a stationary terrestrial mode as well as in terrestrial mobile modes or on a UAV (unattended aerial vehicle). Herein, 3D point clouds are an important result of optical 3D measurement techniques, describing object geometry with a very large number of (quasi-randomly distributed) 3D points. Such a 3D point cloud will mostly not be the final result, but form a basis for change analysis or to derive vectorized geometry data to be used for applications such as Digital Twins. Multi-temporal 3D point clouds also form a basis for change monitoring.

2.1 Static and Mobile Laser Scanning

Laser scanning (or LiDAR) techniques are based on laser distance measurement with a scanning device and can be subdivided into airborne and terrestrial laser scanning, with the latter further sectioned in static and mobile modes. Airborne LiDAR using an airplane platform will deliver data which rarely exceed a point density of 20 points/m^2 (corresponding to a point spacing of 20–25 cm) which may for instance be sufficient to add bridges to 3D city models on the basis of LiDAR and cadastral data /Goebbels 2021/. The use of a light-weight laser scanning device on an UAV allows for much higher point densities due to lower flying heights. As airborne LiDAR instruments typically come with range measurement precision in the order of 1–2 cm, the use of differential Global Navigation Satellite System (GNSS) solutions is essential to provide platform position data with an appropriate accuracy. In /Gaspari et al. 2022/ an off-the-shelf system consisting of a DJI Matrice 300 UAV equipped with Zenmuse L1 scanner is used to capture 3D data of a bridge and reported an accuracy of 5–10 cm, with some degradation caused by GNSS signal obstruction underneath the bridge.

Terrestrial laser scanning (TLS) instruments typically come with a distance measurement precision down to a few millimetres and a maximum range of several hundred meters, which makes them very suitable for bridge data acquisition. TLS data for a bridge information model are used by /Mohammadi et al. 2022/, generating CAD models by a slicing technique and achieving accuracies in the millimetre range. As integrated devices, there are also total stations (see Section 2.4) with integrated cameras and scanning devices. TLS data capture comes with some effort due to the necessity of tripod-based data acquisition from many instrument positions, even though multiview fusion is completely automated. Here, vehicle-based mobile laser scanning (MLS) systems may offer more efficient solutions. For instance /Lueangvilai & Chaisomphob 2022/ use a MLS system for annual inspection of a bridge on live traffic.

Even higher flexibility is offered by backpack-based personal laser scanning systems /Liang et al. 2014/. They often come with SLAM (simultaneous localization and mapping) techniques for the geometric registration of successive scans, which are based on the matching of features in successive scans, often supported by IMU (inertial measurement unit), GNSS and possibly by camera measurements /Karam et al. 2020/. Camera data can be used for both, SLAM solutions and the determination for 3D point cloud colour attributes. Personal Laser Scanning (PLS) systems may provide accuracies in the millimetre-range, slightly below the accuracy level obtained in static TLS. In /Gollob et al. 2020/ tripod-borne TLS and PLS are compared in forest inventory applications, which are also characterized by high complexity, showing the PLS data capture is 5.4 times faster. A combination of TLS, PLS and image-based for generating a 3D model of a bridge to be enriched with inspection and load test data is shown by /Previtali et al. 2020/.

Regardless of the platform used, modern laser scanner systems generate very large data sets due to their very high point rates, up to several million points per second. Removing non-object points, reducing the amount of data and deriving the required geometries requires extensive post-processing. In order to reduce the interactive workload, intelligent methods are increasingly used, which are presented in Sections 3, 4 and 5.

2.2 Image-based 3D Object Data Capture Techniques

Photogrammetric data acquisition allows for almost arbitrary scaling of data acquisition devices, ranging from very lowcost devices such as single-chip computer cameras or amateur digital cameras to

Fig. 2 | Dense colorized 3D point cloud (13 Mpts) of a bridge from UAV data and SfM processing /Mader et al. 2015/

specialized high-resolution or high-speed cameras. Cameras may be used hand-held, on a tripod or on UAVs (*Fig. 2*). The fast 3D reconstruction of a historic masonry bridge from low-cost UAV data is shown in /Pepe & Costantino 2021/, in /Tatsuro Yamane & Honda 2022/ 3D models of a bridge are generated using SfM techniques and are used to geo-locate the damage that has been detected by a deep learning approach.

Panchromatic camera images are better suited for pure geometric measurement purposes, as they avoid colour pattern effects in the image data and come with a higher light sensitivity. Nevertheless, most image data used in photogrammetry is RGB imagery due to the nature of off-the-shelf cameras, providing RGB attributes for point cloud colorization and a basis for multi-spectral classification techniques. Beyond RGB, other ranges of the electromagnetic spectrum allow for manifold analyses, for instance, NIR (near infrared) sensors for moss detection and pitting corrosion, or TIR (thermal infrared) sensors for humidity-induced damages. In /Mader et al. 2016/ a study on the potential of RGB, NIR and TIR cameras onboard a UAV for damage mapping on buildings and bridges is shown. Beyond multi-spectral information, /Isfort 2022/ show a study on the use of polarization cameras for the determination of concrete surface moisture.

Proper geometric and stochastic modeling is crucial to exploit the accuracy potential of cameras used as measurement devices /Luhmann et al. 2016/. Subpixel accuracy image analysis techniques allow for image space measurement accuracies down to the order of 0.02 – 0.05 pixels under good conditions. Translated into 3D object space, this corresponds to an accuracy in the order of 1 : 100 000 of object size. However, most cameras come with very significant systematic errors caused by the effects of camera optics, electronics, and mechanics. These errors will often be 1 – 2 orders of magnitude larger than image space measurement accuracy and need to be compensated using proper sensor modeling and simultaneous calibration techniques /Luhmann et al. 2016/. Only in pure 2D monitoring tasks with a static camera, where points to be observed only move over a few pixels in image space, camera calibration can be omitted. An application in a pilot study on 2D bridge deformation measurement, where even accuracies of less than 0.01 pixel, corresponding to 1 : 200 000 of the largest object dimension and validated by inductive gauges, could be achieved by /Albert et al. 2002/.

2.3 Range Cameras and Mobile Devices

As a kind of hybrid between LiDAR devices and cameras, range images allow for the acquisition of panchromatic image data plus a distance value for each pixel, employing modulated LEDs and phase-based time-of-flight measurement techniques /Oggier et al. 2004/. Unlike stereoscopic multi-camera systems, range cameras deliver depth information in real-time without the necessity of stereo-matching, and unlike laser scanning, depth is determined for the whole scene simultaneously /Maas 2008/, however with a maximum range often limited to several meters and an accuracy rather in the order of a centimeter /Westfeld & Maas 2013/. Meanwhile, some high-end mobile phones also come with LiDAR sensors. In combination with the integrated GNSS, gyroscopes and accelerometers, this allows to use them as standalone devices for kinematic 3D data acquisition of complex scenes. The potential of several such devices, yielding an accuracy potential of a few centimeters under good conditions is examined by /Costantino et al. 2022/. Mobile device cameras may also be used for flexible stereo photogrammetric image data acquisition, but come with considerable instabilities of their interior orientation cameras, which require extra effort in calibration /Elias et al. 2020/. Moreover, integrated AI-based techniques for image sharpening may affect the measurement accuracy.

2.4 Total and Multi Stations

In addition to the area-based methods described above, single-point or hybrid methods are also used to generate 3D geometry data as a basis for BIM models. Tachymeters or total stations are well known from conventional surveying in construction. They enable the precise acquisition of geometry data in the form of discrete points. The general workflow of a total station-based data collection from a BIM perspective is shown in /Blankenbach 2018/. While in terrestrial laser scanning the required geometry data is extracted from a large amount of 3D points in the post-processing phase, in tachymetric surveying the geometric discretization of the object must be carried out prior to the survey, as the points required for the reconstruction are measured directly or indirectly. Data acquisition can be divided into three different methods. (i) the acquisition of 3D structure edges, (ii) the direct acquisition of sections, and (iii) 3D modeling. The third method is the most relevant for BIM data acquisition. Recent total stations are characterized by a high degree of digitalization. Integrated coaxial camera systems, target tracking and extensive on-board software are now standard features. These innovations are particularly useful for measuring and documenting damaged areas for BIM model enrichment. In /Lienhart et al. 2017/ an approach for monitoring vibrations on bridges using an image-assisted total station is shown. Camera integration has the advantage of eliminating the need for prisms. In addition to modern total stations, there are also hybrid systems, i. e. high-performance total stations with additional 3D scanning capability. In terms of pure point measurement rate, they are less efficient than current terrestrial laser scanners, but they are more accurate and can be easily stationed in an existing network of checkpoints. This is confirmed by /Fagandini et al. 2017/ in their study on the use of a multi-station in structural monitoring. In particular, the elimination of special targets for alignment and the high accuracy are seen as very positive.

2.5 2D Plans and Stock Data

An additional source of data for data capture of bridges may be digitized historic 2D plans. A drawback herein is in the fact that these plans will only deliver as-planned data. If multiple views are available, such plans may also be used for 3D reconstruction, either standalone or in combination with 3D point clouds or vectorized data captured by laser scanning or stereo photogrammetry. A workflow for the generation of 3D models for BIM models of historical bridges from 2D plans based on image processing for corner detection is described in /Poku-Agyemang & Reiterer 2023b/. An advantage of using plans in the generation of BIM models is in the fact, that they also contain details, that are inaccessible to laser scanning or image-based data capture. Beyond as-built data acquisition, such plans may also support point cloud segmentation and classification tasks /Humblot-Renaux et al. 2023/.

3 DATA PRE-PROCESSING AND FUSION

Data pre-processing involves a range of techniques and operations that are applied to raw data before it is utilized for analysis or other purposes. We chose to name this step pre-processing since the algorithms described in our workflow specifically serve this purpose. However, these algorithms can also be employed in later stages of the process. The objective of data pre-processing is to clean, convert, and organize the data in a manner that is more suitable for subsequent analysis or processing steps. This encompasses the fusion of different data sources or various available point clouds. Ensuring high data quality is crucial as it profoundly affects the dependability and accuracy of the findings.

3.1 Data Fusion

The complementarity of the above-mentioned optical 3D measurement techniques favors the use of data fusion techniques. The data fusion process combines the different data sources to overcome the limitations of the individual 3D measurement techniques. For instance, stereo imaging techniques will usually provide better accuracies in the lateral direction, while LiDAR techniques will be better in the radial direction. Moreover, different sensors will provide different attributes to 3D points, such as RGB or thermal information. The data sources are transformed into the same reference system through the process of registration. There is a wide range of registration algorithms. These range from classical methods such as Iterative Closest Point (ICP) /Besl & McKay 1992/ and its extensions which, for example, introduce the geometric features into the error function to achieve accurate registration /He et al. 2017/. Other possible registration algorithms use high-order graph matching /Zhang & Wang 2018/, Normal Distribution Transform /Biber & Straßer 2003/, /Zaganidis et al. 2017/ or 3D deep learning /Ao et al. 2021/ to solve the registration problem. The registered point cloud can be additionally transformed into a geodetic coordinate reference system using GNSS positions of keypoints within the point cloud /Otepka et al. 2013/. When fusing different data sources it is essential to do so, while maximizing the completeness of the resulting point cloud. Quantifying the level of completeness depending on the use case has been proposed in /Rebolj et al. 2017/. Furthermore, planning the point cloud acquisition positions and methods to maximize completeness has been the subject of multiple studies, where a comprehensive review can be found in /Aryan et al. 2021/.

In /Dabous & Feroz 2020/, a study is presented on the condition monitoring of concrete bridges, including the detection of rebar corrosion, delamination and cracking, integrating close-range photogrammetry, terrestrial laser scanning, infrared thermography and ground-penetrating radar. Various sources, including different laser scanners and analog 2D plans, are utilized in /Poku-Agyemang et al. 2023a/ to construct a comprehensive point cloud of a bridge, going beyond outer shell geometries.

3.2 Data Pre-processing

Before utilizing the point cloud, there exist various methods to pre-process it. A frequently utilized pre-processing technique involves downsampling to decrease the number of data points. The two most commonly used techniques are random downsampling and voxel-grid downsampling. The former method is characterized by high speed and the ability to generate a point cloud with a fixed number of points. On the other hand, the latter technique produces a more uniform point cloud by utilizing a voxel grid with a constant voxel size and a single point per voxel.

To account for potential object or scene rotation in relation to a particular coordinate system, the process of axis alignment is used. Many computational geometry problems can be solved more efficiently when dealing with axis-oriented objects, like axis-aligned rectangles or line segments /Martens & Blankenbach 2020/. These objects have a clear and simple structure that enables faster processing. Axis alignment can be achieved with the use of PCA, as discussed in Section 5.1. The largest eigenvector of the matrix corresponds to the direction of the principal axes and can be utilized to rotate the object to align with the global axes.

Another crucial step is to identify and eliminate any outlier points. Outlier points, in this context refer to data points that deviate significantly from the overall pattern or distribution and are attributed to

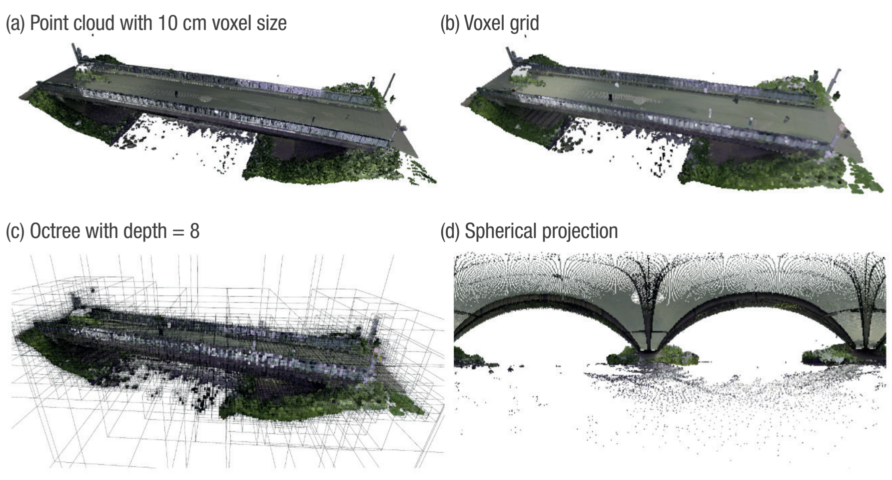

Fig. 3 | Visualization of point clouds and different possible representations environmental factors, measurement errors or, in the case of optical sensors, reflective or transparent surfaces. The Statistical Outlier Removal /Grubbs 1969/ method involves computing the mean and standard deviation of distances between every data point and its k nearest neighbors. Any points with distances exceeding a specific threshold are then removed. The Local Outlier Probabilities /Kriegel et al. 2009/ algorithm computes a density-based result for every data point based on the density of its k nearest neighbors. Data points with significantly lower density than their neighbors are identified as outliers and eliminated. Both methods utilize the k nearest neighbors, which are commonly approximated with a kd-tree for purposes of efficiency /Muja & Lowe 2009/. Another frequently utilized method for detecting outliers is RANSAC, which will be further discussed in Section 5.3.

It is important to note that there are several ways to represent point clouds. Alternatively, instead of using the point cloud, which is essentially a set of points, the point cloud can be transformed into a 3D grid of voxels. The point cloud can be projected onto a plane and displayed as a 2D image. Various representations are depicted in *Fig. 3*. Due to the variety of available representations, alternative algorithms may be utilized with significant benefits. Particularly in the context of 2D projection, these algorithms can be accelerated through the reduction of one dimension.

4 SEMANTIC SEGMENTATION

Semantic segmentation assigns a label to each pixel or point, indicating the category, type or class of object to which it belongs. This labeling offers a thorough comprehension of the scene content, critically important in tasks such as autonomous driving, scene understanding, and object recognition. The deep understanding and interpretation of the visual scene that semantic segmentation demands make it a challenging process. It entails the identification and differentiation of objects or regions within an image or point cloud that may differ in shape, size, orientation, and color. Another crucial point to consider is the differentiation between semantic segmentation and instance segmentation. While for the former, only a label needs to be assigned to each point showing its class, for the latter, it is necessary to distinguish between different objects belonging to the same class. The introduction of additional individual instances makes the whole process even more complex, and it is often divided into two steps. Therefore, we will first describe semantic segmentation and possible approaches to it, and finally, in Section 4.4, we will discuss the instantiation. There are multiple methods for accomplishing 3D segmentation and often different representations, as described in Section 3.2, are used. These different representation options result in different approaches to performing the segmentation. While some of the methods described in this section are specific to bridge infrastructure, most are applied in the general context of point clouds. For an overview of their application to bridges we describe particular workflows in Section 7.

4.1 Unsupervised Rule-based Approaches

In rule-based approaches, the segmentation process is based on a set of predetermined conditions or thresholds. These conditions are often derived from empirical observations about the data being segmented or domain knowledge. The rules are designed to capture certain patterns, features or relationships which can either be used as input for supervised machine learning models or on its own to distinguish different objects or regions of interest in an unsupervised way.

4.1.1 Primitive Fitting for Segmentation

Although primitive fitting is mainly a technique for finding geometric shapes in point clouds, if the scene and objects of interest consist primarily of primitive shapes, it can also serve as a segmentation method. Therefore, in the following, we describe such use cases. Applications for geometric modeling are then discussed in Section 5. Primitive fitting involves approximating a geometric form, such as a plane, sphere, cylinder, or cone, to a group of 3D points in a point cloud. It frequently serves as the initial phase in object recognition and segmentation from point clouds. The objective is to recognize straight-forward geometric shapes that can symbolize the object and then exploit this information to extract the object from the point cloud. A standard way for object inlier determination are RANSAC-based methods, which we will discuss in more detail in Section 5.3.

In /Maalek et al. 2019/ a framework is proposed for extracting primary structural components, which include columns, slabs, and rebars, from point clouds in regular rectangular reinforced concrete structures. Major bridge components are detected within point clouds in /Lu et al. 2019/. The deck assembly is separated from the pier assemblies using a slicing algorithm. Afterwards, the pier caps and girders are detected and segmented based on their surface normal, oriented bounding boxes, and density histograms. This approach achieves segmented and labeled point clusters in a top-down manner.

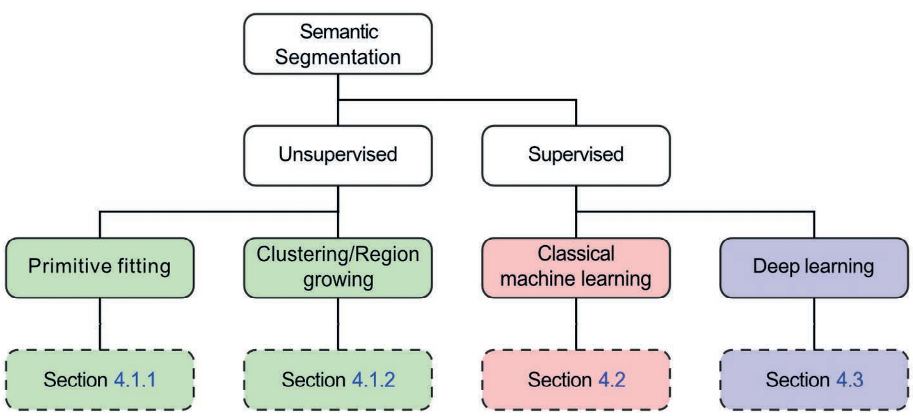

Fig. 4 | General semantic segmentation overview

4.1.2 Clustering

Clustering algorithms, in general, are powerful tools in data analysis as they group similar data points into coherent clusters, revealing inherent structures in complex data sets. In the field of the analysis of 3D point clouds, clustering algorithms play a crucial role in 3D segmentation. By detecting points that share common geometric attributes such as proximity, orientation, curvature, or colour, clustering algorithms identify points that correspond to specific components, such as pillars, for instance. The resulting clusters represent underlying geometric classes, enabling the reconstruction of structural elements with increased detail. For the clustering of point clouds into meaningful parts many different algorithms are available, however, here we will only focus on a small selection of widely used algorithms, such as k-means, Density-based spatial clustering of applications with noise (DBSCAN), or region growing.

k-Means is a centroid-based clustering algorithm that divides a data set into k distinct, non-overlapping clusters, see /Mirkin 2005/. It starts by randomly selecting cluster centers and then assigns each data point to the nearest cluster center based on a chosen distance metric, e.g. the Euclidean distance. After assignment, the cluster centers are recalculated as the mean of the data points within each cluster. This process of assignment and center updating iterates until convergence. k-Means is widely used for point cloud segmentation by partitioning points into clusters based on spatial proximity, making it effective for extracting geometric primitives like planar surfaces.

DBSCAN, proposed by /Ester et al. 1996/, is a density-based clustering algorithm that categorizes data points into clusters based on the point density and connectivity. It is particularly effective in handling point clouds with varying densities and irregularly shaped clusters, making it a valuable tool for 3D point cloud analysis in asset reconstruction and other scenarios. Two key parameters are defined: a parameter ϵ that specifies the radius of a neighborhood with respect to a core point and the minimum number of points *minPts* within a distance ϵ to form a core point. DBSCAN starts with a core point and expands the cluster by connecting neighboring core points within the distance ϵ. Points that are not core points but fall within ϵdistance of a core point are assigned to the cluster as well. For 3D asset reconstruction, DBSCAN can reveal irregular structural components such as pillars, railings, decorative elements, or other structural details that may not conform to standard shapes. /Czerniawski et al. 2018/, for example, presents an approach for the extraction of planar objects using DBSCAN in a six dimensional clustering space. Region growing is a seed-based clustering algorithm that iteratively expands clusters from seed points, see /Bali & Singh 2015/ for a general overview of region based clustering methods. It starts with a set of seed points and iteratively adds neighboring points that satisfy predefined similarity criteria. These criteria are often based on geometric features such as normal vectors, curvature, colour, or spatial proximity. If a neighboring point meets the criteria, it is added to the cluster and the process continues. The growth of the region is stopped when no more points can be added to the cluster based on the chosen criteria. This algorithm is suitable for segmenting point clouds into coherent regions and thus for identifying planar surfaces, edges and other geometric features. Employing algorithms like k-means, DBSCAN, or region growing, the segmentation of point clouds into meaningful segments is facilitated and thus enabling accurate geometric extraction and asset reconstruction.

4.2 Classical Machine Learning

While the previous methods focused entirely on unsupervised methods the following sections deal with methods which utilize labeled datasets. This is the case if for each point in a point cloud $\boldsymbol{x}_i \in \mathbb{R}^{\text{features}}$ a corresponding semantic class $\boldsymbol{y}_i \in \mathbb{N}^{\text{classes}}$ is provided by an annotator. The goal of the machine learning model then becomes to learn a function which maps unseen examples to the observed semantic classes. Therefore, supervised methods have the benefit of not only providing segments of points with similar geometric properties but also assigning these with a semantic label. In principle, supervised methods are divided into "classical" that require the prior manual feature extraction and "deep learning", which automate feature extraction. In the following, we start by describing classical approaches, while the subsequent Section 4.3 describes deep learning methods.

Due to their reliance on precomputed features, research in the semantic segmentation of point clouds has mainly focused on extracting rich feature sets which are able to efficiently describe the local geometry. Some of the information needed is already contained in the individual point and is often used without the need for further computation – this includes the absolute point height, color information captured by cameras and signal intensity in the case of LiDAR. In /Mallet et al. 2011/ this is extended by fitting a parametric function to the full-waveform LiDAR data. In order to capture the geometric properties of surfaces and objects, the context surrounding individual points becomes relevant. For this purpose the general workflow can be divided into 4 steps: 1) Neighborhood selection, 2) Feature extraction, 3) Feature selection, 4) Supervised classification. The neighborhood is the set of points surrounding the point of interest which are going to be included in the computation and is mostly defined as either all points within a certain radius r or a fixed number of the k-nearest neighbors. Based on these neighbors a description of the local geometry can then be either computed using covariance-based features /Kawashima et al. 2012/, /Weinmann et al. 2015/ or histogram-based features such as spinning images /Johnson & Hebert 1999/, SHOT descriptors /Hutchison et al. 2010/, /Xu et al. 2018/ and 2D-accumulation maps /Monnier et al. 2012/. Covariance-based features use statistical methods, described in Section 5.1, to capture geometrical characteristics of the neighborhood, which are often perceived as interpretable qualities – e.g. planarity, linearity, sphericity, etc. Histogram-based features instead discretize the relative position and orientation /Hutchison et al. 2010/ of surrounding points in order to describe the pattern of the neighborhood.

Covariance-based features are heavily dependent on the size of the neighborhood and a fixed radius may be sub-optimal for datasets with objects of various scales /Weinmann et al. 2015/. They propose an optimal neighborhood strategy which adapts its size based on the eigenentropy of mutually exclusive geometric features. The same paper goes further to perform feature selection, resulting in run time reductions, as well as provide an overview of the performance of all typical classifiers on the resulting feature sets such as Random Forrest, Naive Bayes, K-Nearest Neighbor, Support Vector Machines, and others.

Even though information from neighbors is captured by the feature extraction step, the points are still treated individually by the classifier which may result in noisy or erroneous segmentation maps. To circumvent this, statistical contextual models can take into account the relationships between points by modeling them as conditional probabilities – in /Niemeyer et al. 2017/ the predictive performance is improved by modeling mid and long-range relationships between points using Conditional Random Fields.

4.3 Deep Learning

Deep neural networks (deep learning) are currently the most prominent supervised segmentation methods. A general overview of deep learning applied to point clouds can be found in /Guo et al. 2020/. The approaches can be divided into the above representations and fusion-based methods that use a combination of different representations. *Fig. 5* illustrates the subdivision into the different representations and the corresponding approaches.

4.3.1 Projection-based

An intuitive way of transferring the success of convolutional neural networks (CNNs) with images to point clouds is to view the point clouds as projections on a surface. This enables the use of already established 2D-CNNs architectures and multi-view feature fusion. The most common projections are the spherical and the bird's eye view. The spherical view is often referred to as a depth image because the Cartesian coordinates of each point (X, Y, Z) are transformed into spherical coordinates (θ, ϕ, d) and the angles are used as pixel positions holding the depth value d. /Milioto et al. 2019/, /Cortinhal et al. 2020/ achieved quite remarkable results with this projection on the /Behley et al. 2019/ data. The bird's eye view is a top-down perspective, similar to how the scene would appear if viewed from above, like a bird flying overhead. This view is created by projecting the 3D points onto a horizontal plane, effectively removing depth or height information and preserving spatial relationships only along the X and Y axes. In /Zhang et al. 2018/ an occupancy grid is created and the gravitational axis serves as the feature channel. PolarNet /Zhang et al. 2020/ adopts the approach by using a polar grid instead of a Cartesian one. In addition to other possible projections, there is also the possibility of combining several viewing angles or even different projections /Gerdzhev et al. 2020/, /Alnaggar et al. 2020/. However, through the projection process, some of the spatial information is lost. Furthermore, the choice of projection planes may heavily influence recognition performance and occlusion in 3D may impede accuracy. One advantage, however, is their overall small parameter size and quick inference which makes them applicable for self-driving where latency is an issue.

4.3.2 Discretization-based

Some authors quantize the point clouds in discrete 3D voxel grid which allows the use of 3D convolutions but often leads to massive computational and memory costs due to many voxels being naturally empty. Some solutions exist that aim to mitigate this: e.g., by using sparse convolutions, where the kernel is only applied to occupied voxels /Graham et al. 2018/ or by dividing the scene into hierarchical partitions with octrees /Riegler et al. 2017/. The authors of /Zhu et al. 2020/ propose an alternative to the conventional voxel grid by introducing a 3D cylindrical partition. They utilize an asymmetric 3D convolution that focuses on improving the horizontal and vertical responses while aligning with the distribution of object points. Nevertheless, these designs may still suffer from some detail information loss depending on the choice of the voxel size.

4.3.3 Point-based

A rather successful approach has been the design of deep network architectures, that operate on continuous points directly. Those make use of permutation-invariant operators, such as shared multilayer perceptron (MLPs) and pooling to aggregate features over a set – PointNet /Qi et al. 2016/. Additionally, these also include a learned transformation that aligns point clouds to a canonical space, making the network transformation invariant. One drawback of the shared MLP approach is that they handle each point independently, which limits their capability to capture local features. The MLP method was developed further in PointNet++ /Qi et al. 2017/ to compute point-wise features by applying it recursively on small neighborhoods around sampled points, progressively abstracting the point cloud.

Graph convolution-based architectures aim to solve the problem of local features, that is inherent to the classical shared MLP by interpreting the point cloud as a graph. Dynamic graph CNNs create an operation – EdgeConv, which generates permutation invariant edge features based on each point's relationship to its local neighbors /Wang et al. 2019/. The

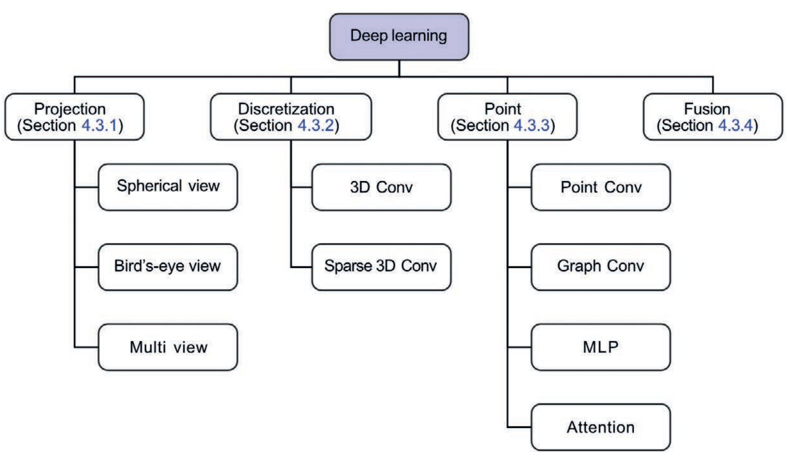

Fig. 5 | Overview of the types of deep learning architectures for different 3D representations modularity and straightforward implementation make it easy to integrate into the basic version of PointNet.

Point convolution is another family of recent architectures, which apply continuous convolutions directly to a 3D point set. Inspired by image convolutions KPConv /Thomas et al. 2019/ defines the kernel explicitly as a set of kernel points with either a flexible or rigid configuration. The lack of a spatial data structure, required for the convolution operator to compute correspondences between its kernel and the input point cloud is solved by computing a weighted average of each kernel point's neighbors' features based on the Euclidean distance between them. The resulting matrix of input features is multiplied by a learned kernel matrix which encodes the density of the kernel's neighborhood as a feature vector of the point the kernel currently operates on. The authors argue that their method generates superior local features, due to the ability of KPConv to better encode local surface deformation. In PointConv /Wu et al. 2018/, convolution kernels are treated as nonlinear functions based on the local coordinates of 3D points.

These functions include both weight and density components. The weight functions are acquired through multi-layer perceptron networks, while the density functions are acquired through kernel density estimation, all in relation to a specific point. Efficient computation of weight functions enables the network to enhance its performance and scalability.

The self-attention operation has recently emerged from the field of natural language processing, due to its ability to capture long-range context and thus create a richer feature representation /Vaswani et al. 2017/. In general, attention is a set of learned relevance scores which describe the importance between all elements in a sequence for the interpretation of any other given element of the same sequence. It is therefore more flexible than the convolutional operator which only focuses on a direct neighborhood but also comes at the cost of high memory and computational demands. Further works have shown that treating pixels in an image as a sequence makes the application of attention layers possible which has been shown to produce state-of-the-art results in image analysis /Yuan et al. 2021/. In the field of point cloud processing attention mechanisms have also been used to produce high-performing models with the main issue being a trade-off between large receptive field and computational demand /Engel et al. 2021/, /Lai et al. 2022/, /Park et al. 2022/. Engel et al. first introduce local vector self-attention applied on the k-nearest neighbors of each points /Engel et al. 2021/. Lai et al. enlarge the receptive field by including sparsely sampled key points far away from the query point /Lai et al. 2022/. Most recently voxel hashing has been applied in conjunction with local self-attention to relieve a large amount of the computational overhead /Park et al. 2022/.

4.3.4 Fusion-based

To balance the disadvantages of one type of representation with the advantages of another, there are different ways to combine them. The transfer of hidden knowledge at both the point and voxel representation occurs through knowledge distillation from Point-to-Voxel /Hou et al. 2022/. To improve the learning efficiency of affinity distillation, they proposed the supervoxel partition and the difficulty-based sampling strategy. The combination of complementary projections with a learnable fusion based on the point-based KPConv approach is presented in /Kellner et al. 2022/. In /Xu et al. 2021/ a fusion framework with multiple and mutual information interaction between all three representations presented is introduced. To enable the interaction between the different representations, a special Range-Point-Voxel indexing system is developed using hash mapping.

4.4 Instance Segmentation

Fig. 6 clarifies the difference between semantic segmentation and instance segmentation. Semantic segmentation involves pixel-level classification, while instance segmentation distinguishes between individual objects in addition to pixel labeling. This is why the abutment and railing, for instance, are colored differently, since they belong to different instances in the second image.

As previously noted, the process of object instantiation is often divided into two separate stages, although there are several methods that propose it in an end-to-end, learnable way. In /Wong et al. 2020/, a bird's eye view is processed by a 2D convolutional feature pyramid network to create a category-agnostic embedding space to cluster points into instances using DBSCAN. The segmentation is based on the spherical projection and the instances are created using DBSCAN with a weighted distance function in /Chang & Chen 2021/. SGPN /Wang et al. 2018/ constructs a feature similarity matrix, based on the feature output generated by PointNet++, to group points of similar features into instances. PointGroup /Jiang et al. 2020/ utilizes a 3D U-Net backbone for obtaining semantic labels and proposes point clustering with dual coordinate sets. Additionally, PointGroup creates ScoreNet to predict instance scores. Using a 3D U-Net backbone to extract semantics or point-wise features is a common strategy. All of the following examples use this basic structure, even if it is slightly modified, but customize the subsequent instantiation step. The SoftGroup /Vu et al. 2022/ module enables multiple class associations for each point to alleviate issues arising from semantic prediction errors. OccuSeg /Han et al. 2020/ performs graph-based clustering guided by object occupancy signal for more accurate segmentation outputs. HAIS /Chen et al. 2021b/ extends PointGroup and introduces a set aggregation and intra-instance prediction to refine the instance at the object level. Mask3D /Schult et al. 2023/ and SPFormer /Sun et al. 2022/ make use of Transformer decoders to generate object instances. Within the

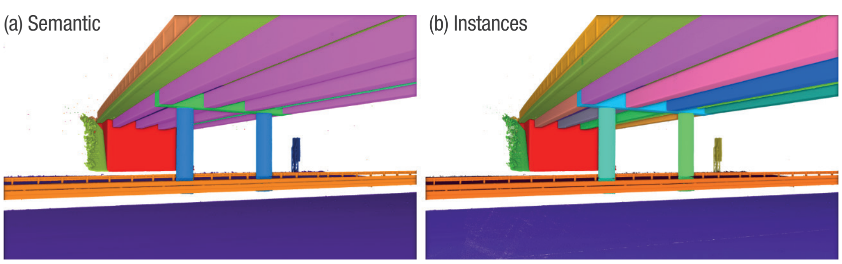

Fig. 6 | Visualization of point clouds with semantic and instance segmentation former, instance queries are learned by iteratively attending to features in the point cloud at multiple scales. Combined with point characteristics generated by the backbone, the queries for the specific instance generate all instance masks simultaneously. The latter, first use a superpoint pooling layer to pool potential point-wise features into superpoints, followed by a query decoder where learnable query vectors can capture instance information by superpoint cross-attention.

4.5 Comparison and Discussion

Rule-based approaches are often intuitive and computationally efficient compared to more complex methods such as deep learning, for example, because they do not require as much training data. However, they may have limitations in dealing with complex or ambiguous scenarios, and their performance depends heavily on the quality and appropriateness of the predefined rules. Furthermore, inferring the necessary domain knowledge is often non-trivial. In /Ponciano et al. 2021/, the rule knowledge-based approach and deep learning are compared and their advantages and disadvantages are discussed in more detail.

Moreover, it is noteworthy that approaches already exist that mitigate the weaknesses of each approach while exploiting the strengths of the other. In /Landrieu & Boussaha 2019/, a neural network with low complexity learns deep embeddings of local geometry and radiometry for 3D points. These embeddings are used to oversegment the point cloud by formulating a graph partitioning problem. Alternatively, a neural network approach has been proposed to decide whether to include a point in a new region, allowing for the implementation of a region growing method for object segmentation without any assumptions of their shapes and sizes /Chen et al. 2021a/.

5 GEOMETRY EXTRACTION

Geometry extraction is a central processing step in which relevant geometric objects in 3D point clouds are detected as automatically as possible and parameterised for further processing. This step plays a crucial role in the detection of contours and special features of individual structural elements and thus serves as a basis for creating comprehensive BIM entities that are as accurate, complete and reliable as possible. In the following, we will present several common methods for geometry extraction in the automated creation of BIM models from 3D point clouds in order to explain this crucial step in more detail. We will primarily concentrate on techniques such as Principal Component Analysis (PCA), Hough Transform, Random Sample Consensus (RANSAC), and the use of primitive fitting to extract geometric properties, see *Fig. 7*.

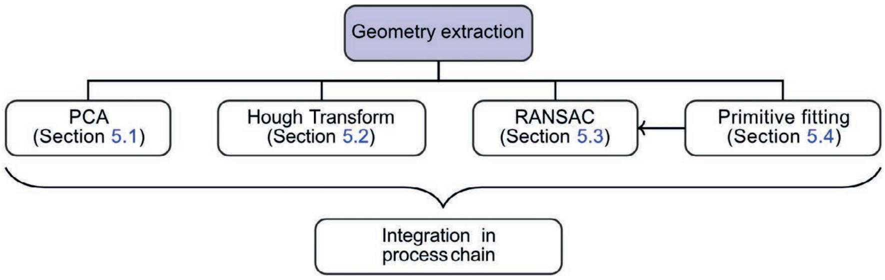

Fig. 7 | Geometry extraction techniques from 3D point clouds

5.1 Principal Component Analysis

Principal component analysis (PCA) was first introduced by /Pearson 1901/ and later refined by /Hotelling 1933/ and is an analytical method utilized to gather statistical information in large data sets. In this method, the number of variables in the data is reduced by replacing them with a smaller set of uncorrelated variables. These new variables, linear combinations of the original data, are carefully selected to maximize variance while ensuring accurate data representation and are commonly referred to as principal components. They summarize the essence of the data set and can be interpreted as approximations of lines, planes, or hyperplanes (in k-dimensional space) that span the original data and provide spatial orientation. The principal components of a data set can be found by solving an eigenvalue/eigenvector problem using, for example, a Singular Value Decomposition (SVD). In general, PCA is an indispensable tool for reducing the dimensionality of data, which is widely used in fields such as computer vision, remote sensing or, as already mentioned, for the 3D reconstruction of complex structures. /Maalek et al. 2019/ employs embedded PCA to extract planar and linear elements in their approach for automatically recognizing structural objects.

PCA is a useful tool for extracting distinctive geometric features from a local set of points /West et al. 2004/. These properties are obtained by establishing a link between the eigenvalues λ_i acquired from the eigenvalue decomposition, where $\lambda_1 > \lambda_2 > \lambda_3$. The geometric features are explicitly defined in /Weinmann et al. 2013/ and presented in *Tab. 1* and visualized in *Fig. 8*.

Planarity	f_p	$(\lambda_2 - \lambda_3)/\lambda_1$
Linearity	f_l	$(\lambda_1 - \lambda_2)/\lambda_1$
Sphericity	f_s	λ_3/λ_1
Surface Variation	f_v	$\lambda_1/(\lambda_1 + \lambda_2 + \lambda_3)$
Sum of eigenvalues	f_Σ	$\lambda_1 + \lambda_2 + \lambda_3$
Omnivariance	f_o	$(\lambda_1 \cdot \lambda_2 \cdot \lambda_3)^{1/3}$
Eigentropy	f_e	$-\sum^{3} \lambda_i \cdot \ln(\lambda_i)$
Anisotropy	f_a	$(\lambda_1 - \lambda_3)/\lambda_1$

Tab 1 | Geometrical features calculated by the eigenvalues of the covariance matrix

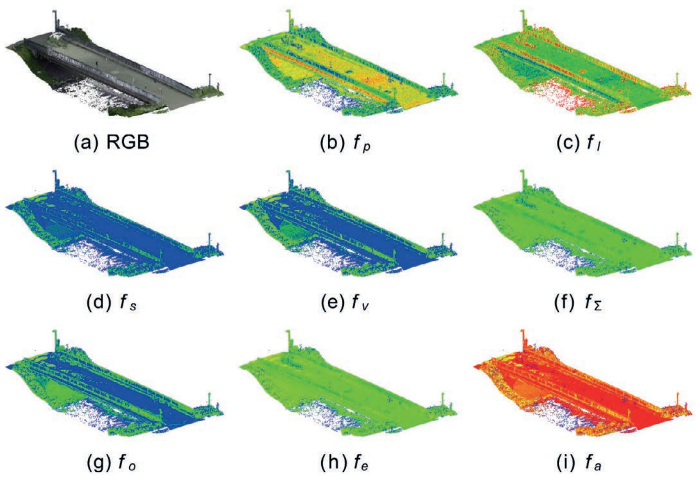

Fig. 8 | Calculated features from *Tab. 1*. The color is scaled from blue, green, yellow to red for the feature value of range [0, 1]. We used the implementation from /Kellner et al. 2023/ for calculation.

Since point cloud properties are dependent on the chosen neighborhood size, /Demantké et al. 2011/ investigates the optimal neighborhood size for each point. Additionally, they demonstrate that the approach can be utilized for classification and segmentation purposes and is effective for point clouds acquired from various sensor systems. /Poux & Billen 2019/ expands on this concept by utilizing geometric properties as a foundation with relationships and topology between voxel entities to generate a knowledge-based decision tree to allow for indoor classification. Meanwhile, /Hackel et al. 2016/ uses these characteristics to identify contour points that highlight the shape of a curve or surface in which there is an abrupt change in direction or curvature. Based on calculated properties, a binary random forest classifier predicts whether a point is a contour point.

5.2 Hough Transform

The Hough transform, a pioneering technique in image and signal processing, has its origins in the 1960s. Developed by Hough /Hough 1962/ to detect lines in binary images, this transformation method has evolved significantly over the years and is used not only in traditional image analysis, but also in more complex scenarios such as the analysis of 3D point cloud data. Originally, the Hough transform was developed for the detection of lines, but its versatility and adaptability have allowed it to be used for the detection of other geometric shapes such as circles, ellipses, and planes. In general, it provides a robust approach to detecting linear and planar structures within the data. By representing each point in the cloud as a parameterized entity in a so-called parameter space or Hough space, the Hough transform accumulates votes for specific parameters that correspond to the presence of lines or planes. In the case of line detection, for instance, each data point generates a curve in the Hough space based on angle and distance parameters. Points lying on the same line in the original space intersect at common points in the Hough space. The point with the most intersections of curves in Hough space represents the most probable line. By selecting appropriate threshold values, the line can be detected and extracted. This technique is versatile and extends to detecting various shapes by altering the parameterization. In the context of 3D point clouds, the concept remains analogous, yet the parameter space expands to encompass extra dimensions.

For the detection of planes in 3D point clouds, for instance, we get a 3D Hough space, wherein each point of the point clouds generates a sinusoidal plane. It is important to note that while the Hough Transform is a powerful tool, it can be computationally expensive, especially in high-dimensional spaces. Various optimizations and adaptations have been proposed over the years to make it more efficient for different applications and types of data. The approach for the 3D reconstruction of interior wall surfaces presented in /Adan & Huber 2011/, for example, extract planar wall surfaces using the Hough transform. Also /Okorn 2010/ apply the Hough transform to extract walls for automatic modeling of floor plans. A detailed overview of the plane detection in 3D point clouds using the Hough transform is given by /Borrmann et al. 2011/.

5.3 RANSAC

The Random Sample Consensus (RANSAC), introduced in /Fischler & Bolles 1981/, is an iterative method for eliminating outliers in a dataset as a pre-processing step to estimate the parameters of a mathematical model. Its basic working principle involves iteratively selecting subsets of data points, fitting a model to each subset, and evaluating the quality of the fit. Therefore, the data points are divided into inliers and outliers based on a specified tolerance. If a certain predefined threshold is met by the number of inliers, the model is accepted to describe the data well. Otherwise, another sub-sample will be tested.

RANSAC is particularly well suited for detecting primitive shapes in 3D point clouds containing a significant amount of noise and/or outliers. Many approaches have been derived from its basic concept. There are quite a few adaptations regarding runtime, accuracy or robustness, and a general overview is given in /Choi et al. 2009/. Here, we focus on examples where RANSAC is used to obtain geometric objects. In /Schnabel et al. 2007/, for example, a high performance RANSAC algorithm is developed to extract a variety of different geometric primitives, such as planes, spheres, cylinders or cones, which is directly implemented in the approach to extract walls for an automatic reconstruction of building models proposed by /Ochmann et al. 2016/. In addition, /Bassier & Vergauwen 2020/ extract wall axes from 3D point clouds using RANSAC to reconstruct BIM wall objects. Fitting as many planes as possible with RANSAC and creating surfaces from them using the intersections is done by /Nan & Wonka 2017/. An optimal subset of the candidate surfaces is selected by a binary linear programming formulation, producing a lightweight, manifold, and watertight reconstructed model. In addition, there are first approaches that combine neural networks with RANSAC to guide the sampling of the minimal set /Brachmann & Rother 2019/.

One of the drawbacks of RANSAC is the considerable computational costs, depending on the model chosen, the ratio of inliers to outliers, and the desired probability of identifying a model that is free of outliers. Accordingly, it is not possible to determine complex geometric objects with a high outlier-to-inlier ratio. Point clouds in infrastructure scenarios typically consist of millions of points, only a fraction of which are associated with actual geometric structures. In this context, deep learning techniques play a role in assisting RANSAC methods by minimising the number of potential candidate points.

5.4 Fitting Geometric Primitives

The concept of fitting primitives, as explained in Section 4.1.1, can also enable the extraction of geometric models. For instance, one can segment points within a basic shape while simultaneously providing a parametric description of the shape. Primitive geometric shapes including lines, planes, spheres, cylinders, and cones can be easily described using a set of specific parameters and are frequently found in human-created objects. Some examples of geometric primitives are shown in *Fig. 9*. The parameters required to describe the geometry are highlighted in red.

As mentioned above, RANSAC is a commonly used method to determine the inliers and outliers corresponding to a defined object. Knowing the inliers allows the model parameters to be determined in a subsequent step using, for instance, least squares approaches.

However, there are alternative approaches available for parameter determination. A neural network based on PointNet++ is proposed in /Li et al. 2019/ to predict per-point properties, and then the primitive parameters are estimated using a differentiable model estimator. In /Tulsiani et al. 2017/, the problem is further simplified by describing each object with volumetric primitives (rigidly transformed cuboids). A neural network based on 3D convolutions is then trained to assemble arbitrary 3D objects from the given primitives. In this way, 3D shapes can be described in a highly abstracted and simplified way.

5.5 Fitting Parametric Surfaces

Parametric surfaces use parameterized equations to describe points on a surface in three-dimensional space. The parameters vary over a range, and the equations define how the coordinates of points on the surface change as the parameters change. Bézier or B-spline curves, for example, are used to describe such freeform curves and surfaces. An n-degree Bézier curve is a parametric curve defined by $n + 1$ control points. Adjusting one control point affects the whole curve. This is the main difference to B-splines, where a change only has a local effect and only part of the curve needs to be recalculated. Non-Uniform Rational B-spline (NURBS) are an extension using the B-spline basis function and allow for non-uniform spacing of weighted control points.

Fitting such curves is more complex than fitting primitives as described in Section 5.4. For example, the number of control points needed to achieve the required accuracy is not known in advance. For this reason, least squares methods are usually used to obtain them /Piegl & Tiller 1996/. However, this does not allow the curve to

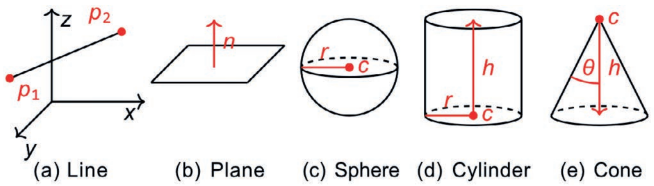

Fig. 9 | Geometric primitives and corresponding parameters

be optimized locally. For this reason, an iterative process is often used to adjust the number of control points /Deng & Lin 2014/. In /Park & Lee 2007/, so-called dominant points are used to implement the concept of adaptive curve refinement, which means that fewer curve segments are generated for flat regions, but more for complex regions. Instead of optimizing the B-spline curve, the weights are iteratively optimized with the least square method in /Wang et al. 2022/.

How well NURBS can match bridge structures can be seen in /Barazzetti et al. 2016/. Although the generation is not automated, it demonstrates that the traditional BIM approach for buildings can also be applied to a complex medieval bridge. Other use cases include bridge deformation analyses in terms of a distance metric of two models recorded at different times. In /Kermarrec et al. 2020/ the automatic fitting of a B-Spline is additionally supported by a stochastic model of the laser scanner. The mathematical approximation of the surface accounts for variations in the noise levels of the scanner, thereby allowing for a more accurate distance metric between the two models, compared to a point cloud-based distance.

5.6 Discussion

The algorithms used to extract geometric primitives, such as RANSAC, PCA, Hough transform, or clustering algorithms, are usually not used individually but are often integrated in a process chain. While the methods themselves can produce good results, they are often not scalable to large and complex scenes containing many different objects. For example, a semantically segmented point cloud can drastically reduce the inlier-outlier ratio, which has a direct impact on the complexity of the RANSAC algorithm. A semantically segmented beam, for instance, can be fitted much more easily by a geometric primitive than using the whole scene. In the same way, the misclassified points can be detected as outliers. Therefore, by combining these techniques with other methods like deep learning approaches, researchers and engineers can achieve more accurate and comprehensive reconstructions of the 3D scenes or objects from point cloud data.

6 MODEL CREATION

Once the relevant objects have been identified, instantiated if necessary, and geometrically modeled, the object must be converted to the desired data representation. There is a wide range of possible ways to represent 3D data, that all serve different purposes in fields such as computer graphics, computer-aided design (CAD) or computer vision, each with a different way of expressing and modeling

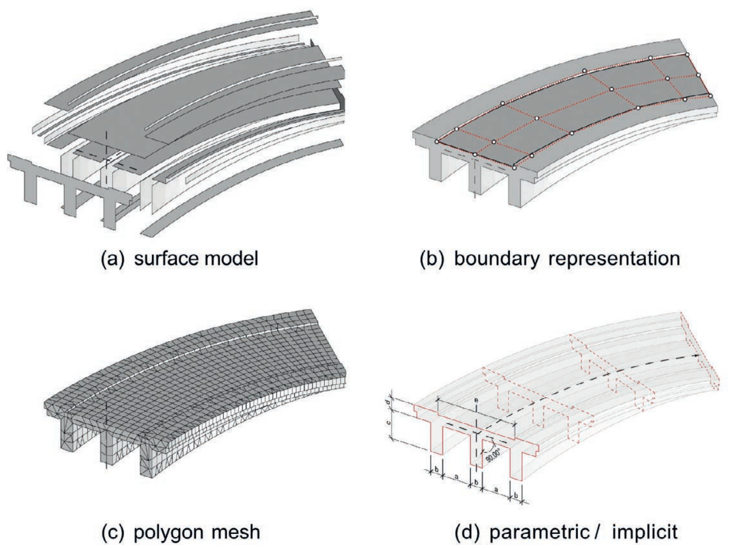

Fig. 10 | Different geometric representations of a bridge deck

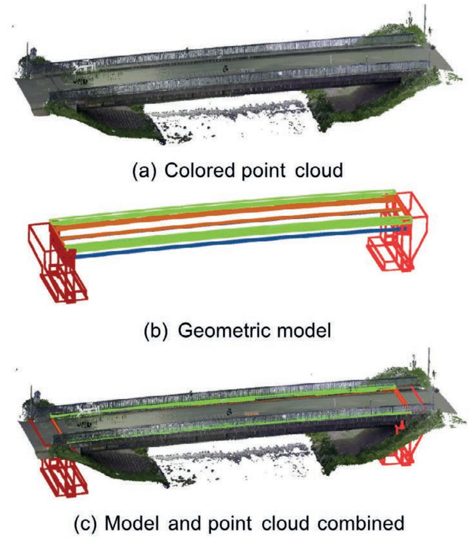

Fig. 11 | Visualization of point cloud and geometric model

shapes, structures or objects in a mathematical or computational form. Specifically, standard CAD geometric representations are often categorized into explicit and implicit where the choice of representation is dependent on the use case /Borrmann et al. 2018/. In the following, we briefly describe each with its relevance to the Scan2BIM workflow, followed by some concrete examples of bridge infrastructure.

6.1 Explicit Representation

The most common way of representing geometry is by explicitly defining a wire-frame, a surface or a solid. Wireframe modeling uses only lines and points to define edges and vertices. In this way, a skeleton of the object, without surfaces or volumes, describes the object. Connecting multiple edges in a closed loop can be used to form surfaces, which is the basis of surface modeling (*Fig. 10a*). Further connecting surfaces to form a watertight shape is called solid modeling and is the preferred representation for multiple BIM-related downstream tasks, such as quantity takeoff. The most widely used forms to represent surfaces and solids include Polygonal representations (Mesh) – *Fig. 10c* and Boundary representations (BRep) – *Fig. 10b*. Polygon representations follow the idea of using flat polygons with straight edges (usually triangles or quads) to approximate the surfaces of an object. *Fig. 11* shows a simple model describing the main components of the bridge. In this example, each bridge component is modeled by using the wireframe representation.

Mesh representations are used as the basis for numerical analyses, such as finite element models, and are beneficial due to their simplicity and integration into many software pipelines. Their drawback, however, is the increase in required storage capacity, when curved surfaces are to be represented accurately.

Boundary representations (BReps) can be viewed as an extended variant of polygonal geometry, which allows for an analytical description of edges and surfaces e. g. planes or polynomial functions (see Sections 5.4 and 5.5). Compared to meshes, smooth curved surfaces described by BReps often have a significantly lower amount of faces with a comparable or better adhesion to a desired surface. They are therefore the preferred method in BIM applications and have been included in multiple international standards, such as STEP /ISO 10303-21:1994/ and IFC /ISO 16739-1:2018/ (see Section 6.3).

6.2 Implicit/Parametric Representations

An alternative way of representing geometry is by specifying the basic modeling steps needed to construct the model. As the actions taken to create a desired geometry are preserved, subsequent changes to the base geometry are made possible, which is a favorable quality in iterative workflows such as design. A simple standardized type of implicit representation is that of Constructive Solid Geometry (CSG), which can define hard-surface objects as a binary tree, where leaves are composed of primitive shapes (see Section 5.4) and nodes represent the performed boolean operations (intersection, union, and difference). Other variants include proprietary representations in CAD software, which often store the complete history of complex geometric operations, such as revolution, sweeping, lofting and extrusion, while additionally enabling the use of parameterized 2D drawings used in the process /Autodesk n. d. a/. As discussed by /Sacks et al. 2008/ a parametrized object is one whose geometry and associated data are defined by parameter values and it is overall constrained to follow a certain set of predefined rules. The rules are pre-established according to the type and properties of the modeled component. Some geometric rules that can be established are dimensional horizontal, vertical, perpendicular or coincident constraints /Mafipour et al. 2023a/. The representation consistency is assured as automatic checks and updates in the whole model take place, whenever a component of the model changes the parameters that define it.

Although implicit representations are important for planning workflows they introduce an additional modeling and representation complexity and lead to a limited interoperability between software, which restricts their use in Scan2BIM to only certain edge cases. Some examples include the definition of bridges and roads as a series of cross-sections and a corresponding alignment curve – *Fig. 10d*, where the alignment curve can serve for better relative positioning of building parts or assets. Specific implementations can be found in Section 7.

6.3 Data Model Representation

The already described steps of data acquisition, pre-processing, semantic segmentation and geometry fitting enable an accurate and whenever necessary flexible geometry representation of existing infrastructure assets. However, the BIM process strives to facilitate numerous workflows throughout the lifecycle of the assets, which either require or produce additional non-geometric, alphanumeric information, the administration of which requires a data model to enable full and sustainable compatibility. These data models commonly have a hierarchical object-oriented structure, that aims to group objects spatially and semantically. The geometric representation of building elements often serves the additional role of containing and mediating this information and is therefore associated with a specific semantic type.

A vendor-neutral standardized /ISO 16739-1:2018/ schema that enables a detailed and holistic digital description of the built environment is the Industry Foundation Classes (IFC) schema authored by the buildingSMART International (bSI) organization /buildingSMART International 2023/. A schema in this context can be viewed as an agreement on common concepts that prescribe the instantiation of models and objects /Cerovsek 2011/. A particular object type may be associated with a predefined set of alphanumeric properties but is not restricted to a specific geometric representation, allowing a single instantiated object to have multiple alternative representations, depending on the use case. The IFC schema may be expressed using several formats among which the STEP physical file format with extension .ifc is the most widely used in practice /buildingSMART International 2022a/. Although IFC was initially proposed for buildings, as of version IFC4.2, the schema explicitly supports the description of bridge structures too /Borrmann et al. 2019/, /buildingSMART International 2022b/. The extension of the schema introduces some key methods for the localization and grouping of elements, such as support for georeferencing, positioning of building elements along an alignment curve and object containers such as “substructure” and “superstructure”. As reported in the official initial proposal of the IFC extension for bridges /Borrmann et al. 2019/, it was decided that the IFCBridge extension considered the most common types of bridges such as slab bridges or girder bridges. Although as mentioned in the proposal, bridges of truss or arch type were expected to be representable as well. This can be verified in the work of /Justo et al. 2023/, where several IFC entities model a section of a bridge of truss type, principally considering the IFC4.2 version reported byits authors. More recently, additional support for further infrastructure projects, such as road, rail and ports has been added in version 4.3, which as of 2024 has been standardized within the /ISO 16739-1:2024/.

7 PROTOTYPICAL IMPLEMENTATIONS OF A SCAN2BIM APPROACH

There are already a number of approaches to what a possible workflow might look like to transform the captured 3D data into a BIM model. All share the same idea of abstracting the captured data into a digital model that represents the physical elements, geometry, relationships and attributes of the building in a detailed and parametric manner, that would allow for further adjustments and be compatible with established BIM standards. The traditional workflow can be broken down into two parts: the survey workflow and the BIM modelling workflow, of which the second is associated with the predominant portion of the workload /Rocha et al. 2020/, motivating research into the direction of automation.

A BIM model is built upon a library of predefined object types (also sometimes called families), which are then instantiated within the context of the model. An object instance may inherit some of its properties from its type but often carries additional information that is specific to its occurrence and state. Furthermore, single building parts are often related to other parts of the model through objectified relationships, which allow the definition of a spatial hierarchy as well as the semantics of structural connections. The result of this process is a geometric-semantic model, which can be enriched throughout an asset’s life cycle with data from different damage detection methods (see *Fig. 12*) and be used as the basis for structural analysis in the form of Finite Element Analysis (FEA) models /Fedorik et al. 2016/. Specific approaches which automate the modeling process are mostly domain-specific due to the required geometry parametrization, modeling precision and semantics. Therefore, in the following we look into processes, that focus on the generation of BIM models for bridges in specific.

While typical building construction is mostly limited to flat planar regions, bridges tend to have more complex geometry, due to e.g. non-straight horizontal alignments or complex girder types. The application of BIM in the sector of bridge design and documentation is still in development and some approaches focus solely on non-automated bridge BIM model generation processes. In /Mohammadi et al. 2022/ a case study is presented for the creation and quality evaluation of a Scan2BIM of a complex cable-stayed bridge. The process of setting the alignment, profile slicing and modeling was achieved within the Tekla Structures Software, following a dimensional comparison with the as-planned CAD model. In /Girardet & Boton 2021/ a parametric definition is created using the Grasshopper visual programming language /Grasshopper 3D n.d./. The definition aims to describe and model a large variety of bridge designs with a single parametric program.

When faced with the challenge of automating the Scan2BIM workflow, some authors address a holistic end-to-end approach, while most focus on single steps (see *Tab. 2*). Based on the

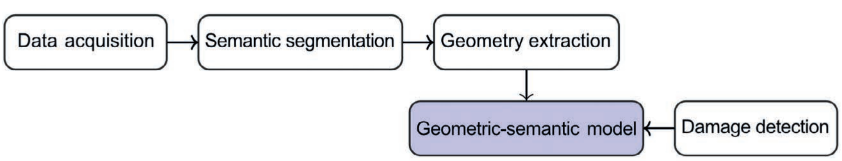

Fig. 12 | Digital twin workflow for infrastructure

	Step 1	Step 2	Step 3	Step 4	Step 5
/Lee et al. 2020/		✓	✓	✓	
/Quin et al. 2021/	✓	✓	✓	✓	✓
/Truong et al. 2022/	✓				
/Mafipour et al. 2023b/	✓	✓	✓	✓	
/Lu et al. 2019/	✓				
/Girardet et al. 2021/					✓
/Yan et al. 2021/	✓	✓	✓	✓	
/Lu & Brilakis 2019/		✓	✓	✓	✓
/Martens et al. 2023/	✓				
/Xia et al. 2023/	✓				
/Mohamaddi et al. 2023/		✓	✓	✓	✓

Tab. 2 | Overview of Scan2BIM approaches in the bridge domain

commonalities of the analysed methods the BIM modelling workflow for bridge infrastructure can be summarised in the following sequential steps:

- Step 1: Semantic and/or instance segmentation;
- Step 2: Definition of main axis or alignment per object;
- Step 3: Slicing of object along said axis;
- Step 4: Estimation of 2D-profiles at sliced positions;
- Step 5: BIM model generation based on profiles.

Differences in the overall approach to the first step (semantic segmentation) allow the described methods to be roughly grouped into bottom-up rule-based, top-up rule-based and deep learning approaches. Bottom-up approaches start from low-level features, from which higher-level features, such as fitted planes, normals, etc. can be derived. Top-down approaches instead break down a complex segmentation problem into a series of sequential segmentation tasks which are simpler and can be solved by heuristic algorithms. Lastly, deep learning methods learn to derive high-level features based on low-level features from the input.

7.1 Bottom-up Rule-based Approaches

After aligning a chosen bridge to the world coordinate system /Lee et al. 2020/ use a variation of the RANSAC algorithm to fit planes to the pre-segmented point cloud of the deck. The relative distances and orientations of the planes to each other are used to derive the optimal values for the parameters of a standard bridge cross-section. /Truong-Hong & Lindenbergh 2022/ approach semantic segmentation as a two-step process consisting of surface extraction and surface classification. The surface extraction is achieved by means of voxel and cell-based region growing, while the classification of individual surfaces is approached by introducing contextual knowledge, such as the probable shape, orientation, location and minimal dimension of each semantic class. /Qin et al. 2021/ combine a top-down pre-segmentation of ground from bridge points, followed by primitive fitting of cuboids and cylinders. To achieve the latter, the authors simplify the 3D fitting problem into a 2D problem by determining the extrusion axis of primitive shapes and slicing the geometry along these axes. The parameters of rectangles are determined through corner detection on the slices, while the circle center and radius are determined using least-squares. As a final step, the profiles are imported into the visual programming environment “Dynamo” /Autodesk n.d. b/ to create a parametric BIM model, making use of the semantics provided by the software. However, the method notably relies on a restrictive assumption of only modeling primitive shapes.

7.2 Top-down Rule-based Approaches

In /Lu et al. 2019/ known topological constraints in reinforced concrete slab and beam-slab bridges are utilized in order to perform top-down semantic segmentation. The bridge point cloud is aligned to the world coordinate system using PCA. Afterwards, slices are extracted, which are then classified based on extracted geometric features such as the relative height of the slice. The process is performed once to segment pier assemblies and again to segment single piers. Subsequently, density histograms are used to segment girders from the deck. Similarly, /Yan & Hajjar 2021/ restrict their search space on steel girder bridges but don’t assume a straight shape. Instead, an iterative algorithm is presented which can derive a curved horizontal alignment constructed of linear segments that follow the shape of the bridge. This alignment then serves as a reference for subsequent steps, such as slicing the bridge perpendicularly to the direction of the alignment at regular intervals and for distinguishing between longitudinal and transverse elements based on their relative orientation to the alignment axis. Sub-clouds are projected on the slices and a 2D RANSAC is used to fit lines to the projected points (see Section 4.3.1) to form a line drawing of the cross-section. This allows finding the best fitting candidate from a set of predefined I-profiles by minimizing their deviation to the cross-section. In /Lu & Brilakis 2019/ the authors assume the semantic segmentation has been accomplished by an upstream process and solely focus on extracting an IFC model from the segmented point cloud. After aligning the principal direction to the x-axis the horizontal and vertical alignments are each estimated by fitting a second-degree parabola using least squares. Slicing methods are then used to estimate the cross-sections of the slab, piers, pier caps, and girders. For most components, a 2D-concave hull is used to describe the cross-section, with the exception of the girders, where template matching with existing pre-cast catalogs is preferred instead.

7.3 Deep Learning Approaches

While each of the previously described rule-based approaches focused on a predefined type of bridge, deep learning models can normally be applied to a dataset with a much greater variance. This is achieved by stacking multiple hidden layers which each learns progressively higher-level features (see Section 4.3). Deep learning methods, however, come with the drawback of requiring a large amount of training data in order to create a robust classification model.

In /Martens et al. 2023/ the authors address the lack of widely available datasets in the point cloud domain and construct a hybrid approach for semantic segmentation, which combines point cloud geometry with images. They make use of a pre-trained Mask R-CNN deep neural network for image segmentation and project semantic labels onto the point cloud. Subsequently, region growing and feature-based filtering are applied based on geometric information in order to improve the quality of the result. Other methods, such as /Xia et al. 2022/ apply multi-layer perceptron after extracting hand-crafted features based on the Signature of Histograms of OrienTations (SHOT) /Salti et al. 2014/. Finally, the segmentation results are refined using DBSCAN (see Section 4.1.2). In /Mafipour et al. 2023b/ the deep learning architecture RandLA-Net is used for the semantic segmentation with a custom-made spatial encoder. The authors introduce a parametric prototype model (PPM), which takes inspiration from a parametric BIM family, where parts of the geometry are controlled by parameters, that can take a range of continuous or discrete values. From the output of the semantic segmentation, the most appropriate PPM is chosen, after which a particle swarm optimization algorithm is used to derive the best parameters of the PPM, based on the projected point cloud at each slice. In contrast to primitive fitting, the PPMs can define a geometric profile with arbitrary complexity while maintaining interoperability with BIM conventions.

A different approach is taken in /Hu et al. 2021/, where Hu et al. demonstrate an end-to-end deep learning method for deriving a BIM model of a cable-stayed bridge from a combination of a photogrammetry-based point cloud and the images used as an input to the structure from motion algorithm (see Section 2.2). In this instance a combination of a simple PointNet backbone is used together with a multi-view CNN in order to extract a joint embedding of the data, which is then decoded recursively by a binary tree network, resulting in a reconstructed geometric model, consisting of primitives for simple shapes and meshes for complex parts. Notably, the linear, rotational, and symmetric repetition of a single type of object is captured by the method which is valuable for describing structures such as steel trusses in bridges.

8 CONCLUSION AND RESEARCH PERSPECTIVES

The automatic creation of a geometric-semantic model is a challenging problem, which shows great potential for improvement, especially in light of recent technological developments. Approaches for semantic segmentation are currently predominantly rule-based, which come in either a top-down or bottom-up form. These methods achieve good results but deal with design restrictions, such as focusing on a certain construction type or assuming a straight horizontal alignment. So far, rule-based approaches promise robustness when dealing with data coming from different sensors and possible occlusions /Yan & Hajjar et al. 2021/. Although the creation of an actual BIM model was rarely implemented it could be seen, how the results of intermediate steps used for the semantic segmentation could be useful additions to the BIM modeling workflow, such as the grouping of elements in assemblies /Lu et al. 2019/ and the derivation of the horizontal alignment /Yan & Hajjar et al. 2021/, which is a crucial element for the geometric representation of a bridge according to the IFC standard /ISO 16739-1:2018/. On the other hand, deep-learning-based approaches so far mostly focus on semantic segmentation and although some high-level features are extracted, these are not human-readable and don't currently contribute to the creation of any relationships or to the geometry extraction, which have to be handled in isolation by subsequent processes. Nevertheless, promising potential has been demonstrated for decoding the latent information, captured by deep learning models in order to facilitate end-to-end semantic-geometric modeling.

Even though deep learning methods present a promising approach for even the most complex infrastructure assets, there remains a challenge with their application in practice, due to their lack of interpretability and sometimes unexpected results, especially given the high variety of data sources and available bridge types. Furthermore, another challenge is the need to combine different types of data, such as images, point clouds and 2D design drawings.

8.1 EXplainable Artifical Intelligence (XAI)

The lack of understanding of the capabilities of a system, including the AI-based ones, can effectively hinder its adoption. For certain AI-based systems, the expectations from their deployment include not only achieving a high performance but also getting some understanding of their functionality to verify its results. This broader model characterization is particularly crucial in safety-critical domains where malfunctions can lead to severe consequences /Castelvecchi 2016/, /Weld & Bansal 2018/, /European Parliament 2023/. Enhancing our understanding of the functional capabilities of the presented AI-based semantic segmentation approaches would refine our characterization of the discussed Scan2BIM workflow, unlocking the benefits of its adoption.

EXplainable Artificial Intelligence (XAI) refers to the group of techniques aimed to clarify or detail the learned functions of AI-based approaches /Arrieta et al. 2019/. A common strategy towards this goal is the attribution of importance to the inputs for a given model output. An interesting implementation for point cloud understanding is that of /Tan & Kotthaus 2022/. The implementation extends the LIME /Ribeiro et al. 2016/ technique by approximating the decision boundary of the PointNet /Qi et al. 2016/ with the use of perturbed inputs. Regularized linear regression is used to get a model easier to understand from the human perspective compared to the original model. This work proposes too a verification of the simpler model by observing the correlation between the prediction score and the inclusion of positive or negative deemed relevant points in a given point cloud classification task.

In /Lapuschkin et al. 2016/ several benefits from the adoption of XAI techniques are discussed. For example, the gained knowledge from the functionality of the AI-based models could eventually be determinant in the choice among models with similar performances. Another possible use of XAI techniques is that of finding biases within datasets when the model learns from spurious signals. Lastly, XAI techniques could also be used to debug models improving their generalization ability. Overall, the adoption of XAI techniques presents itself as a worthwhile research opportunity.

8.2 Uncertainty Estimation

Another essential quality for the implementation of data-driven systems is the ability to recognize when the decision-making process inevitably fails. Especially with deep learning models this has become an increasingly demanded topic, as it has been shown, that the uncertainty metric associated with a prediction or decision of a neural network is often inaccurate /Hein et al. 2019/, particularly for inference cases, which are dissimilar from the training data. This has motivated research from other safety-critical fields such as autonomous driving /Kendall & Gal 2017/ and medical imaging /Kwon et al. n. d./ to explore uncertainty quantification of data-driven approaches. These methods may provide a more reliable uncertainty metric and for separating between data uncertainty (or aleatoric) from model (or epistemic) uncertainty, providing additional insights into where possible improvements can be made or which inference cases can be trusted.

Uncertainty estimation is also relevant for the construction and infrastructure sectors, where scenes can have a high diversity, and various types of data acquisition methods are used (e. g. laser scanning, photogrammetry, etc.). Therefore, if the model is blindly trusted some of the scanned objects may be falsely interpreted or missed /Vassilev et al 2024/. Alternatively, if an under-confident uncertainty metric is trusted some objects which may have been correctly classified are dropped. Because as-is" models play a vital role in the functions of a digital twin, it becomes vital that their generation does not trade efficiency for safety and reliability.

8.3 Multi-modality

3D point clouds and image data may be enriched with various other data to maximize the information content of a as-is BIM model or a digital twin. An obvious step is using RGB information from geometrically referenced images as attributes of a 3D point cloud. Beyond RGB values, information from the non-visible image spectrum may be of large interest, especially thermal image information. An extension of this is impulse thermography where variations of the decay of a thermal pulse in a thermal camera image sequence are analyzed to detect damages underneath a surface. Ground penetration radar systems may also be used for this purpose.

To derive wall thickness of structures, it is necessary to measure interiors or hollow spaces as well as the facade. Terrestrial laser scanning for outdoor areas and mobile respectively personal laser scanning for indoor areas may be combined effectively for this purpose. In addition, inaccessible construction areas at high altitudes may be mapped using UAV-borne sensors.

8.4 Integrating Heterogeneous Data Sources

The fusion of different data sources, such as 3D point clouds, 2D plan data, CAD drawings, measurements for specific components, and 3D bridge models, is a difficult and complex task in terms of Scan2BIM approaches, but one that should not be neglected. This integration process presents several significant challenges that require specific methodologies.

One of the challenges in the integration of heterogeneous data sources is the registration into a common reference frame. Matching features, such as geometric primitives (see Section 4.1.1), across different types of data sets are a prerequisite for the registration process. Notably, the extraction of these features varies for each data source and requires different processing of the data respectively. The feature matching is succeeded by determining the approximate values for the transformation parameters in a combined adjustment using the method of least squares.

Precision measures play a critical role in this process. While stochastic information is readily available for some data sources, such as point clouds obtained from terrestrial laser scanning, they may be absent or less precise for others, e. g. 2D plan data. In these cases, assumptions must be made about non-statistical uncertainties. This can be done by assigning individual variance components to each group of observations. These components are iteratively adjusted to determine relative precision information. Thus, on the basis of the stochastic information of 3D point clouds, the accuracy of fit and reliability of the other data sources can be determined.

Reliability measures play an important role in assessing the quality of the adjustment results. This implies the detection of significantly erroneous or incomplete input data using statistical test procedures well-known from the field of geodesy as well as removing such erroneous data. On this basis, a geometric model reconstruction can be performed, including semantic information regarding individual components of the model, derived from statistical tests.

ACKNOWLEDGMENTS

This project has received funding by DFG (German Research Foundation) within the framework of “SPP 2388: Hundert plus – Verlängerung der Lebensdauer komplexer Baustrukturen durch intelligente Digitalisierung” and by the BMDV as part of the mfund-project “mdfBIM+ – Teilautomatisierte Erstellung objektbasierter Bestandsmodelle mittels Multi-Daten-Fusion”.

REFERENCES

Adan, A.; Huber, D. (2011): 3D reconstruction of interior wall surfaces under occlusion and clutter. In: 2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, 275–281.

Albert, J.; Maas, H.-G.; Schade, A.; Schwarz, W. (2002): Pilot studies on photogrammetric bridge deformation measurement. In: 2nd IAG Com. IV Symposium on Geodesy for Geotechnical and Structural Engineering, 2002.

Alnaggar, Y. A.; Afifi, M.; Amer, K.; Elhelw, M. (2020): Multi projection fusion for real-time semantic segmentation of 3D LiDAR point clouds. https://arxiv.org/abs/2011.01974.

Ao, S.; Hu, Q.; Yang, B.; Markham, A.; Guo, Y. (2021): Spinnet: Learning a general surface descriptor for 3D point cloud registration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2021, 11753–11762.

Arrieta, A. B.; Rodríguez, N. D.; Ser, J. D.; Bennetot, A.; Tabik, A.; Barbado, A.; García, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; Chatila, R.; Herrera, F. (2919): Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. https://arxiv.org/abs/1910.10045.

Aryan, A.; Bosché, F.; Tang, P. (2021): Planning for terrestrial laser scanning in construction: A review. In: Automation in Construction 125(2021), 103551.

Autodesk, Inc. (n. d. a): Fusion 360 – 3D CAD, CAM, CAE, & PCB Cloud-Based Software. https://www.autodesk.com/products/fusion-360/overview (31. 10. 2023).

Autodesk, Inc. (n. d. b): Revit.

Azhar, S.; Khalfan, M. M. A.; Maqsood, T. (2012): Building information modelling (BIM): now and beyond. In: Australasian Journal of Construction Economics and Building 12(2012), 15 – 28.

Bali, A.; Singh, S. N. (2015): A review on the strategies and techniques of image segmentation. In: 2015 Fifth International Conference on Advanced Computing and Communication Technologies, 113 – 120.

Barazzetti, L.; Banfi, F.; Brumana, R.; Previtali, M.; Roncoroni, F. (2016): BIM from laser scans … Not just for buildings: Nurbs-based parametric modeling of a medieval bridge. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences III-5(2016), 51 – 56. https://isprs-annals.copernicus.org/articles/III-5/51/2016.

Bassier, M.; Vergauwen, M. (2020): Unsupervised reconstruction of building information modeling wall objects from point cloud data. In: Automation in Construction 120(2020), 103338. https://www.sciencedirect.com/science/article/pii/S0926580520309183.

BASt – Bundesamt für Straßenwesen (2023): Zustandsnoten der Brücken. https://www.govdata.de/web/guest/daten/-/details/zustandsnoten-der-bruckenfe8db (29. 09. 2023).

Behley, J.; Garbade, M.; Milioto, A.; Quenzel, J.; Behnke, S.; Stachniss, C.; Gall, J. (2019): A dataset for semantic segmentation of point cloud sequences. https://arxiv.org/pdf/1904.01416.

Besl, P.; McKay, N. D. (1992): A method for registration of 3-d shapes. In: IEEE Transactions on Pattern Analysis and Machine Intelligence 14(1992)2, 239 – 256.

Biber, P.; Straßer, W. (2003): The normal distributions transform: a new approach to laser scan matching. In: Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453) 3(2003), 2743 – 2748. https://api.semanticscholar.org/CorpusID:18365794.

Blankenbach, J. (2018): Building surveying for as-built modeling. In: Borrmann, A.; König, M.; Koch, C.; Beetz, J. (Eds.): Building Information Modeling: Technology Foundations and Industry Practice. Springer, Cham, Switzerland, 393 – 411.

Borrmann, D.; Elseberg, J.; Lingemann, K.; Nüchter, A. (2011): The 3D hough transform for plane detection in point clouds: A review and a new accumulator design. 3DR Express 2(2011), 3. https://link.springer.com/article/10.1007/3DRes.02(2011)3.

Borrmann, A. König, M.; Koch, C.; Beetz, J. (Eds.) (2018): Building Information Modeling: Technology Foundations and Industry Practice. Springer, Cham, Switzerland.

Borrmann, A.; Muhic, S.; Hyvarinen, J.; Chipman, T.; Jaud, S.; Castaing, C.; Dumoulin, C.; Liebich, T.; Mol, L. (2019): The IFC-bridge project – extending the ifc standard to enable high-quality exchange of bridge information models. In: Proceedings of the 2019 European Conference on Computing in Construction.

Bosché, F.; Ahmed, M.; Turkan, Y.; Haas, C. T.; Haas, R. C. G. (2015): The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and bim: The case of cylindrical map components. In: Automation in Construction 49(2015), 201 – 213.

Brachmann, E.; Rother, C. (2019): Neural-Guided RANSAC: Learning where to sample model hypotheses. https://arxiv.org/abs/1905.04132.

buildingSMART International (2022a): IFC formats. https://technical.buildingsmart.org/standards/ifc/ifc-formats (31. 10. 2023).

buildingSMART International (2022b): IFC release notes. https://technical.buildingsmart.org/standards/ifc/ifc-schema-specifications/ifc-release-notes (13. 11. 2023).

buildingSMART International (2023): Industry foundation classes (IFC). https://technical.buildingsmart.org/standards/ifc (31. 10. 2023).

BVBS – Bundesministerium für Verkehr, Bau und Stadtentwicklung (2013): Bauwerksprüfung nach DIN 1076 – Bedeutung, Organisation, Kosten. Berlin.

Castelvecchi, D. (2016): Can we open the black box of AI? In: Nature 538(2016), 20 – 23.

Cerovsek, T. (2011): A review and outlook for a ‘building information model’ (bim): A multi-standpoint framework for technological development. In: Advanced Engineering Informatics 25(2011), 224 – 244.

Chang, T.; Chen, T.-H. (2021): RangeSeg: Range-aware real time segmentation of 3D LiDAR point clouds. In: IEEE Transactions on Intelligent Vehicles (June 2021). DOI: 10.1109/TIV.2021.3085827.

Chen, J.; Kira, Z.; Cho, Y. (2021a): LRGNet: Learnable region growing for class-agnostic point cloud segmentation. In: IEEE Robotics and Automation Letters 6(2021)2, 2799 – 2806. https://ieeexplore.ieee.org/document/9364361.

Chen, S.; Fang, J.; Zhang, Q.; Liu, W.; Wang, X. (2021b): Hierarchical aggregation for 3D instance segmentation. https://arxiv.org/abs/2108.02350.

Choi, S.; Kim, T.; Yu, W. (2009): Performance evaluation of ransac family. In: Proceedings British Machine Vision Conference, Sept. 7 – 10, 2009, London, UK.

Cortinhal, T.; Tzelepis, G.; Erdal Aksoy, E. (2020): SalsaNext: Fast, Uncertainty-Aware Semantic Segmentation of LiDAR Point Clouds. In: Lecture Notes in Computer Science 12510, 207 – 222.

Costantino, D.; Vozza, G.; Pepe, M.; Alfio, V. S. (2022): Smartphone LiDAR technologies for surveying and reality modelling in urban scenarios: Evaluation methods, performance and challenges. In: Applied System Innovation 5(2022)4, 63, https://doi.org/10.3390/asi5040063.

Czerniawski, T.; Sankaran, B.; Nahangi, M.; Haas, C.; Leite, F. (2018): 6D dbscan-based segmentation of building point clouds for planar object classification. In: Automation in Construction 88(2018).

Dabous, S. A.; Feroz, S. (2020): Condition monitoring of bridges with non-contact testing technologies. In: Automation in Construction 116(2020). https://doi.org/10.1016/j.autcon.2020.103224.

Dawkins, O.; Dennett, A.; Hudson-Smith, A. (2018): Living with a digital twin: Operational management and engagement using IoT and mixed realities at UCL’s here east campus on the Queen Elizabeth Olympic Park. In: GISRUK 2018, Leicester, UK.

Demantké, J.; Mallet, C.; David, N.; Vallet, B. (2011): Dimensionality based scale selection in 3D LiDAR point clouds. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII-5/W12(2011), 97 – 102. https://isprs-archives.copernicus.org/articles/XXXVIII-5-W12/97/2011.

Deng, C.; Lin, H. (2014): Progressive and iterative approximation for least squares b-spline curve and surface fitting. In: Computer-Aided Design 47(2014), 32 – 44. https://www.sciencedirect.com/science/article/pii/S0010448513001528.

Elias, M.; Eltner, A.; Liebold, F.; Maas, H.-G. (2020): Assessing the influence of temperature changes on the geometric stability of smartphone and raspberry pi cameras. In: Sensors 20(2020)3, 643. https://doi.org/10.3390/s20030643.

Engel, N.; Belagiannis, V.; Dietmayer, K. (2021): Point transformer. In: IEEE Access 9(2021), 134826 – 134840.

Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. (1996): A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. AAAI Press, Washington, DC, 226 – 231.

European Parliament (2023): Artificial intelligence act. https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html (06.10.2023).

Fagandini, R.; Federici, B.; Ferrando, I.; Gagliolo, S.; Pagliari, D.; Passoni, D.; Pinto, L.; Rossi, L.; Sguerso, D. (2017): Evaluation of the laser response of Leica Nova Multistation MS60 for 3D modelling and structural monitoring. In: Gervasi, O. et al. (Eds): Computational Science and Its Applications – ICCSA 2017. Lecture Notes in Computer Science 10407. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-319-62401-3_8.

Fedorik, F.; Makkonen, T; Heikkilä, R. (2016): Integration of BIM and FEA in Automation of Building and Bridge Engineering Design. In: 33th International Symposium on Automation and Robotics in Construction, July 2016, Auburn, AL.

Fischler, M. A.; Bolles, R. C. (1981): Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. In: Communications of the ACM 24(1981)6, 381–395. https://doi.org/10.1145/358669.358692.

Gaspari, F.; Ioli, F.; Barbieri, F.; Belcore, E.; Pinto, L. (2022): Integration of UAV-Lidar and UAV-Photogrammetry for Infrastructure Monitoring and Bridge Assessment. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLIII-B2-2022, 995–1002. https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-995-2022.

Gerdzhev, M.; Razani, R.; Taghavi, E.; Liu, B. (2020): Tornado-net: Multiview total variation semantic segmentation with diamond inception module. https://arxiv.org/abs/2008.10544.

Girardet, A.; Boton, C. (2021): A parametric BIM approach to foster bridge project design and analysis. In: Automation in Construction 126(2021), 103679.

Goebbels, S. (2021): 3D reconstruction of bridges from airborne laser scanning data and cadastral footprints. In: Journal of Geovisualization and Spatial Analysis 5(2021)1. https://doi.org/10.1007/s41651-021-00076-9.

Gollob, C.; Ritter, T.; Nothdurft, A. (2020): Forest inventory with long range and high-speed personal laser scanning (PLS) and simultaneous localization and mapping (SLAM) technology. In: Remote Sensing 12(2020)9, 1509. https://doi.org/10.3390/rs12091509.

Graham, B.; Engelcke, M.; Van Der Maaten, L. (2018): 3D Semantic Segmentation with Submanifold Sparse Convolutional Networks. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018, 9224–9232.

Grasshopper 3D (n.d.): https://www.grasshopper3d.com (21.02.2024).

Grubbs, F. E. (1969): Procedures for detecting outlying observations in samples. In: Technometrics 11(1969), 1–21. https://api.semanticscholar.org/CorpusID:120469443.

Guo, Y.; Wang, H.; Hu, Q.; Liu, H.; Liu, L.; Bennamoun, M. (2020): Deep learning for 3D point clouds: A survey. In: IEEE Transactions on Pattern Analysis and Machine Intelligence. https://arxiv.org/abs/1912.12033.

Hackel, T.; Wegner, J. D.; Schindler, K. (2016): Contour detection in unstructured 3D point clouds. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1610–1618.

Haardt, P.; Holst, R. (2008): The German approach to bridge management: current status and future development. In: Transportation Research Circular E-C128: Tenth International Conference on Bridge and Structure Management. Transportation Research Board of the National Academies, Buffalo, New York, 3–15.

Han, L.; Zheng, T.; Xu, L.; Fang, L. (2020): Occuseg: Occupancy-aware 3D instance segmentation. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2937–2946.

He, Y.; Liang, B., Yang, J.; Li, S.; He, J. (2017): An iterative closest points algorithm for registration of 3D laser scanner point clouds with geometric features: In: Sensors 17(2017)8, 1862. https://doi.org/10.3390/s17081862.

Hein, M.; Andriushchenko, M.; Bitterwolf, J. (2019): Why ReLU networks yield high-confidence predictions far away from the training data and how to mitigate the problem. https://arxiv.org/abs/1812.05720.

Hotelling, H. (1933): Analysis of a complex of statistical variables into principal components. In: Journal of Educational Psychology 24(1933), 498–520. https://api.semanticscholar.org/CorpusID: 144828484.

Hou, Y.; Zhu, X.; Ma, Y.; Loy, C. C.; Li, Y. (2022): Point-to-voxel knowledge distillation for LiDAR semantic segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition, 2022, 8479–8488.

Hough, P. V. (1962): Method and means for recognizing complex patterns. https://www.osti.gov/biblio/4746348.

Hu, F.; Zhao, J.; Huang, Y.; Li, H. (2021): Structure-aware 3D reconstruction for cable-stayed bridges: A learning-based method. In: Computer-Aided Civil and Infrastructure Engineering 36(2021)1, 89–108.

Humblot-Renaux, G.; Jensen, S. B.; Møgelmose, A. (2013): From cad models to soft point cloud labels: An automatic annotation pipeline for cheaply supervised 3D semantic segmentation. In: Remote Sensing 15(2013)14. https://www.mdpi.com/2072-4292/15/14/3578.

Hutchison, D.; Kanade, T.; Kittler, J.; Kleinberg, J. M.; Mattern, F.; Mitchell, J. C.; Naor, M.; Nierstrasz, O.; Pandu Rangan, C.; Steffen, B., Sudan, M.; Terzopoulos, D.; Tygar, D.; Vardi, M. Y.; Weikum, G.; Tombari, F.; Salti, S.; Di Stefano, L. (2010): Unique Signatures of Histograms for Local Surface Description: In: Daniilidis, K.; Maragos, P.; Paragios, N. (Eds.): Computer Vision – ECCV 2010. Springer, Berlin/Heidelberg, 356–369.

Isfort, S.; Maiwald, F.; Mulsow, C.; Maas, H.-G. (2022): On the use of polarization cameras for the determination of concrete moisture. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLIII-B1(2022), 113–119. https://doi.org/10.5194/isprs-archives-xliii-b1-2022-113-2022.

ISO 10303-21:1994: Industrial Automation Systems and Integration Product Data Representation and Exchange Part 21: Implementation Methods: Clear Text Encoding of the Exchange Structure, 1994. International Organization for Standardization, Geneva, Switzerland.

ISO 16739-1:2018: Industry Foundation Classes (IFC) for Data Sharing in the Construction and Facility Management Industries – Part 1: Data Schema, Nov. 2018. International Organization for Standardization, Geneva, Switzerland.

ISO 16739-1:2024: Building Information Modeling (BIM) – Data Schema, Exchange Format Definitions, Std., 2024. International Organization for Standardization, Geneva, Switzerland.

Jiang, L.; Zhao, H.; Shi, S.; Liu, S.; Fu, C.-W.; Jia, J. (2020): Pointgroup: Dual-set point grouping for 3D instance segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

Johnson, A.; Hebert, M. (1999): Using spin images for efficient object recognition in cluttered 3D scenes. In: IEEE Transactions on Pattern Analysis and Machine Intelligence 21(1999)5, 433–449.

Justo, A.; Lamas, D.; Sánchez-Rodríguez, A.; Soilán, M.; Riveiro, B. (2023): Generating ifc-compliant models and structural graphs of truss bridges from dense point clouds. In: Automation in Construction 149(2023).

Karam, S.; Lehtola, V.; Vosselman, G. (2020): Strategies to integrate IMU and LiDAR SLAM for indoor mapping. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences V-1(2020), 223–230. https://doi.org/10.5194/isprs-annals-V-1-2020-223-2020.

Kawashima, K.; Kanai, S.; Date, H. (2012): Automatic recognition of piping system from large-scale terrestrial laser scanned point cloud. In: Seimitsu Kogaku Kaishi/Journal of the Japan Society for Precision Engineering 78(2012)8, 722–729.

Kellner, M.; Stahl, B.; Reiterer, A. (2022): Fused projection-based point cloud segmentation. In: Sensors 22(2022)3. https://www.mdpi.com/1424-8220/22/3/1139.

Kellner, M.; Stahl, B.; Reiterer, A. (2023): Reconstructing geometrical models of indoor environments based on point clouds. In: Remote Sensing 15(2023)18. https://www.mdpi.com/2072-4292/15/18/4421.

Kendall, A.; Gal, Y. (2017): What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? https://arxiv.org/abs/1703.04977.

Kermarrec, G.; Kargoll, B.; Alkhatib, H. (2020): Deformation Analysis Using B-Spline Surface with Correlated Terrestrial Laser Scanner Observations – A Bridge Under Load. In: Remote Sensing 12(2020)5, 829.

Kriegel, H.-P.; Kröger, P.; Schubert, E.; Zimek, A. (2009): Loop: Local outlier probabilities. In: Proceedings of the 18th ACM Conference on Information and Knowledge Management, Nov. 2009, New York, NY, 1649 – 1652. https://doi.org/10.1145/1645953.1646195.

Kwon, Y.; Won, J.-H.; Kim, B. J.; Paik, M. C. (n.d.): Uncertainty quantification using Bayesian neural networks in classification: Application to ischemic stroke lesion segmentation. https://openreview.net/pdf?id=Sk_P2Q9sG.

Lai, X.; Liu, J.; Jiang, L.; Wang, L.; Zhao, H.; Liu, S.; Qi, X.; Jia, J. (2022): Stratified Transformer for 3D Point Cloud Segmentation. https://arxiv.org/abs/2203.14508.

Landrieu, L.; Boussaha, M. (2019): Point cloud oversegmentation with graph-structured deep metric learning. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Los Alamitos, CA, 7432 – 7441. https://doi.ieeecomputersociety.org/10.1109/CVPR. 2019.00762.

Lapuschkin, S.; Binder, A; Montavon, G.; Muller, K.-R.; Samek, W. (2016): Analyzing classifiers: Fisher vectors and deep neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, 2912 – 2920.

Lee; J. H.; Park, J. J.; Yoon, H. (2020): Automatic Bridge Design Parameter Extraction for Scan-to-BIM. In: Applied Sciences 10(2020)20, 7346.

Li, L.; Sung, M.; Dubrovina, A.; Yi, L.; Guibas, L. (2019): Supervised fitting of geometric primitives to 3D point clouds. https://arxiv.org/abs/1811.08988.

Liang, X.; Kukko, A.; Kaartinen, H.; Hyyppä, J.; Yu, X.; Jaakkola, A.; Wang, Y. (2014): Possibilities of a personal laser scanning system for forest mapping and ecosystem services. In: Sensors 14(2014)1, 1228 – 1248. https://doi.org/10.3390/s140101228.

Lienhart, W.; Ehrhart, M.; Grick, M. (2017): High frequent total station measurements for the monitoring of bridge vibrations. In: Journal of Applied Geodesy 11(2017)1, 1 – 8.

Lu, R.; Brilakis, I. (2019a): Digital twinning of existing reinforced concrete bridges from labelled point clusters. In: Automation in Construction 105(2019), 102837.

Lu, R.; Brilakis, I.; Middleton, C. R. (2019b): Detection of structural components in point clouds of existing RC bridges: Detection of bridge components in point clouds. In: Computer-Aided Civil and Infrastructure Engineering 34(2019)3, 191 – 212.

Lueangvilai, E.; Chaisomphob, T. (2022): Application of mobile mapping system to a cable-stayed bridge in Thailand. In: Sensors 22(2022)24, 9625. https://doi.org/10.3390/s22249625.

Luhmann, T.; Fraser, C.; Maas, H.-G. (2016): Sensor modelling and camera calibration for close-range photogrammetry. In: ISPRS Journal of Photogrammetry and Remote Sensing 115(2016), 37 – 46. https://doi.org/10.1016/j.isprsjprs.2015.10.006.

Maalek, R.; Lichti, D. D.; Ruwanpura, J. Y. (2019): Automatic recognition of common structural elements from point clouds for automated progress monitoring and dimensional quality control in reinforced concrete construction. In: Remote Sensing 11(2019)9.

Maas, H.-G. (2008): Close range photogrammetry sensors. In: Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences, ISPRS Congress Book, 63 – 72. https://doi.org/10.1201/9780203888445.ch5.

Mader, D.; Blaskow, R.; Westfeld, P.; Maas, H.-G. (2015): UAV-based acquisition of 3D point cloud – A comparison of a low-cost laser scanner and SFM-tools. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-3/W3(2015), 335 – 341. https://doi.org/10.5194/isprsarchives-xl-3-w3-335-2015.

Mader, D.; Blaskow, R.; Westfeld, P.; Weller, C. (2016): Potential of UAV-based laser scanner and multispectral camera data in building inspection: In: ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLI-B1(2016), 1135 – 1142. https://doi.org/10.5194/isprsarchives-xli-b1-1135-2016.

Mafipour, M. S.; Vilgertshofer, S.; Borrmann, A. (2023a): Automated geometric digital twinning of bridges from segmented point clouds by parametric prototype models. In: Automation in Construction 156(2023).

Mafipour, M. S.; Vilgertshofer, S.; Borrmann, A. (2023b): Heuristic optimization for digital twin modeling of existing bridges from point cloud data by parametric prototype models. In: Proceedings of the ASCE International Conference on Computing in Civil Engineering, June 2023.

Mallet, C.; Bretar, F.; Roux, M.; Soergel, U.; Heipke, C. (2011): Relevance assessment of full-waveform LiDAR data for urban area classification: In: ISPRS Journal of Photogrammetry and Remote Sensing 66(2011)6, S71 – S84.

Martens, J.; Blankenbach, J. (2020): An evaluation of pose-normalization algorithms for point clouds introducing a novel histogram-based approach, Advanced Engineering Informatics 46 (2020), 101132, ISSN 1474-0346. https://doi.org/10.1016/j.aei.2020.101132

Martens, J.; Blut, T.; Blankenbach, J. (2023): Cross domain matching for semantic point cloud segmentation based on image segmentation and geometric reasoning. In: Advanced Engineering Informatics 57(2023), 102076.

Milioto, A.; Vizzo, I.; Behley, J.; Stachniss, C. (2019): Rangenet ++: Fast and accurate LiDAR semantic segmentation: In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 4213 – 4220.

Mirkin, B. (2005): Clustering for data mining: A data recovery approach. Chapman and Hall/CRC, New York. https://doi.org/10.1201/9781420034912.

Mohammadi, M.; Rashidi, M.; Mousavi, V.; Yu, Y.; Samali, B. (2022): Application of TLS method in digitization of bridge infrastructures: A path to BIM development. In: Remote Sensing 14(2022)5, 1148. https://doi.org/10.3390/rs14051148.

Monnier, F.; Vallet, B.; Soheilian, B. (2012): Trees detection from laser point clouds acquired in dense urban areas by a mobile mapping system. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences I-3(2012), 245 – 250.

Moore, M.; Phares, B. M.; Graybeal, B. A.; Rolander, D. D.; Washer, G. A. (2001): Reliability of visual inspection for highway bridges. FHWA Report Nos. FHWA-RD-01-020 and FHWA-RD-01-21, Federal Highway Administration, U.S. Department of Transportation, June 2001.

Moufti, S. A.; Zayed, T. M.; Dabous, S. A. (2014): Defect-based condition assessment of concrete bridges. In: Transportation Research Record 2431(2014), 88 – 96.

Muja, M.; Lowe, D. G. (2009): Fast approximate nearest neighbors with automatic algorithm configuration. In: International Conference on Computer Vision Theory and Applications, 2009. https://api.semanticscholar.org/CorpusID:7317448.

Nan, L.; Wonka, P. (2017): Polyfit: Polygonal surface reconstruction from point clouds. In: Proceedings of the IEEE International Conference on Computer Vision, 2017, 2353 – 2361.

NIBS – National Institute of Building Sciences (2023): What is a BIM? https://www.nationalbimstandard.org/faqs#faq1 (14.09.2023).

Niemeyer, J.; Rottensteiner, F.; Soergel, U. (2014): Contextual classification of LiDAR data and building object detection in urban areas. In: ISPRS Journal of Photogrammetry and Remote Sensing 87(2014), 152 – 165.

Ochmann, S.; Vock, R.; Wessel, R.; Klein, R. (2016): Automatic reconstruction of parametric building models from indoor point clouds. In: Computers and

Graphics 54(2016), 94 – 103. https://www.sciencedirect.com/science/article/pii/S0097849315001119.

Office of Transportation Maintenance, New York State Department of Transportation (2008): Fundamentals of Bridge Maintenance and Inspection. Albany, New York.

Oggier, T.; Lehmann, M.; Kaufmann, R.; Schweizer, M.; Richter, M.; Metzler, P.; Lang, G.; Lustenberger, F.; Blanc, N. (2004): An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger). In: SPIE Proceedings 5249(2004). https://doi.org/10.1117/12.513307.

Okorn, B.; Xiong, X.; Akinci, B. (2010): Toward automated modeling of floor plans. https://api.semanticscholar.org/CorpusID:3334144.

Otepka, J.; Ghuffar, S.; Waldhauser, C.; Hochreiter, R.; Pfeifer, N. (2013): Georeferenced point clouds: A survey of features and point cloud management. In: ISPRS International Journal of Geo-Information 2(2013)4, 1038 – 1065.

Park, C.; Jeong, Y.; Cho, M.; Park, J. (2022): Fast Point Transformer. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, 16928 – 16937.

Park, H.; Lee, J.-H. (2007): B-spline curve fitting based on adaptive curve refinement using dominant points. In: Computer-Aided Design 39(2007)6, 439 – 451. https://www.sciencedirect.com/science/article/pii/S0010448507000024.

Pearson, K. (1901): Liii. On lines and planes of closest fit to systems of points in space. In: The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2(1901)11, 559 – 572. https://doi.org/10.1080/14786440109462720.

Pepe, M.; Costantino, M. (2021): UAV photogrammetry, 3D modelling of complex architecture for maintenance purposes: the case study of the masonry bridge on the Sele River, Italy. In: Periodica Polytechnica Civil Engineering 65(2021)1, 191 – 203. https://pp.bme.hu/ci/article/view/16398.

Piegl, L.; Tiller, W. (1996): The NURBS Book. 2nd Ed. Springer, New York, NY.

Poku-Agyemang, K. N.; Kellner, M.; Schmitt, A.; Reiterer, A. (2023a): Creating 3D models of bridges using different data sources and machine learning methods. In: Proceedings of the 33rd European Safety and Reliability Conference, Sept. 3 – 8, 2023, Southampton, UK. http://dx.doi.org/10.3850/978-981-18-8071-1_P289-cd.

Poku-Agyemang, K. N.; Reiterer, A. (2023b): 3D reconstruction from 2d plans exemplified by bridge structures. In: Remote Sensing 15(2023)3, 677. https://doi.org/10.3390/rs15030677.

Ponciano, J.-J.; Roetner, M.; Reiterer, A.; Boochs, F. (2021): Object semantic segmentation in point clouds – Comparison of a deep learning and a knowledge-based method. In: ISPRS International Journal of Geo-Information 10(2021)4. https://www.mdpi.com/2220-9964/10/4/256.

Poux, F.; Billen, R. (2019): Voxel-based 3D point cloud semantic segmentation: Unsupervised geometric and relationship featuring vs deep learning methods. In: ISPRS International Journal of Geo-Information 8(2019)5. https://www.mdpi.com/2220-9964/8/5/213.

Previtali, M.; Brumana, R.; Banfi, F. (2020): Existing infrastructure cost effective informative modelling with multisource sensed data: TLS, MMS and photogrammetry. In: Applied Geomatics 14(2020)S1, 21 – 40. http://dx.doi.org/10.1007/s12518-020-00326-3.

Qi, C. R.; Su, H.; Mo, K.; Guibas, L. J. (2016): PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. https://arxiv.org/abs/1612.00593.

Qi, C. R.; Yi, L.; Su, H. Guibas, L. J. (2017): PointNet++: Deep hierarchical feature learning on point sets in a metric space. https://arxiv.org/abs/1706.02413.

Qin, G.; Zhou, Y.; Hu, K.; Han, D.; Ying, C. (2021): Automated Reconstruction of Parametric BIM for Bridge Based on Terrestrial Laser Scanning Data. In: Advances in Civil Engineering (2021), 1 – 17. https://doi.org/10.1155/2021/8899323.

Rebolj, D.; Pučko, Z.; Babič, N. Č.; Bizjak, M.; Mongus, D. (2017): Point cloud quality requirements for Scan-vs-BIM based automated construction progress monitoring. In: Automation in Construction 84(2017), 323 – 334.

Ribeiro, M. T.; Singh, S.; Guestrin, C. (2016): “Why should i trust you?” Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135 – 1144.

Riegler, G.; Ulusoy, A. O.; Geiger, A. (2017): OctNet: Learning deep 3D representations at high resolutions. In: Proceedings 30th IEEE Conference on Computer Vision and Pattern Recognition, 6620 – 6629.

Rocha, G.; Mateus, L.; Fernández, J.; Ferreira, V. (2020): A Scan-to-BIM Methodology Applied to Heritage Buildings. In: Heritage 3(2020)1, 47 – 67.

Sacks, R.; Eastman, C. M.; Lee, G.; Teicholz, P. M. (2018): BIM handbook: A guide to building information modeling for owners, managers, designers, engineers and contractors, and Facility Managers. John Wiley & Sons, Hoboken, NJ.

Salti, S.; Tombari, F.; Di Stefano, L. (2014): SHOT: Unique signatures of histograms for surface and texture description. In: Computer Vision and Image Understanding 125(2014), 251 – 264.

Schnabel, R.; Wahl, R.; Klein, R. (2007): Efficient ransac for point-cloud shape detection. In: Computer Graphics Forum 26(2007)2, 214 – 226. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2007.01016.x.

Schult, J.; Engelmann, F.; Hermans, A.; Litany, O.; Tang, S.; Leibe, B. (2023): Mask3D: Mask Transformer for 3D Semantic Instance Segmentation. https://arxiv.org/abs/2210.03105.

Sun, J.; Qing, C.; Tan, J.; Xu, X. (2022): Superpoint transformer for 3D scene instance segmentation. https://arxiv.org/abs/2211.15766.

Tan, H.; Kotthaus, H. (2022): Surrogate model-based explainability methods for point cloud NNs. In: 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 2927 – 2936.

Tatsuro Yamane, P.-J. C.; Honda, R. (2022): Detecting and localising damage based on image recognition and structure from motion, and reflecting it in a 3D bridge model. In: Structure and Infrastructure Engineering 20(2022)4, 594 – 606. https://doi.org/10.1080/15732479.2022.2131845.

Thomas, H.; Qi, C. R.; Deschaud, J. E.; Marcotegui, B.; Goulette, F.; Guibas, L. (2019): KPConv: Flexible and deformable convolution for point clouds. In: Proceedings of the IEEE International Conference on Computer Vision, 6410 – 6419.

Truong-Hong, L.; Lindenbergh, R. (2022): Automatically extracting surfaces of reinforced concrete bridges from terrestrial laser scanning point clouds. In: Automation in Construction 135(2022), 104127.

Tulsiani, S.; Su, H.; Guibas, L. J.; Efros, A. A.; Malik, J. (2017): Learning shape abstractions by assembling volumetric primitives. In: Computer Vision and Pattern Regognition (CVPR), 2017.

Vassilev, H.; Laska, M.; Blankenbach, J. (2024): Uncertainty-aware point cloud segmentation for infrastructure projects using Bayesian deep learning. Automation in Construction 164 (2024), 105419, ISSN 0926-5805. https://doi.org/10.1016/j.autcon.2024.105419

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; Polosukhin, I. (2017): Attention is all you need. https://arxiv.org/abs/1706.03762.

Vu, T.; Kim, K.; Luu, T. M.; Nguyen, X. T.; Yoo, C. D. (2022): SoftGroup for 3D instance segmentation on 3D point clouds. https://arxiv.org/abs/2203.01509.

Wang, T.-R.; Liu, N.; Yuan, L.; Wang, K.-X.; Sheng, X.-J. (2022): Iterative least square optimization for the weights of nurbs curve. In: Mathematical Problems in Engineering (2022). https://onlinelibrary.wiley.com/doi/10.1155/2022/5690564.

Wang, W.; Yu, R.; Huang, Q.; Neumann, U. (2018): Sgpn: Similarity group proposal network for 3D point cloud instance segmentation. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2569 – 2578.

Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S. E.; Bronstein, M. M.; Solomon, J. M. (2019): Dynamic Graph CNN for Learning on Point Clouds. In: ACM Transactions on Graphics 38(2019)5, 146, 1–12. https://doi.org/10.1145/3326362.

Weinmann, M.; Jutzi, B.; Hinz, S.; Mallet, C. (2015): Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. In: ISPRS Journal of Photogrammetry and Remote Sensing 105(2015), 286–304.

Weinmann, M.; Jutzi, B.; Mallet, C. (2013): Feature relevance assessment for the semantic interpretation of 3D point cloud data. In: ISPRS Workshop Laser Scanning, 2013, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences II-5/W2(2013), 313–318.

Weld, D. S.; Bansal, G. (2018): The challenge of crafting intelligible intelligence. In: Communications of the ACM 62(2018), 70–79.

West, K.; Webb, B.; Lersch, J.; Pothier, S.; Triscari, J.; Iverson, E. (2004): Context-driven automated target detection in 3D data. In: Proceedings of SPIE – The International Society for Optical Engineering 5426(2004).

Westfeld P.; Maas, H.-G. (2013): Integrierte Selbstkalibrierung distanzmessender 3D-Kameras. In: Photogrammetrie – Fernerkundung – Geoinformation (2013)6, 589–602. https://doi.org/10.1127/1432-8364/2013/0202.

Weygant, R. S. (2011): BIM content development: standards, strategies, and best practices. John Wiley & Sons, Hoboken, NJ.

Wong, K.; Wang, S.; Ren, M.; Liang, M.; Urtasun, R. (2020): Identifying unknown instances for autonomous driving. In: Proceedings of the Conference on Robot Learning, ser. Proceedings of Machine Learning Research, Oct. 30 – Nov 01, 2020, 384–393. https://proceedings.mlr.press/v100/wong20a.html.

Wu, W.; Qi, Z.; Li, F. (2018): Pointconv: Deep convolutional networks on 3D point clouds. http://arxiv.org/abs/1811.07246.

Xia, T.; Yang, J.; Chen, L. (2022): Automated semantic segmentation of bridge point cloud based on local descriptor and machine learning. In: Automation in Construction 133(2022), 103992.

Xu, J.; Zhang, R. Dou, J.; Zhu, Y.; Sun, J.; Pu, S. (2021): Rpvnet: A deep and efficient range-point-voxel fusion network for LiDAR point cloud segmentation. https://arxiv.org/abs/2103.12978.

Xu, Y.; Tuttas, S.; Hoegner, L.; Stilla, U. (2017): Reconstruction of scaffolds from a photogrammetric point cloud of construction sites using a novel 3D local feature descriptor. In: Automation in Construction 85(2017), 76–95.

Yan, Y.; Hajjar, J. F. (2021): Automated extraction of structural elements in steel girder bridges from laser point clouds. In: Automation in Construction 125(2021), 103582.

Yuan, L.; Chen, Y.; Wang, T.; Yu, W.; Shi, Y.; Jiang, Z.; Tay, F. E.; Feng, J.; Yan, S. (2021): Tokens-to-Token ViT: Training Vision Transformers from Scratch on ImageNet. https://arxiv.org/abs/2101.11986.

Zaganidis, A.; Magnusson, M.; Duckett, T; Cielniak, G. (2017): Semantic-assisted 3D normal distributions transform for scan registration in environments with limited structure. In: Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS), Sep. 2017, 4064–4069.

Zhang, C.; Luo, W.; Urtasun, R. (2018): Efficient convolutions for real-time semantic segmentation of 3D point clouds. In: 2018 International Conference on 3D Vision, 399–408.

Zhang, R.; Wang, W. (2018): Second and high-order graph matching for correspondence problems. In: IEEE Transactions on Circuits and Systems for Video Technology 28(2018)10, 2978–2992. https://doi.org/10.1109/tcsvt.2017.2718225.

Zhang, Y.; Zhou, Z.; David, P.; Yue, X.; Xi, Z.; Foroosh, H. (2020): Polarnet: An improved grid representation for online LiDAR point clouds semantic segmentation. https://arxiv.org/abs/2003.14032.

Zhu, X.; Zhou, H.; Wang, T.; Hong, F.; Ma, Y.; Li, W.; Li, H.; Lin, D. (2020): Cylindrical and asymmetrical 3D convolution networks for LiDAR segmentation. https://arxiv.org/abs/2011.10033.

AUTHORS

Maximilian Kellner

UNIVERSITY OF FREIBURG
DEPARTMENT OF SUSTAINABLE SYSTEMS ENGINEERING (INATECH) & FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM

Georges-Koehler-Allee 301 | 79110 Freiburg | Germany
maximilian.kellner@ipm.fraunhofer.de | ORCID: 0000-0002-1184-5346

M. Sc. Hristo Vassilev

RWTH AACHEN UNIVERSITY
CHAIR OF COMPUTING IN CIVIL ENGINEERING AND GIS AND GEODETIC INSTITUTE

Mies-van-der-Rohe-Str. 1 | 52074 Aachen | Germany
hristo.vassilev@gia.rwth-aachen.de | ORCID: 0009-0001-5809-7197

M. Sc. Antonia Busch

TU BERLIN
INSTITUTE OF GEODESY

Kaiserin-Augusta-Allee 104-106 | 10553 Berlin | Germany
antonia.busch@tu-berlin.de | ORCID: 0009-0007-9074-3017

Dipl.-Ing. Robert Blaskow

TUD DRESDEN
UNIVERSITY OF TECHNOLOGY
INSTITUTE OF PHOTOGRAMMETRY AND REMOTE SENSING

Helmholtzstr. 10 | 01069 Dresden | Germany
robert.blaskow@tu-dresden.de | ORCID: 0009-0004-7248-9631

M. Sc. Mariana Ferrandon Cervantes

UNIVERSITY OF FREIBURG
DEPARTMENT OF SUSTAINABLE SYSTEMS ENGINEERING (INATECH) & FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM

Georges-Koehler-Allee 301 | 79110 Freiburg | Germany
mariana.ferrandon@inatech.uni-freiburg.de | ORCID: 00090000-0003-4875-1498

Kwasi Nyarko Poku-Agyemang
UNIVERSITY OF FREIBURG
DEPARTMENT OF SUSTAINABLE SYSTEMS ENGINEERING (INATECH) & FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM
Georges-Koehler-Allee 301 | 79110 Freiburg | Germany
kwasi.poku-agyemang@inatech.uni-freiburg.de | ORCID: 0000-0002-9492-6589

Annette Schmitt
UNIVERSITY OF FREIBURG
DEPARTMENT OF SUSTAINABLE SYSTEMS ENGINEERING (INATECH) & FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM
Georges-Koehler-Allee 301 | 79110 Freiburg | Germany
annette.schmitt@inatech.uni-freiburg.de | ORCID: 0000-0002-0788-1387

Sven Weisbrich
TU BERLIN
INSTITUTE OF GEODESY
Kaiserin-Augusta-Allee 104-106 | 10553 Berlin | Germany
s.weisbrich@tu-berlin.de | ORCID: 0000-0001-6786-9008

Prof. Dr. habil. Hans-Gerd Maas
TUD DRESDEN
UNIVERSITY OF TECHNOLOGY
INSTITUTE OF PHOTOGRAMMETRY AND REMOTE SENSING
Helmholtzstr. 10 | 01069 Dresden | Germany
hans-gerd.maas@tu-dresden.de | ORCID: 0000-0001-9034-3469

Prof. Dr. Frank Neitzel
TU BERLIN
INSTITUTE OF GEODESY
Kaiserin-Augusta-Allee 104-106 | 10553 Berlin | Germany
frank.neitzel@tu-berlin.de | ORCID: 0000-0001-7241-0656

Prof. Dr. Alexander Reiterer
UNIVERSITY OF FREIBURG
DEPARTMENT OF SUSTAINABLE SYSTEMS ENGINEERING (INATECH) & FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM
Georges-Koehler-Allee 301 | 79110 Freiburg | Germany
alexander.reiterer@ipm.fraunhofer.de | ORCID: 0000-0002-3196-3876

Prof. Dr. Jörg Blankenbach
RWTH AACHEN UNIVERSITY
CHAIR OF COMPUTING IN CIVIL ENGINEERING AND GIS AND GEODETIC INSTITUTE
Mies-van-der-Rohe-Str. 1 | 52074 Aachen | Germany
blankenbach@gia.rwth-aachen.de | ORCID: 0000-0002-5700-8818

Manuskript eingereicht: 21.2.2024 | Im Peer-Review-Verfahren begutachtet