



R&D Project

Registering and visualizing point cloud data with existing 3D CityGML Models

Alan Orlando Gomez Torres

Submitted to Hochschule Bonn-Rhein-Sieg,
Department of Computer Science
in partial fulfilment of the requirements for the degree
of Master of Science in Autonomous Systems

Supervised by

Prof. Dr Paul G. Plöger
Dr. Stefan Rilling
Alex Mitrevski

September 2020

I, the undersigned below, declare that this work has not previously been submitted to this or any other university and that it is, unless otherwise stated, entirely my own work.

Date

Alan Orlando Gomez Torres

Abstract

- INSERT ABSTRACT

Acknowledgements

Thanks to

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Challenges and Difficulties	2
1.3	Problem Statement	2
2	State of the Art	5
2.1	Point-to-point registration	6
2.1.1	Global methods	6
2.1.2	Local methods	7
2.1.3	Hybrid methods	8
2.2	Point-to-model registration	9
2.2.1	Sampling methods	9
2.2.2	Local methods	9
2.3	Others	10
2.4	Limitations of previous work	11
3	Methodology	13
3.1	Data	13
3.2	Setup	14
3.3	Registration Problem in 3D	15
3.4	Registration Problem in 2D	15
3.5	Registration Problem as Mixed Integer Linear Program	15
4	Solution	17
4.1	Proposed algorithm	17
4.1.1	Projection of the point cloud onto the xy-plane	18
4.1.2	Projection the 3D model onto the xy-plane	18
4.1.3	Detection of line segments	19
4.1.4	Detection of line intersections and their angles	19
4.1.5	Identification of possible correspondences between angles detected	21
4.1.6	Alignment of both projections onto the xy-plane	21
4.1.7	Alignment on the z-axis	23
4.2	Implementation details	23
5	Evaluation and Results	25

6 Conclusions	27
6.1 Contributions	27
6.2 Lessons learned	27
6.3 Future work	27
Appendix A Design Details	29
Appendix B Parameters	31
References	33

List of Figures

1.1	A-DRZ architecture system (Image provided by the Fraunhofer IAIS).	3
1.2	System diagram.	4
2.1	Classification of the State of the Art.	6
3.1	Front view of a point cloud and a model before registration.	14
3.2	Back view of a point cloud and a model before registration.	14
4.1	Projection of the source onto the xy-plane.	18
4.2	Projection of the target onto the xy-plane.	19
4.3	Source image after filter.	20
4.4	Source lines and intersections.	20
4.5	Target lines and intersections.	21
4.6	Initial position of the source image and the target image.	22
4.7	Source image and target image aligned.	22
5.1	Front view of a point cloud and a model after registration.	25
5.2	Back view of a point cloud and a model after registration.	26

List of Tables

1

Introduction

1.1 Motivation

Every day people are exposed to have some kind of accident and the worst kind of accidents are those that put their lives at risk. Fortunately, for those kinds of accidents, we have rescue services like firefighters. But firefighters are humans and put their lives at risk trying to save other lives. According to the Deutscher Feuerwehrverband (German Firefighters Association), the number of firefighters missions in Germany in 2017 ascends to more than 3 million, and the number of deaths in fires in 2016 was approximately 350 [3].

What if the firefighters are supported by rescue robots? With the use of rescue robots that explore the accident areas, we can help firefighters do their job more safely. Rescue robots can explore buildings to find entrances, exits, and other key areas that help the firefighters to rescue people in a faster and more efficient way. But robots cannot find their exact position inside a building with a single GPS, because this kind of technology is unreliable under certain conditions (e.g. inside buildings).

In this work, we try to enable rescue robots to explore buildings with the use of LiDAR systems and 3D models of the buildings. With the use of LiDAR systems, we can obtain representations of the actual status of buildings at risk without endangering any lives. These kinds of representations are known as point clouds.

With the use of CityGML models [22], we have access to the 3D representation of many buildings in Germany. CityGML models can be presented in different Levels of Detail (LOD), from the coarsest level LOD0 to the most refined level LOD4. In LOD0, buildings can be represented by a simple footprint. While in LOD4, buildings contain information about windows, doors, stairs, and other internal structures. Unfortunately, not all models are LOD4. Even if they were, the information might be outdated. For example, it may be that due to some collapse, some door or window is blocked.

Using modern registration techniques, one can find a match between the point cloud given by the LiDAR system and the 3D model of the building. The registration result can be then used to find the exact position of the robot with respect to the building. The registration result can also help to locate relevant parts of the building like windows, doors, stairs, or corridors. This information can be presented to the firefighters in a way that enables them to elaborate on a rescue plan, with the only objective of supporting firefighters in their mission to rescue lives and safeguard theirs.

1.2 Challenges and Difficulties

Most of the registration algorithms are focused on working for two point clouds. Just a few registration algorithms address the registration of 3D models with point clouds. And between those registration algorithms, only a couple are created to work with CityGML models, as described in chapter 2.

The registration process is normally done manually, due to the fact that it is a simple task for a human since we all learn to match patterns when we are just kids. The automation of registration is complicated because for a computer it is difficult to find features and match them between two different types of 3D data.

An additional challenge is the data itself, the point clouds available are partial and they do not represent the whole 3D model, normally they just capture part of the walls of the buildings and the surface outside the buildings. And vice versa, the 3D models do not have information of the surface outside the buildings. The only information available in both the 3D models and the point cloud is some of the walls of the buildings.

1.3 Problem Statement

This work intends to implement an automated registration method for a point cloud with a CityGML model. This automated registration method will then be used as part of the A-DRZ (Aufbau des Deutschen Rettungsrobotik-Zentrums) [1] project of the Fraunhofer IAIS. Figure 1.1 shows the system architecture of the A-DRZ project. The final objective of the automated registration in the A-DRZ project is to be able to look inside the CityGML models, in order to find the location and status of relevant parts of the buildings (e.g. doors, windows, stairs, and corridors) that can be used for rescue task.

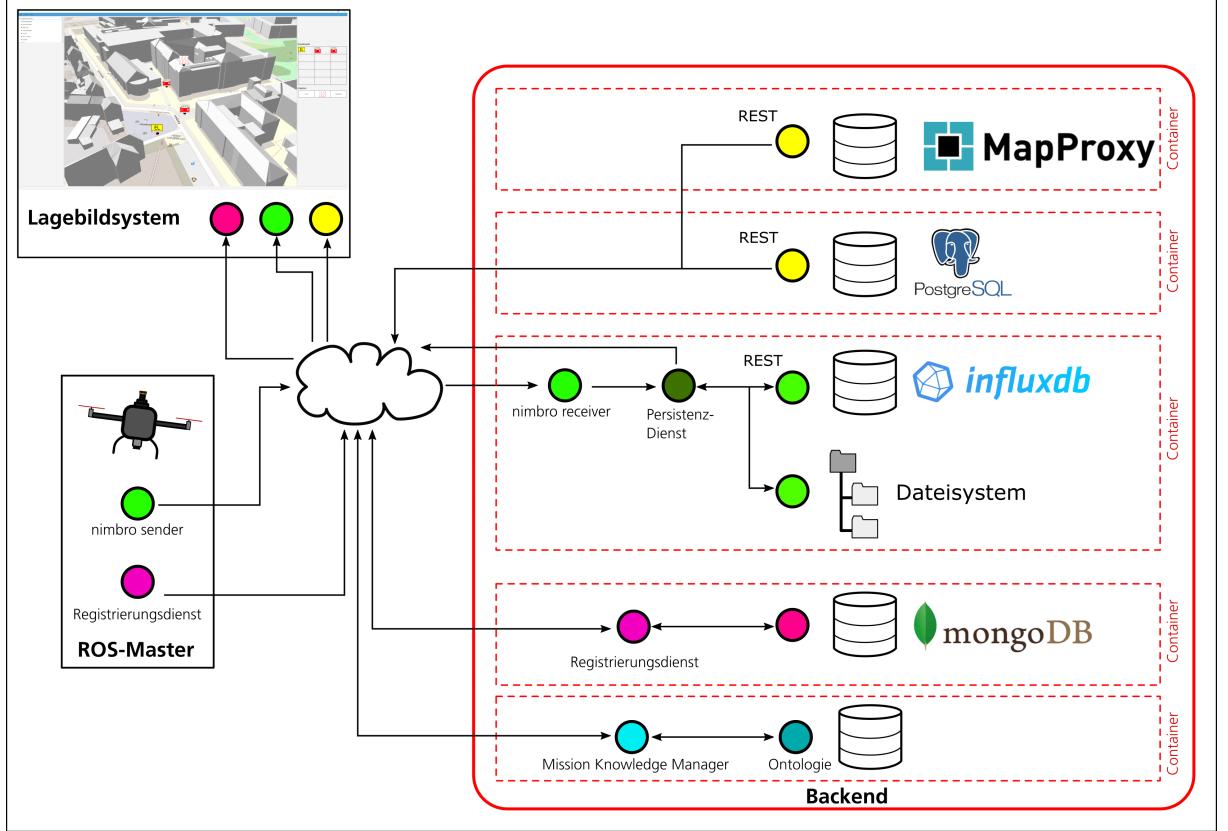


Figure 1.1: A-DRZ architecture system (Image provided by the Fraunhofer IAIS).

The current automated approaches are mainly focused on point-to-point registration. Therefore, the exploration of point-to-model automated registration methods is necessary to implement a solution that aims to work in real-time. Moreover, there is no automated process implemented in A-DRZ project yet. However, as an assistance system, the A-DRZ project should avoid human interaction for performing the registration.

Figure 1.2 shows the general workflow of the needed method. The input of the method is a point cloud together with the corresponding CityGML model of a building (the process to obtain this is outside of the scope of this work). The provided input is then preprocessed and used for the registration algorithm. The output of the registration algorithm is a transformation that maps the point cloud to the CityGML model. Finally, the output transformation and the input data are used for visualization of the point cloud on the CityGML model.

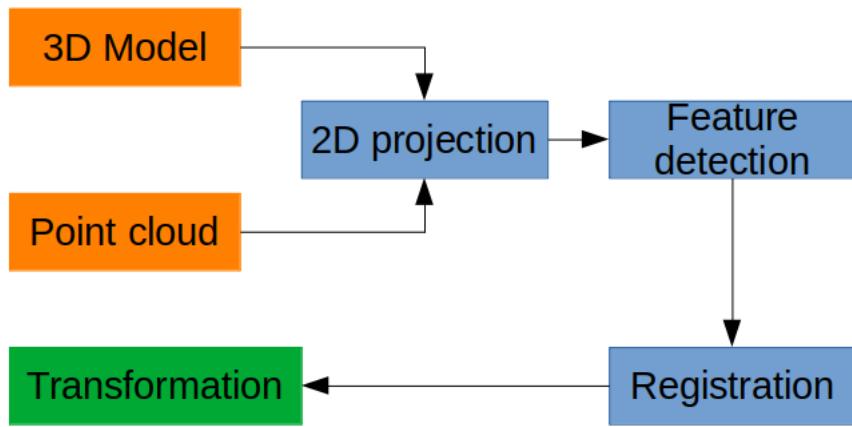


Figure 1.2: System diagram.

Firstly, the implementation will be tested individually. Secondly, the implementation will be tested with its integration on a visualization software of the Fraunhofer IAIS. This visualization software is a simple 3D rendering of CityGML models limited to a certain geographical area. Thirdly, the implementation will be tested with its integration on the A-DRZ project.

2

State of the Art

Point cloud registration algorithms are normally classified into two major categories: global (coarse) registration and local (fine) registration [36, 26]. Coarse registration algorithms attempt to find a transformation that approximately maps the source point cloud to the target point cloud. Fine registration algorithms attempt to refine an initial transformation to map the source point cloud to the target point cloud. There are some approaches that combine both of these methods, they first apply a global method and then refine the results with the help of a local method. We can classify these approaches into a 3rd category of hybrid methods.

For the purpose of this work we can classify point cloud registration algorithms into two major categories: point-to-point and point-to-model. Point-to-point registration algorithms attempt to align two point clouds, while point-to-model registration algorithms attempt to align a point cloud with a 3D model.

The approaches that do not work with point cloud data or 3D models, but are still related to the registration problem, are classified into the category of others.

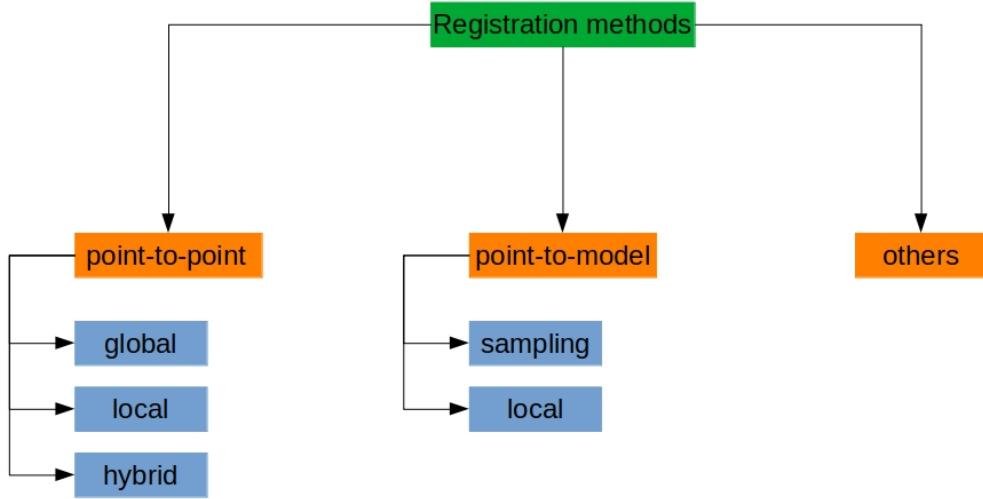


Figure 2.1: Classification of the State of the Art.

Figure 2.1 shows an overview of the papers reviewed.

2.1 Point-to-point registration

2.1.1 Global methods

RANSAC [17] can be used to perform the registration task but is time-consuming due to the randomness of the RANSAC algorithm. Therefore, Pankaj et al. [33] present a guided sampling improvement to RANSAC, to reduce the registration time, sorting the correspondences obtained from a previous 3D feature matching according to their quality.

Following the same line, S. Quan and J. Yang [36] propose a method that helps the RANSAC algorithm sample the best point correspondences. Their approach first detects keypoints using Harris3D, then it matches the keypoints by using the Local Voxelized Structure (LoVS) descriptor. After that, the set of correspondences is reduced using the Nearest Neighbor Similarity Ration (NNSR) and the compatibility score of the remaining correspondences with the rigidity and DSP constraint. Finally, the correspondences with the best compatibility scores are used to generate transformations and the best one (the one with the maximum number of inliers) is the final result. Unfortunately, the RANSAC-based approaches are sensitive to outliers.

Another less common approach for coarse registration is proposed by Sakakubara et al. [37]. This coarse registration approach uses Mixed Integer Linear Programming to find pairs of corresponding points

that are then used to find a coarse transformation that aligns two different point clouds. This approach finds a global optimal result without using the values of invariant features, it adjusts the error tolerance depending on the accuracy of the given data and it finds the best correspondences between the points. Nevertheless, Mixed Integer Linear Programming problems are NP-hard. Therefore, this approach is excessively time-consuming.

Others have proposed the use of neural networks to attempt to solve the global registration problem. This is the case of Wang and Solomon [41]. They propose a learning-based method called Deep Closest Point (DCP). The approach proposed first uses PointNet [34] or DGCNN [42] to embed the point clouds. Then, it uses an attention-based module to encode the output of the first part. Finally, it estimate the transformation using a differentiable SVD layer.

Sarode et. al [39] present a registration framework that also uses PointNet to extract features from the point clouds. The PointNet features are then aligned using another network. The process is then repeated a given number of times to improve the registration result.

Ding and Feng [15] present another deep learning method called DeepMapping, whose main structure is constituted of two deep neural networks. The first network is a localization network, whose output is the pose estimation for the input point clouds, and the second network is a map network that estimates the occupancy status of the input locations to compute a loss function that reveals the registration quality.

More recently, Huang et al. [25] propose a learning method that can be trained in a semi-supervised or unsupervised way. The proposed framework is composed of an encoder that extracts features from the input point clouds and a decoder that decodes the extracted features to compute a projection error that serves as input for an algorithm that estimates a transformation increment. This increment is then employed to update the parameters of the transformation and transform one of the point clouds. The whole process is repeated until the loss computed is less than some desired value.

2.1.2 Local methods

Besl and N. D. McKay [9] propose the Iterative Closest Point (ICP) algorithm to perform registration between point sets, line segments, implicit curves, parametric curves, triangles, implicit surfaces, and parametric surfaces. ICP iteratively finds and updates correspondences between the desired sets based on an initial approximation to the solution and the spatial distance.

According to Segal et al. [40], the ICP algorithm can be summarized in two main steps: the computation of correspondences between the two scans and the computation of a transformation that minimizes the distance between the corresponding points. They present a generalized version of the ICP algorithm by attaching a probabilistic model to the second step. This generalization makes the algorithm more robust to outliers without modifying the simplicity. However, it does not improve the speed of the original ICP.

ICP and its variants are still the standard algorithms for point cloud registration because of their simplicity.

2.1.3 Hybrid methods

The results of the global methods can be refined with the use of a local method, most of the time a variation of the popular ICP. In this category, some of these methods can be found.

Go-ICP is a branch-and-bound (BnB) based approach proposed by Yang et al. [44] to address the susceptibility of ICP to local minima. The main idea of Go-ICP is that BnB and ICP work iteratively until a global minimum is reached: BnB finds a solution and this is refined by ICP. Go-ICP guarantees global optimality regardless of the initialization and it is recommended for scenarios where real-time performance is not critical.

The approach proposed by Jun Lu et al. [29] extracts keypoints from the point clouds, then it obtains four optimal consistent points from the key points by using Super4PCS [32], afterwards it creates a neighborhood ball with each point as its center and obtains the overlapping regions as the intersections of these neighborhood balls. Finally, ICP is applied to the overlapping regions. The main drawback of this method is that there overlapping regions should exist to perform the registration task.

The approach proposed by Jun Lu et al. [29] extracts keypoints from the point clouds, then it obtains four optimal point correspondences from the key points by using Super4PCS [32], i.e., the four-point correspondences that minimize the number of outliers. Afterward, it creates a neighborhood ball with each point as its center and obtains these neighborhood balls' overlapping regions. Finally, it applies ICP to the overlapping regions. The main drawback of this method is that overlapping regions should exist to perform the registration task.

The approach proposed by Jun Lu et al. [2] extracts keypoints from the point clouds, then it obtains four optimal point correspondences from the key points by using Super4PCS [5], i.e., the four-point correspondences that minimize the number of outliers. Afterward, it creates a neighborhood ball with each point as its center and obtains the overlapping regions as these neighborhood balls' intersections. Finally, it applies ICP is to the overlapping regions. The main drawback of this method is that there overlapping regions should exist to perform the registration task.

Because of the increasing success of Deep Learning in 2D applications, there have been attempts to use deep neural networks to solve the complete point cloud registration problem (global and local registration). DeepICP [30] is the first end-to-end learning-based point cloud registration framework. DeepICP makes use of other networks to build a complete framework. The first step of DeepICP is the Deep Feature Extraction (PointNet++ [35]), which extracts features from the point clouds. In the second step, the Pont Weighting (3DFeatNet [45]) selects the relevant points. In the third step, the selected points of the target point cloud are sampled using the initial transformations. In the fourth step, Deep Feature Embedding is used to obtain more specific descriptors of the points. In the final step, the corresponding pairs of points are generated with a 3D Convolutional Neural Network to compute an improvement of the initial transformation.

Other approaches does not use a neural network to solve the complete point cloud registration problem but a part of it. LORAX is an algorithm proposed in [16] that selects some sets of points called super-points and uses a low-dimensional descriptor generated by an autoencoder to describe the structure of each

super-point. The sets of superpoints of both pointclouds are then matched using the euclidean distance between the descriptors. Finally, a coarse registration is computed using a RANSAC procedure, followed by a fine registration using ICP. The main contribution of LORAX is the use of an autoencoder to extract features of 3D point clouds. Nevertheless, the implementation is not optimized for real-time performance and it is not proved to be fast for big point clouds.

2.2 Point-to-model registration

2.2.1 Sampling methods

In order to register a 3D model with a point cloud, the simplest idea is to sample points from the 3D model to generate a second point cloud and then perform one of the point-to-point registration methods.

Kim et al. [26] register a 3D CAD model of a construction site with a point cloud by sampling points from the 3D model and then applying a point-to-point registration algorithm. The registration consists of three main steps. First, the point clouds are re-sampled to the same resolution. Second, PCA is used to identify the principal axes of the data to obtain a coarse registration based on these axes. Third, fine registration is performed with the Levenberg-Marquardt iterative closest point (LM-ICP) algorithm [18].

Kim et al. [27] is an extension of their previous work [26]. In this approach, they just add a noise filter step to the pre-processing of the point cloud in order to improve the registration results. The rest of the steps remain the same. However, using PCA for coarse registration will not assure a good performance when the form of the building is close to a cube.

2.2.2 Local methods

Other approaches work directly with the information provided in the 3D models to perform the registration task.

Li and Song [28] propose two extensions to the most popular local registration method, the ICP algorithm, to be able to use it with an STL-file (3D model) and a point cloud. The first of these extensions is called ICP-STL, and its main modification consists of finding the corresponding triangle mesh of the STL file to the points in the point cloud, and then the closest point in the triangle mesh to each point of the point cloud. The rest of the ICP algorithm remains the same. The second modification is called ICP-DAF (ICP dynamic adjustment factor), and it aims to reduce the time of the ICP-STL method. ICP-DAF adds a dynamic adjustment factor to adjust the transformation parameters dynamically. The disadvantage of this method is the same disadvantage present in the ICP algorithm, to perform well it requires a good initial approximation of the solution.

In the specific case of CityGML, Goebbels et al. [19] propose an extension of the well-known ICP to perform point-to-model registration based on point-to-plane registration. This approach only considers points on the perimeter of the buildings. The main contribution of this approach is the speed since point-to-model ICP is faster than the point-to-plane ICP. Nevertheless, the CityGML models used contained no terrain information. Therefore, the ground plane is the lowest point of the building and some walls might

intersect with the real terrain, increasing the error of the algorithm. An additional disadvantage of this method is that it expects the point cloud to be coarsely register with the model.

Another approach by Goebbels et al. [21] explores the possibility to find features that help to perform the registration. The approach first rotates the point cloud to align the walls to the z-axis. Then it projects all the points onto the xy-plane to obtain a 2D image of the point cloud viewed from above. After that, it detects the corners in the 2D image, and having the corners of the CityGML model it applies a Mixed Integer Linear Program to perform the registration. Finally, a Linear Program relaxation is used to fine-tune the coarse result. Nonetheless, the point cloud has to be already coarsely registered in order to obtain a successful and fast result.

An extension of the work in [21] has been made by Goebbels et al. [20], where the steps to perform the registration are the same, but instead of corners, lines are used in the Mixed Integer Linear Program. Furthermore, no fine registration is performed at the end. Unfortunately, it has the same disadvantage as the previous approach: the point cloud has to be already coarsely registered to obtain a successful and fast result. Although, this is not a disadvantage if the desired method will be just used to perform a fine registration.

2.3 Others

There are other approaches that attempt to solve different matching problems with or without correspondences. These approaches cannot be classified in the previous categories because their input data is not composed of two point clouds or one point cloud and a 3D model. Some of these approaches are collected into this category.

Breuel [11] proposes the use of matchlist-based BnB techniques to solve geometric matching problems without correspondences. For example, point and line segment matching in 2D and 3D.

Bazin et. al [8] present an approach that combines Linear Programming and BnB procedures to achieve global optimality. The proposed algorithm receives two different sets of features extracted from two images and outputs the correspondences between the features verifying the appearance similarity and geometric constraints. In the end, it not only computes the feature correspondences but also computes an estimation of a geometric transformation between the features and identifies the inliers/outliers.

Brown et. al [12] extend the idea of the two previous approaches and present a 2D-3D registration method that does not need correspondences, by using a BnB algorithm. They also propose a deterministic annealing algorithm to reduce computation time. The proposed method works with points or lines, but not with a combination of both. Therefore, Brown et. al [13] propose a framework to solve the 2D-3D registration problem without feature correspondences that is able to use points, lines, or a combination of both.

Windheuser et al. [43] propose the use of Integer Linear Programming to find correspondences between non-rigid 3D shapes. They use an Integer Linear Program to minimize the elastic thin-shell energy required to deform one shape into the other one. Their method can be improved by a relaxation of the Integer Linear Program and by including feature descriptors.

2.4 Limitations of previous work

ICP and its variants provide simple and easily-implemented iterative methods, but these algorithms can converge to false local optima [41]. Furthermore, they do not scale well with the number of points and they are not differentiable, therefore they cannot be integrated into end-to-end deep learning pipelines [39]. And the principal disadvantage of these methods is that they require a good initial approximation to the solution to perform well.

RANSAC and its variations cannot guarantee any global optimality in their final results [8]. According to [36], RANSAC has two main issues. The first one is its computational complexity, which makes it time-consuming. The second one is its randomness during sampling that can lead to inaccurate results. These two disadvantages worsen with the number of outliers.

The globally optimal methods such as Linear programming and BnB based methods are very time-consuming. Therefore, their application in real-time tasks is limited [39].

In the last years, there has been a boom in the use of neural networks and deep neural networks to solve 2D problems. Therefore, their use in 3D tasks has been explored too. Unfortunately, methods based on deep learning generally suffer from issues regarding the generalization ability and the requirement of massive training data [36].

In general, one thing that can be noticed is that none of this methods offers a direct solution for the coarse registration of a model with a point cloud. The only solution available for this is to sample a point cloud from the model and apply a global method for two point clouds. Unfortunately, this solution does not leverage the complete information contained in the model. However, some of the ideas of the approaches presented in this chapter can be used to implement alternative coarse registration methods of a point cloud with a CityGML model.

3

Methodology

In this section, a brief description of the Data and the definition of the Registration Problem will be given. The structure of a CityGML model is complex, and their complete explanation is out of the scope of this project, For detailed information about CityGML please refer to Gröger et. al [22] or the official websites.

3.1 Data

The CityGML models are public data in some states in Germany, for example in Nordrhein-Westfalen [5]. Other states of Germany like Hessen [4] or Bayern [2] charge money for their Models. That is not relevant for this project because the CityGML Models are received from the server of the A-DRZ project.

Besides the representation of the building, CityGML models provide other interesting information about the buildings. For this work, the intersection of the model with the terrain is the most relevant information. The terrain intersection is provided as an array of 3D points that represent some points on the walls that intersect with the ground.

The point cloud data is collected and provided by the A-DRZ project partners that develop and operate the actual Robots. These are mainly the University of Bonn [7] and the Technical University of Darmstadt [6], which have integrated corresponding laser scanners and Lidar sensors on their platforms.

Figure 3.1 and Figure 3.2 show a CityGML model and its corresponding point cloud before being registered.

Unfortunately, there is no ground truth for the registration of the data available. Therefore, the results are evaluated manually by a person who decides whether the registration was successful or not.

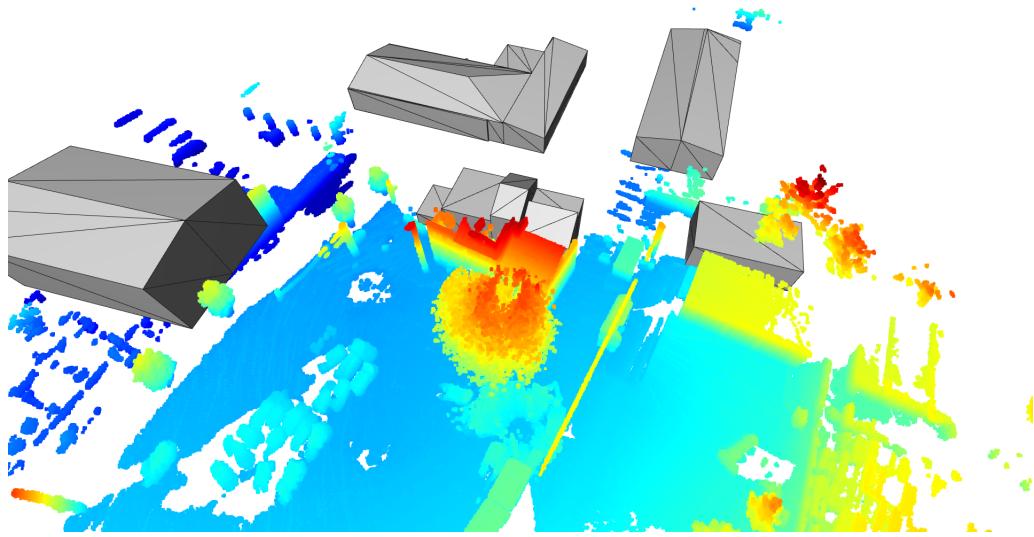


Figure 3.1: Front view of a point cloud and a model before registration.

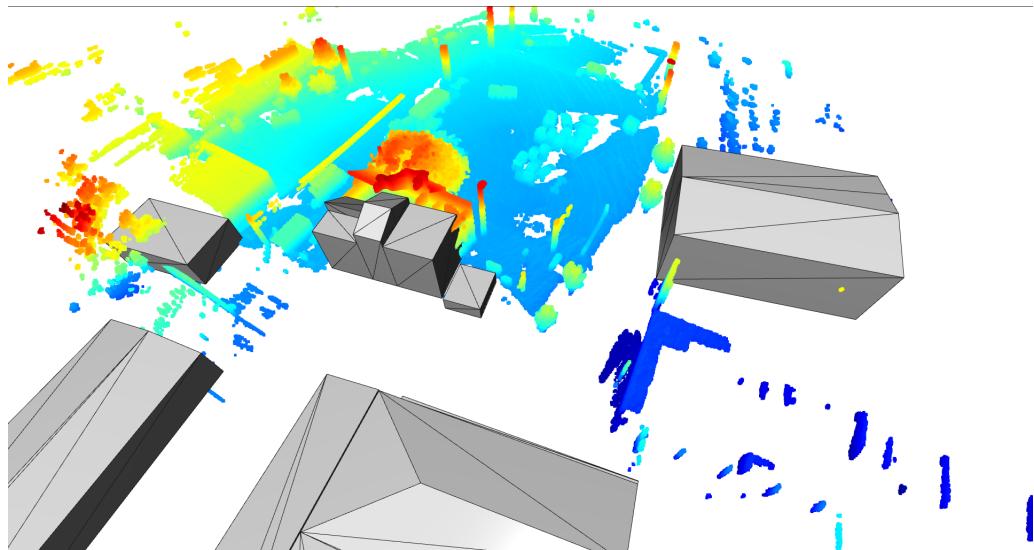


Figure 3.2: Back view of a point cloud and a model before registration.

3.2 Setup

The main setup is depicted in Figure 1.2. The input of the system is a CityGML model and its corresponding point cloud. The output of the registration algorithm is a transformation that aligns the point cloud with the CityGML model. This transformation is then used to visualize the alignment of the input data for its proper evaluation.

3.3 Registration Problem in 3D

In general, the registration task consists of finding the transformation that aligns the input CityGML model with the input point cloud.

More specifically, given a source point set $P = \{p_i\}_{i=1}^N$, where $p_i \in \mathbb{R}^4$, and a target point set $Q = \{q_j\}_{j=1}^M$, where $q_j \in \mathbb{R}^4$, the registration task consists of finding a transformation $T \in SE(3)$ that minimizes:

$$\sum_{i=1}^N \min_{j \in 1 \dots M} d(Tp_i, q_j) \quad (3.1)$$

where $d(Tp_i, q_j)$ is the Euclidian distance between Tp_i and q_j . p_i and q_j are points in homogeneous coordinates. Therefore, their fourth element is 1.

In this case, the point set P corresponds to the point cloud, while the point set Q corresponds to the CityGML model. It is clear that P is formed by the same points that form the point cloud, but the set Q could be a sampled point cloud from the CityGML model or other characteristic points of the CityGML model.

The difficulty of this task is to find correspondences between P and Q that minimizes Equation 3.1. If the correspondences were defined, the registration would be pretty straightforward and correct.

3.4 Registration Problem in 2D

To simplify the problem, one can transform the 3D registration problem into a 2D registration problem by projecting the point cloud and the CityGML model onto the xy-plane.

The registration problem in 2D remains the same as in 3D, but $p_i \in \mathbb{R}^3$, $q_j \in \mathbb{R}^3$, and $T \in SE(2)$. Again, p_i and q_j are points in homogeneous coordinates and their third element is always 1.

3.5 Registration Problem as Mixed Integer Linear Program

The 3D registration problem defined in section 3.3 can be redefined as a Mixed Integer Linear Program. The exact definition is given in [37].

The 2D registration problem defined in section 3.4 can be redefined as a Mixed Integer Linear Program in the same way. The objective is to maximize

$$\sum_i^N \sum_j^M x_{ij} \quad (3.2)$$

subject to

$$\sum_i^N x_{ij} \leq 1 \quad (3.3)$$

$$\sum_j^M x_{ij} \leq 1 \quad (3.4)$$

$$\sum_i^N \sum_j^M l_0 \leq \epsilon_d + b(1 - x_{ij}) \quad (3.5)$$

$$\sum_i^N \sum_j^M -l_0 \leq \epsilon_d + b(1 - x_{ij}) \quad (3.6)$$

$$\sum_i^N \sum_j^M l_1 \leq \epsilon_d + b(1 - x_{ij}) \quad (3.7)$$

$$\sum_i^N \sum_j^M -l_1 \leq \epsilon_d + b(1 - x_{ij}) \quad (3.8)$$

$$\sum_i^N \sum_k^N \sum_j^M \sum_l^M x_{ij} + x_{kl} \leq 1 \quad \text{for } |d(p_i, p_k) - d(q_j, q_l)| \geq \epsilon_l \quad (3.9)$$

$$\begin{aligned} \delta_1^- &\leq t_{11} \leq \delta_1^+ \\ \delta_2^- &\leq t_{12} \leq \delta_2^+ \\ t_x^- &\leq t_{13} \leq t_x^+ \\ t_y^- &\leq t_{23} \leq t_y^+ \\ t_{11} &= t_{22} \\ t_{21} &= -t_{12} \end{aligned} \quad (3.10)$$

$X \in \mathbb{R}^{N \times M}$ is the matrix that represents the correspondences. That means that $x_{ij} = 1$ if the point p_i in the source image corresponds to the point q_j in the target image, otherwise $x_{ij} = 0$.

$L = (q_i - Tp_i)$ and the constant ϵ_d is a threshold to remove outliers. In other words, ϵ_d is the maximum distance allowed between a target point and its corresponding source point after transformation.

The constant b is some large number. For example, b can be two times the maximum distance between points in the source and the target image.

The Equation 3.9 is used to preserve lines between points after the transformation. If $|d(p_i, p_k) - d(q_j, q_l)| \geq \epsilon_l$, then either p_i corresponds to q_j or p_k corresponds to q_l . The constant ϵ_l depends on ϵ_d and according to [37], $\epsilon_l \geq 2\sqrt{3}\epsilon_d$ needs to be satisfied because ϵ_d is the allowed rigid transformation error.

The elements of T are constrained in Equation 3.10, but the rest of the elements are constant: $t_{31} = 0$, $t_{32} = 0$, and $t_{33} = 1$.

4

Solution

Due to causes unrelated to this project, the available data was just one CityGML model and its corresponding point cloud. Therefore, a solution was designed based on this data and its characteristics.

Moreover, at the time this work was started, there was no GPS information of the point clouds used in the registration, just of the CityGML model. Therefore, it was not possible to perform a coarse registration with GPS information and the coarse registration need to be done by other means. All this together made deciding on a specific approach difficult.

Later, in the middle phase of the project, another point cloud was available but its corresponding CityGML model was not. However, instead of a CityGML model, a 3D model of the site was available in Polygon File Format (ply). Because of the different information contained in the PLY file a different preprocessing was needed.

Although (deep) learning methods are very attractive due to their boom in recent years, they are unthinkable to solve our problem. The main reason for this is the data needed to train such an approach.

Goebbels et al. [20, 21] successfully propose the use of Mixed Integer Linear Programming to register a point cloud with a CityGML model, based on lines and points. Mixed Interger Linear Programming has been used to solve other registration problems like the registration between two point clouds [37], registration between 3D shapes [43], and 2D registration [8].

Therefore, the solution proposed is to detect the walls in the point cloud and the 3D model by projecting both onto the xy-plane, and then use the angles of the corners of the walls (where the walls intersect) in a Mixed Integer Linear Program to find correspondences and a transformation that aligns the point cloud with the 3D Model at the same time. The exact details are given in the next section.

From now on, the point cloud could be referred to as the source, and the 3D model as the target.

4.1 Proposed algorithm

The main steps of the proposed algorithm are listed as follows:

1. Projection of the point cloud onto the xy-plane.
2. Projection of the 3D model onto the xy-plane.
3. Detection of line segments.
4. Detection of line intersections and their angles.

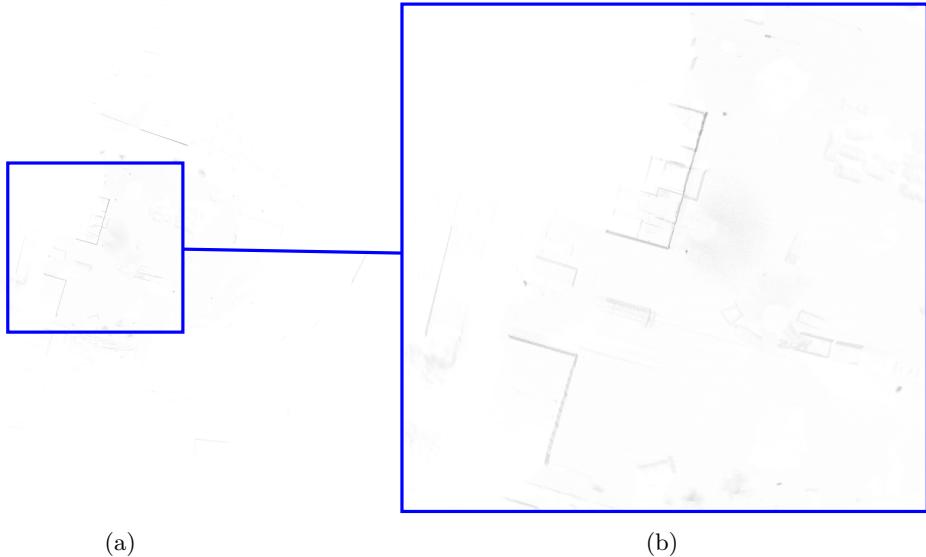


Figure 4.1: Projection of the source onto the xy-plane.

5. Identification of possible correspondences between angles detected.
6. Alignment of both projections onto the xy-plane.
7. Alignment on the z-axis.

4.1.1 Projection of the point cloud onto the xy-plane

To simplify the registration task, this is transported from the 3D space to the 2D space. As we can see in Figure 3.1 and Figure 3.2, the z-axis of the source and the target is already correctly aligned. Which allows for a direct projection of the source onto the xy-plane to detect walls.

To project the source onto the xy-plane, one can create an empty image with enough size for the source and the target to fit. Then, one can sum the number of points in the source with the same x- and y-coordinates and write the result in the corresponding pixel in the image. In this way, we can obtain a view of the source from below where line segments depict walls. Figure 4.1 shows the image obtained after the projection of the source.

4.1.2 Projection the 3D model onto the xy-plane

The difference between using CityGML Models and the PLY models relies on their projection onto the xy-plane. From the PLY model, a point cloud is sampled and projected onto the xy-plane as described in subsection 4.1.1.

Fortunately, the CityGML model already provides a terrain intersection. From this terrain intersection, only the x- and y-coordinates are used to make the projection. In an empty image, line segments are drawn from point to point so that a footprint is obtained.

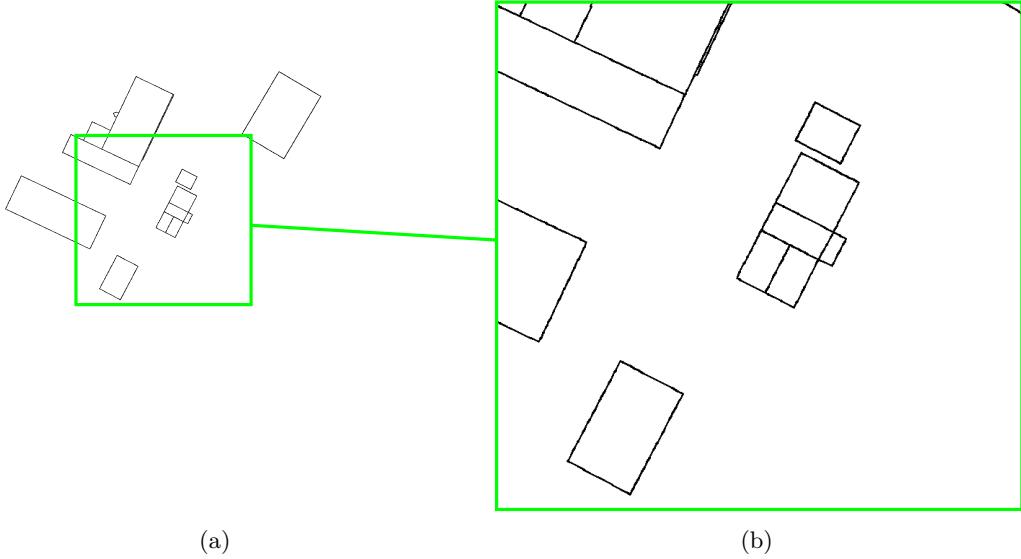


Figure 4.2: Projection of the target onto the xy -palne.

Figure 4.2 shows the image obtained after this process.

4.1.3 Detection of line segments

The projections of the source and the target are now 2D images. The source image could be noisy, but the target is clear and the line segments that represent the walls are very well defined. Therefore, to deal with the noise and remark the line segments of the source image an implementation of the technique proposed by Chaudhuri et al. [14], to detect blood vessels in retinal images, together with a binary threshold is used. The implementation of [14] used is the one provided by DIPlib [31]. Figure 4.3 shows the result after the filter is applied to Figure 4.1.

Line segments in both images are detected using the probabilistic hough line transform of OpenCV [10]. The number of line segments is reduced in four steps. Firstly, short line segments are deleted, i.e. line segments with lengths lower than a given threshold. Secondly, line segment s with similar slopes are identified. Then, from this set of line segments, subsets of line segments with close starting and ending points are created, and only the longest one of each set is preserved. The rest of the line segments are deleted. Thirdly, line segments with similar inclination angles are detected, and the same as in the second step is done. Finally, line segments with similar angles that are too close are merged by eliminating or extending one of the line segments.

4.1.4 Detection of line intersections and their angles

The line segments are converted to their homogeneous representation, and the intersections between them are calculated using the cross product as described in [24]. Only the intersections that are in the

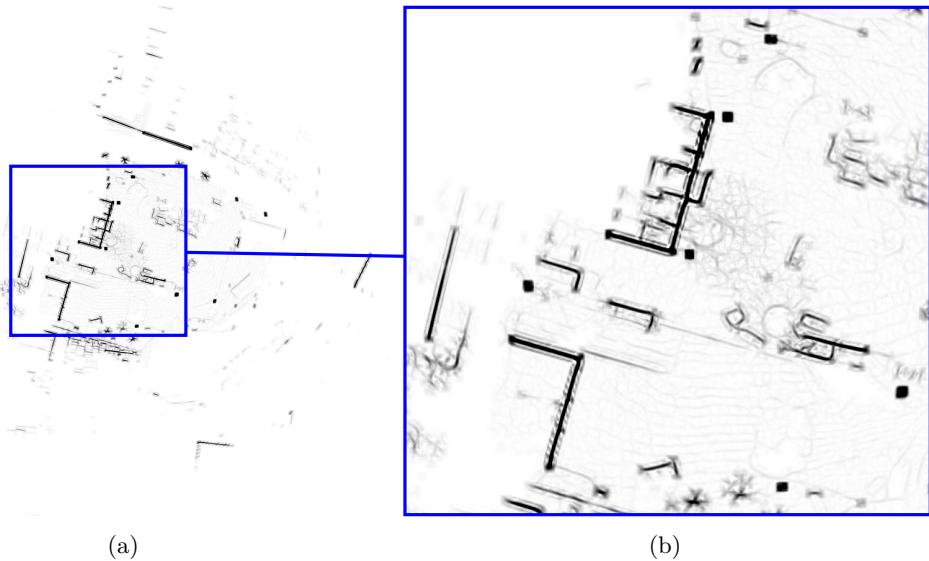


Figure 4.3: Source image after filter.

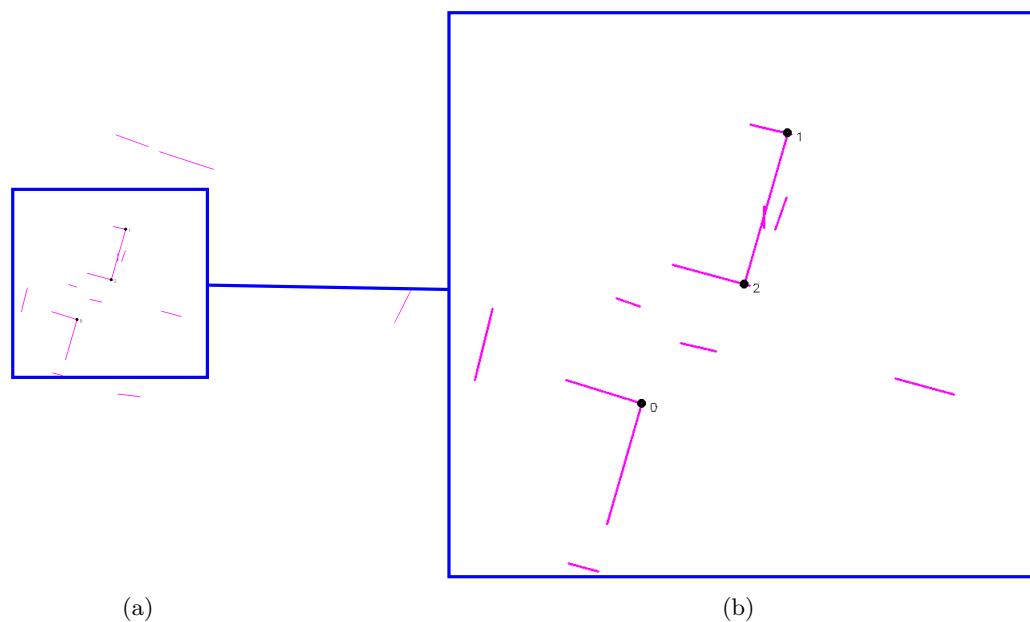


Figure 4.4: Source lines and intersections.

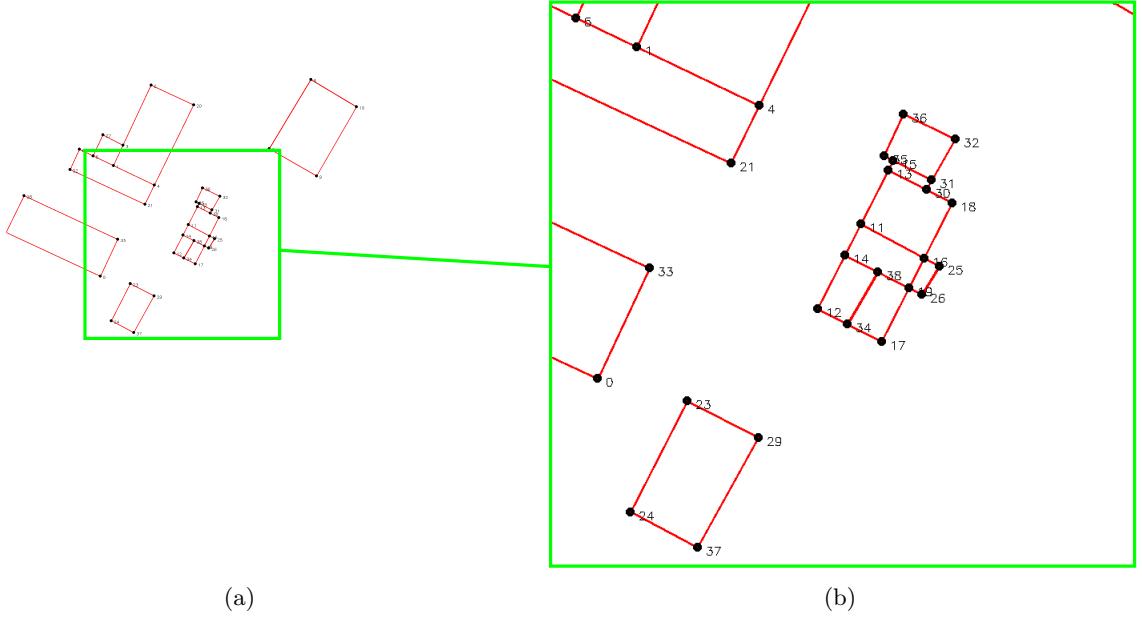


Figure 4.5: Target lines and intersections.

line segments are taken into account. Then, the angles of the intersections are computed.

The intersections detected in the source image and the target image are now the set of points P and Q , respectively, as described in section 3.4. Figure 4.4 shows P and Figure 4.5 shows Q .

4.1.5 Identification of possible correspondences between angles detected

The angles of the intersections in the source image and the target image are compared, and if they are similar to each other, they represent a possible match.

More formally, there is a matrix $C \in \mathbb{R}^{N \times M}$ that represents the possible correspondences. That means that $c_{ij} = 1$ if the angle i in the source image is similar to the angle j in the target image. Otherwise, $c_{ij} = 0$.

4.1.6 Alignment of both projections onto the xy-plane

The Mixed Integer Linear Program of section 3.5 is solved with the intersections detected. The objective is changed to speed up the Mixed Integer Linear Program. The new objective is to maximize

$$\sum_i^N \sum_j^M x_{ij} m_{ij} \quad (4.1)$$

subject to the constraints presented in section 3.5. This change in the objective adds additional information about the correspondences and significantly reduces the computation time.

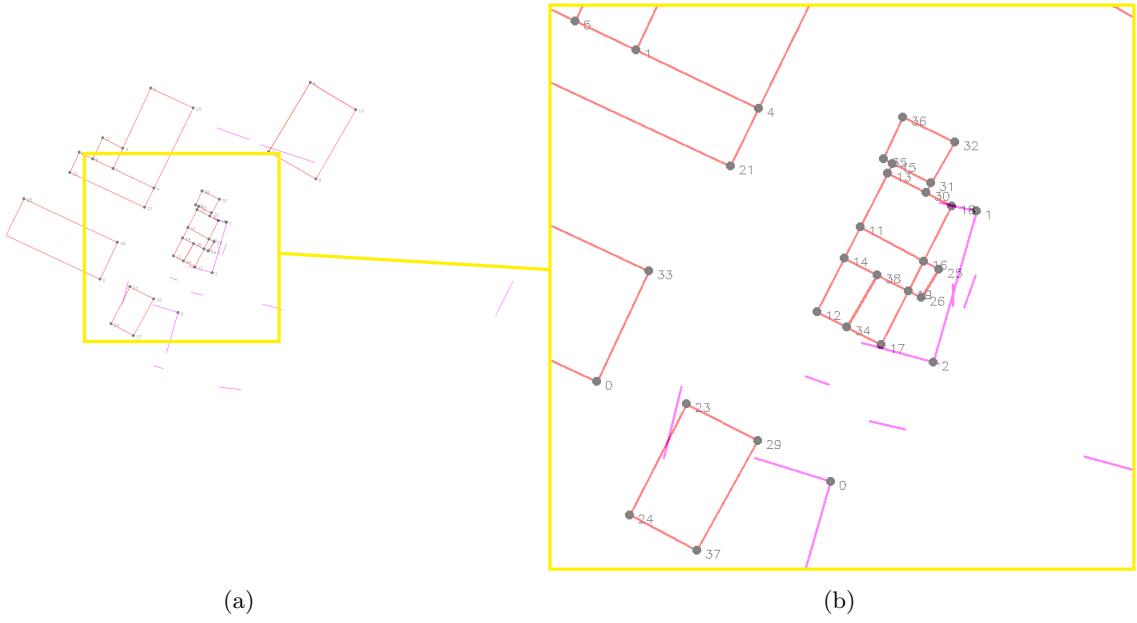


Figure 4.6: Initial position of the source image and the target image.

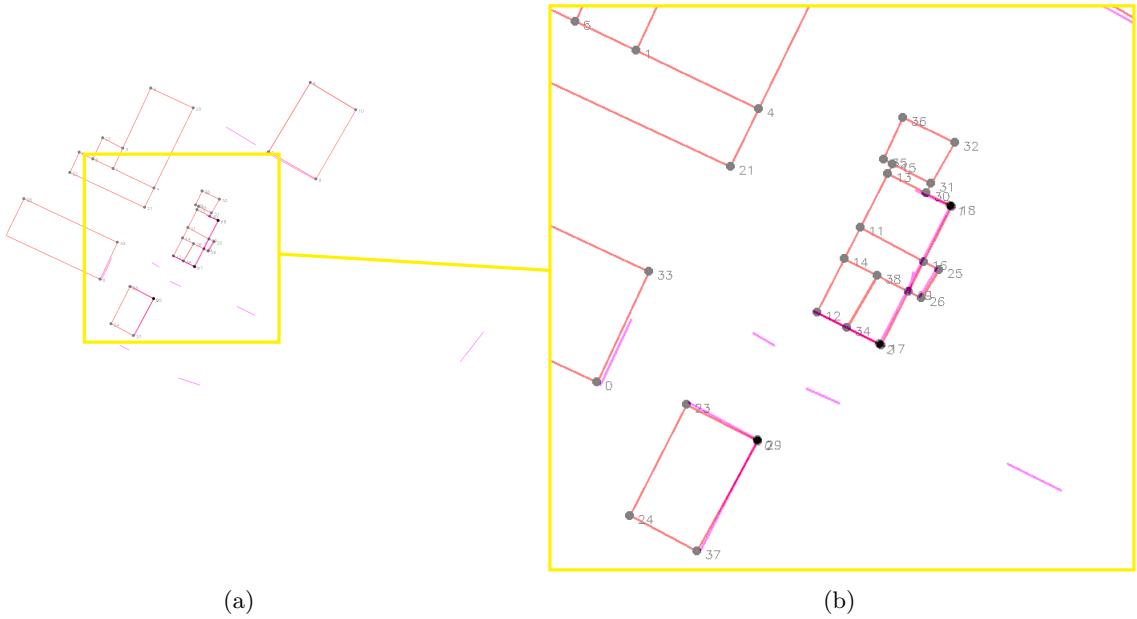


Figure 4.7: Source image and target image aligned.

The Mixed Integer Linear Program is solved and the result is a transformation T that aligns the source image with the target image, and a matrix X that represents the correspondences between P and Q .

Figure 4.6 shows the initial position of the source image and the target image. Figure 4.7 shows the result after the transformation of the source image with T .

Because the source image and the target image are a projection onto the xy-plane of the source and the target, respectively, T is used to align the source and the target on the x-axis and the y-axis by expanding it into the $SE(3)$.

4.1.7 Alignment on the z-axis

Sometimes a translation across the z-axis is needed to align the source and the target completely. The correspondences found in subsection 4.1.6 can be used to find the l

If $x_{ij} = 1$ then the points p_i and q_j are used in the following way. The point p_i is used to find the lowest point p_i^l in the source with x- and y-coordinates similar to those of p_i . And the point q_j is used to find the lowest point q_j^l in the source with x- and y-coordinates similar to those of q_j . Then, the difference along the z-axis between p_i^l and q_j^l is computed. Finally, an average of all the differences computed is the translation t_z , which is then integrated into the transformation T .

4.2 Implementation details

The solution was implemented with Python. Therefore, the standard library "json" was used to read the CityGML models. Numpy [23] was used to compute array operations. Open3d [46] was used for the visualization of 3D models and point clouds. OpenCV [10] was used to process images. PyDIP [31] was used to filter the 2D projection of the model and to make the lines more visible. Python-MIP [38] was used to solve the Mixed Integer Linear Program described in subsection 4.1.6.

The parameter ϵ_d used in Equation 3.5, Equation 3.6, Equation 3.7, and Equation 3.8 was set to 20 cm, and the parameter b was set to two times the maximum distance between points in the source and the target image. Moreover, the parameter ϵ_l was set to $\epsilon_l \geq 2\sqrt{3}\epsilon_d$.

The upper and lower bounds of T showed in Equation 3.10 were set as follows: $\delta_1^- = 0.7$, $\delta_1^+ = 1.3$, $\delta_2^- = -0.3$, $\delta_2^+ = 0.3$, $t_x^- = -b/2$, $t_x^+ = b/2$, $t_y^- = -b/2$, and $t_y^+ = b/2$. However, these bounds can be changed to allow a different transformation (different rotation or different translation). Because the universities that provide the point cloud data agreed to send it in a position and orientation close to the model's position, the parameters chosen allow only a small rotation and a small translation.

5

Evaluation and Results

Implementation and measurements.

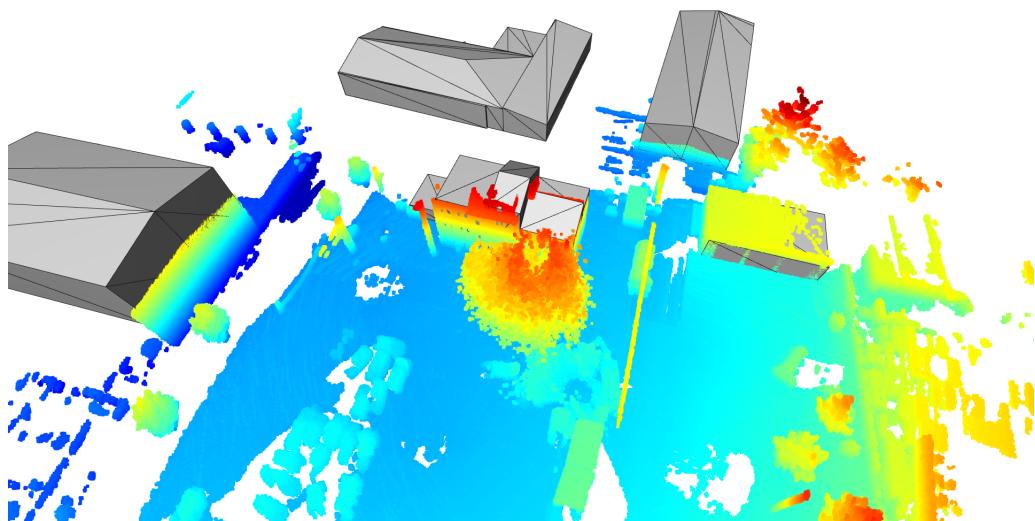


Figure 5.1: Front view of a point cloud and a model after registration.

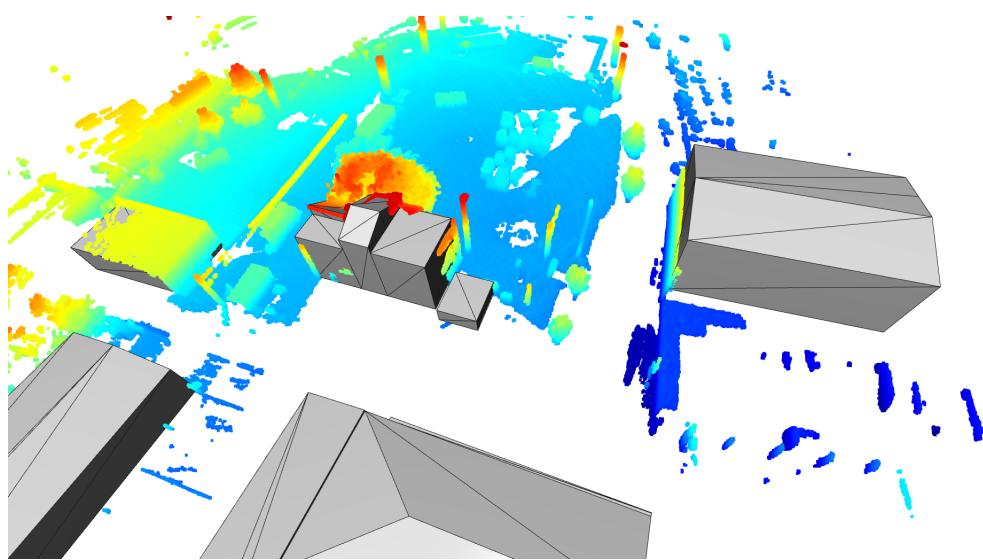


Figure 5.2: Back view of a point cloud and a model after registration.

6

Conclusions

6.1 Contributions

- The solution presented in this work can be used and improved for the registration of a 3D Model with a point cloud using different kinds of features

6.2 Lessons learned

- How to make a research in a specific field.
- How to work in a project that already has many specifications defined.
- How to work with third-party libraries in order to ease the implementation of a project.
- How to adapt yourself to the changes in the requirements of a project.

6.3 Future work

- Improve the angle detection. - The direction of the corners can be added to improve the performance.
- More features can be added. - There are two possible options to improve the performance of the system (reduce the time of registration)
 - One can work with 3D features and find directly a 3D transformation that aligns the CityGML-Model and the point cloud.
 - a Branch-and-Bounding [8, 11, 12, 13] approach could be applied to find a transformation that aligns the projections of the CityGML model and the point cloud onto the XY-plane or to find such a transformation directly with 3D features such as lines or corners.

A

Design Details

Your first appendix

B

Parameters

Your second chapter appendix

References

- [1] Kompetenzzentrum für Rettungsrobotik. URL <https://rettungsrobotik.de/>.
- [2] Bayerische Vermessungsverwaltung - Produkte - 3D-Produkte - 3D-Gebäudemodell. URL <https://www.ldbv.bayern.de/produkte/3dprodukte/3d.html>.
- [3] Deutscher Feuerwehrverband: Feuerwehr-Statistik. URL <https://www.feuerwehrverband.de/presse/statistik/>.
- [4] Cp.Landesvermessung_3DDaten — Hessische Verwaltung für Bodenmanagement und Geoinformation. URL <https://hvbg.hessen.de/geoinformation/landesvermessung/geotopographie/3d-daten>.
- [5] 3D-Gebäudemodell LoD2 - NRW. URL https://www.opengeodata.nrw.de/produkte/geobasis/3dg/1od2{_}gml/.
- [6] SIM.tud : FG Simulation, Systemoptimierung und Robotik / FB Informatik / TU Darmstadt. URL <https://www.sim.informatik.tu-darmstadt.de/index/>.
- [7] University of Bonn, Computer Science VI, Autonomous Intelligent Systems. URL <http://www.ais.uni-bonn.de/>.
- [8] Jean Charles Bazin, Hongdong Li, In So Kweon, Cédric Demonceaux, Pascal Vasseur, and Katsushi Ikeuchi. A branch-and-bound approach to correspondence and grouping problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(7):1565–1576, 2013. ISSN 01628828. doi: 10.1109/TPAMI.2012.264.
- [9] P J Besl and N D McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, feb 1992. ISSN 1939-3539. doi: 10.1109/34.121791.
- [10] G Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- [11] Thomas M. Breuel. Implementation techniques for geometric branch-and-bound matching methods. *Computer Vision and Image Understanding*, 90(3):258–294, jun 2003. ISSN 10773142. doi: 10.1016/S1077-3142(03)00026-2.
- [12] Mark Brown, David Windridge, and Jean Yves Guillemaut. Globally optimal 2D-3D registration from points or lines without correspondences. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 2015 Inter, pages 2111–2119, 2015. ISBN 9781467383912. doi: 10.1109/ICCV.2015.244.

-
- [13] Mark Brown, David Windridge, and Jean Yves Guillemaut. A family of globally optimal branch-and-bound algorithms for 2D–3D correspondence-free registration. *Pattern Recognition*, 93:36–54, sep 2019. ISSN 00313203. doi: 10.1016/j.patcog.2019.04.002. URL <https://doi.org/10.1016/j.patcog.2019.04.002>.
 - [14] Subhasis Chaudhuri, Shankar Chatterjee, Norman Katz, Mark Nelson, and Michael Goldbaum. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Transactions on Medical Imaging*, 8(3):263–269, sep 1989. ISSN 02780062. doi: 10.1109/42.34715.
 - [15] Li Ding and Chen Feng. DeepMapping: Unsupervised Map Estimation From Multiple Point Clouds. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.
 - [16] Gil Elbaz, Tamar Avraham, and Anath Fischer. 3D Point Cloud Registration for Localization Using a Deep Neural Network Auto-Encoder. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
 - [17] Martin A Fischler and Robert C Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, 24(6):381–395, 1981. ISSN 0001-0782. doi: 10.1145/358669.358692. URL <https://doi.org/10.1145/358669.358692>.
 - [18] Andrew W Fitzgibbon. Robust registration of 2D and 3D point sets. *Image and Vision Computing*, 21(13):1145–1153, 2003. ISSN 0262-8856. doi: <https://doi.org/10.1016/j.imavis.2003.09.004>. URL <http://www.sciencedirect.com/science/article/pii/S0262885603001835>.
 - [19] S Goebbels, R Pohle-Fröhlich, and P Pricken. Iterative Closest Point Algorithm for Accurate Registration of Coarsely Registered Point Clouds with CityGML Models. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 201–208, 2019.
 - [20] Steffen Goebbels and Regina Pohle-Froehlich. Line-based Registration of Photogrammetric Point Clouds with 3D City Models by Means of Mixed Integer Linear Programming. In *Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP)*, 2018.
 - [21] Steffen Goebbels, Regina Pohle-Fröhlich, and Philipp Kant. A Linear Program for Matching Photogrammetric Point Clouds with CityGML Building Models. In Natalia Kliewer, Jan Fabian Ehmke, and Ralf Borndörfer, editors, *Operations Research Proceedings 2017*, pages 129–134, Cham, 2018. Springer International Publishing. ISBN 978-3-319-89920-6.
 - [22] Gerhard Gröger, Thomas H Kolbe, Claus Nagel, and Karl-Heinz Häfele. *OGC City Geography Markup Language (CityGML) Encoding Standard*. Open Geospatial Consortium, 2.0.0 edition, 2012.
 - [23] Charles R Harris, K Jarrod Millman, St'efan J van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández

- del R'io, Mark Wiebe, Pearu Peterson, Pierre G'erard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, sep 2020. doi: 10.1038/s41586-020-2649-2. URL <https://doi.org/10.1038/s41586-020-2649-2>.
- [24] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2003.
- [25] Xiaoshui Huang, Guofeng Mei, and Jian Zhang. Feature-metric Registration: A Fast Semi-supervised Approach for Robust Point Cloud Registration without Correspondences. *arXiv preprint arXiv:2005.01014*, 2020.
- [26] Changmin Kim, Joohyuk Lee, Minwoo Cho, and Changwan Kim. Fully automated registration of 3D CAD model with point cloud from construction site. In *28th International Symposium on Automation and Robotics in Construction, Seoul, Korea*, pages 917–922, 2011.
- [27] Changmin Kim, Hyojoo Son, and Changwan Kim. Fully automated registration of 3D data to a 3D CAD model for project progress monitoring. *Automation in Construction*, 35:587–594, 2013. ISSN 0926-5805. doi: <https://doi.org/10.1016/j.autcon.2013.01.005>. URL <http://www.sciencedirect.com/science/article/pii/S0926580513000071>.
- [28] Weimin Li and Pengfei Song. A modified ICP algorithm based on dynamic adjustment factor for registration of point cloud and CAD model. *Pattern Recognition Letters*, 65:88–94, 2015. ISSN 0167-8655. doi: <https://doi.org/10.1016/j.patrec.2015.07.019>. URL <http://www.sciencedirect.com/science/article/pii/S0167865515002287>.
- [29] J Lu, W Wang, H Shao, and L Su. Point Cloud Registration Algorithm Fusing of Super 4PCS and ICP Based on the Key Points. In *2019 Chinese Control Conference (CCC)*, pages 4439–4444, 2019. doi: 10.23919/ChiCC.2019.8866059.
- [30] Weixin Lu, Guowei Wan, Yao Zhou, Xiangyu Fu, Pengfei Yuan, and Shiyu Song. DeepICP: An End-to-End Deep Neural Network for 3D Point Cloud Registration. *CoRR*, abs/1905.0, 2019. URL <http://arxiv.org/abs/1905.04153>.
- [31] Cris Luengo and Wouter Caarls. DIPlib: Quantitative Image Analysis in C++, MATLAB and Python. URL <https://github.com/DIPlib/diplib>.
- [32] Nicolas Mellado, Dror Aiger, and Niloy J Mitra. SUPER 4PCS Fast Global Pointcloud Registration via Smart Indexing. *Computer Graphics Forum*, 33(5):205–215, 2014. doi: 10.1111/cgf.12446.
- [33] Dhanya S. Pankaj and Rama Rao Nidamanuri. A robust estimation technique for 3D point cloud registration. *Image Analysis and Stereology*, 35(1):15–28, 2015. ISSN 18545165. doi: 10.5566/ias.1378.
- [34] Charles Ruizhongtai Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. Technical report, 2017.

-
- [35] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. In I Guyon, U V Luxburg, S Bengio, H Wallach, R Fergus, S Vishwanathan, and R Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 5099–5108. Curran Associates, Inc., 2017. URL <http://papers.nips.cc/paper/7095-pointnet-deep-hierarchical-feature-learning-on-point-sets-in-a-metric-space.pdf>.
 - [36] S Quan and J Yang. Compatibility-Guided Sampling Consensus for 3-D Point Cloud Registration. *IEEE Transactions on Geoscience and Remote Sensing* (2020), pages 1–13, 2020. ISSN 1558-0644. doi: 10.1109/TGRS.2020.2982221.
 - [37] Shizu Sakakubara, Yuusuke Kounoike, Yuji Shinano, and Ikuko Shimizu. Automatic range image registration using mixed integer linear programming. In *Asian Conference on Computer Vision*, pages 424–434. Springer, 2007.
 - [38] H.G. Santos and T.A.M. Toffolo. Python MIP (Mixed-Integer Linear Programming) Tools. URL <https://github.com/coin-or/python-mip>.
 - [39] Vinit Sarode, Xueqian Li, Hunter Goforth, Yasuhiro Aoki, Animesh Dhagat, Rangaprasad Arun Srivatsan, Simon Lucey, and Howie Choset. One Framework to Register Them All: PointNet Encoding for Point Cloud Alignment. *arXiv preprint arXiv:1912.05766*, 2019.
 - [40] Aleksandr V Segal, Dirk Haehnel, and Sebastian Thrun. Generalized-ICP. *Robotics: science and systems*, 2:435, 2009.
 - [41] Yue Wang and Justin M Solomon. Deep Closest Point: Learning Representations for Point Cloud Registration. In *The IEEE International Conference on Computer Vision (ICCV)*, 2019.
 - [42] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *ACM Transactions on Graphics*, 38(5):Article 146, 2019. ISSN 15577368.
 - [43] Thomas Windheuser, Ulrich Schlickwei, Frank R Schimdt, and Daniel Cremers. Large-Scale Integer Linear Programming for Orientation Preserving 3D Shape Matching. *Computer Graphics Forum*, 30(5):1471–1480, 2011. doi: 10.1111/j.1467-8659.2011.02021.x. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1467-8659.2011.02021.x>.
 - [44] J Yang, H Li, D Campbell, and Y Jia. Go-ICP: A Globally Optimal Solution to 3D ICP Point-Set Registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(11):2241–2254, nov 2016. ISSN 1939-3539. doi: 10.1109/TPAMI.2015.2513405.
 - [45] Zi Jian Yew and Gim Hee Lee. 3DFeat-Net: Weakly Supervised Local 3D Features for Point Cloud Registration. In Vittorio Ferrari, Martial Hebert, Cristian Sminchisescu, and Yair Weiss, editors, *European Conference on Computer Vision (ECCV) 2018*, pages 630–646, Cham, 2018. Springer International Publishing. ISBN 978-3-030-01267-0.

References

- [46] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Open3D: A modern library for 3D data processing. *arXiv preprint arXiv:1801.09847*, 2018.