Improving Progress Monitoring by Fusing Point Clouds, Semantic Data and Computer Vision

Alex Braun[a,c], Sebastian Tuttas[b,c], André Borrmann[a,c], Uwe Stilla[b,c]

[a] *Chair of Computational Modeling and Simulation, Technical University of Munich, Germany*
[b] *Chair of Photogrammetry and Remote Sensing, Technical University of Munich, Germany*
[c] *Leonhard Obermeyer Center, TUM Center of Digital Methods for the Built Environment*

Abstract

Automated construction-progress monitoring enables the required transparency for improved process control, and is thus being increasingly adopted by the construction industry. Many recent approaches use Scan-to/vs-BIM methods for capturing the as-built status of large construction sites. However, they often lack accuracy or are incomplete due to occluded elements and reconstruction inaccuracies. To overcome these limitations and exploit the rich project knowledge from the design phase, the authors propose taking advantage of the extensive geometric-semantic information provided by Building Information Models. In particular, valuable knowledge on the construction processes is inferred from BIM objects and their precedence relationships. SfM methods enable 3D building elements to be located and projected into the picture's 2D coordinate system. On this basis, the paper presents a machine-learning-based object-detection approach that supports progress monitoring by verifying element categories compared to the expected data from the digital model. The results show that, depending on the type of construction and the type of occlusions, the detection of built elements can rise by up

to 50% compared to an SfM-based, purely geometric as-planned vs. as-built comparison.

Keywords: Construction progress monitoring, BIM, point clouds, semantic and temporal knowledge, deep learning

---

1. Introduction

1.1. Automated progress monitoring

Construction progress monitoring is currently still performed mostly manually, in a laborious and error-prone non-automated process. To prove that all works have been completed as agreed contractually, all completed tasks must be documented and monitored. Complete and detailed monitoring techniques are required for large construction sites where the entire construction area and the number of subcontractors become too large for manual tracking to be efficient. Detecting possible deviations from the schedule provides a benchmark for the performance of the construction site. Regulatory matters add to the requirement of keeping track of the current status on the site. The ongoing establishment of building information modeling (BIM) technologies in the planning of construction projects facilitate the application of digital methods also in the execution phase. In an ideal implementation of BIM, all relevant information on materials, construction methods, and even the process schedule are interlinked. On this basis, it is possible to estimate project costs and project duration more precisely than with conventional methods [1].

On top of the digitized construction design process, recent advancements for capturing the as-built geometry by laser scanning [2] or photogrammetry [3] allow using the resulting point cloud data to be compared against the as-planned model. Photogrammetry, in particular, has gained more attention with the broader availability of Unmanned Aerial Vehicles (UAVs), making this method more flexible in terms of camera positions [4]. The main idea is not to use laser scanners but conventional camera equipment on construction

sites to capture the current construction state ("as-built"). Since the acquisition from different perspectives is significantly faster than laser scanners, the building can be captured in a comprehensive manner with comparatively low effort. As soon as a sufficient number of images from different points of view are available, a 3D point cloud can be produced using Structure from Motion (SfM) methods [5]. Finally, the point cloud, representing one particular observation time-point, can be compared against the geometry of the Building Information Model.

1.2. Problem statement

Currently, the detection of built elements using SfM methods and other point-cloud-based approaches faces several challenges:

1.2.1. As-planned modeling vs. as-built construction

As introduced in Braun et al. [6], the as-planned model is represented by a 4D building information model (see Figure 1 a)). All 4D construction processes are linked to their associated elements, allowing for statements regarding the expected construction state at any given observation time. As the relevant model and point cloud are co-registered, an initial detection algorithm can compare the model's geometry with the 3D information from the point cloud. During the construction phase, the actual as-built process can deviate from the original as-planned process. To clarify this deviation, Figure 1 depicts the digital model from one of the test sites, a corresponding UAV-aerial image, and the generated point cloud (a), c) and d)). Accordingly, the as-planned 4D model does not represent the real construction sequencing. This problem is also described in Huhnt et al. [7], Tulke [8].

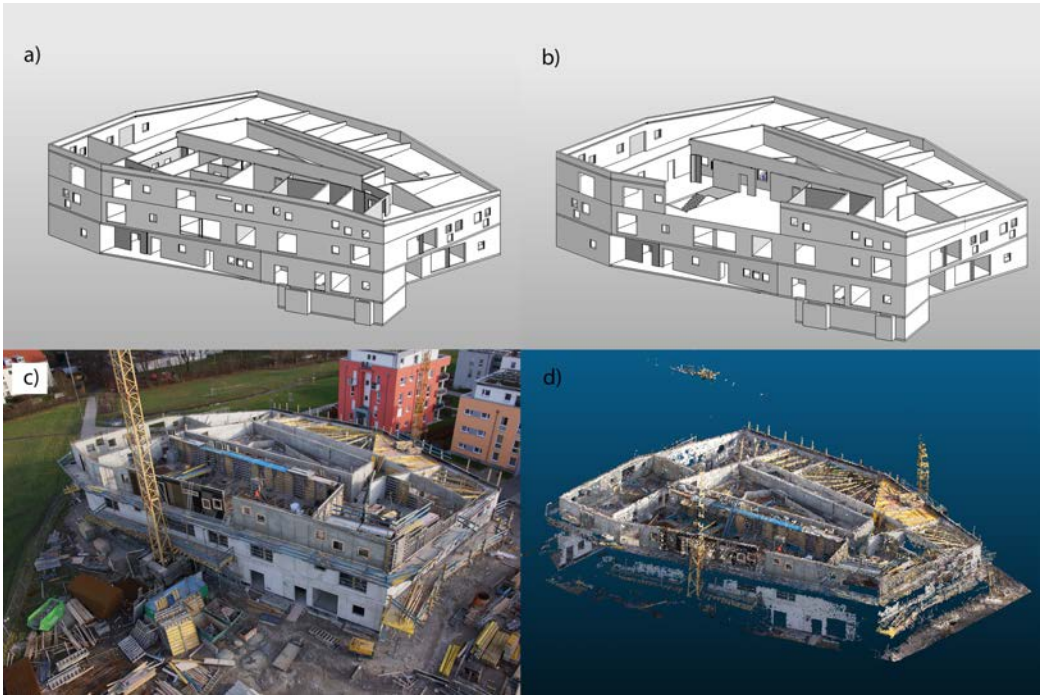


Figure 1: Process modeling problems depicted by a) as-planned modeling, b) as-built modeling, c) as-built image, d) as-built point cloud on sample construction site.

In comparison, Figure 1 b) shows the correct corresponding as-planned model for the given timestamp, with all subsequent elements being removed from the as-scheduled model.

1.2.2. Reconstruction

The monitoring of construction sites by applying photogrammetric methods has become common practice. Currently, several companies (for example, Pix4D or DroneDeploy) provide commercial solutions for end users that permit the generation of 3D meshes and point clouds from UAV or other image-based site observations. All these methods provide proper solutions for finished construction sites or visible elements of interest.

However, UAV-based monitoring of construction sites exhibits several problems. On the one hand, photogrammetric methods are sensitive to low-structured surfaces like monochrome painted walls, or windows. Because of the used method, each element needs to be visible from multiple (at least two) different points of view. Thus, elements inside of a building cannot be reconstructed as they are not visible from a UAV flying outside of the building. Monitoring inside of a building is currently still the subject of much research [9] and not available via an automated method, as localization in such mutable areas like construction sites is hard to tackle. These problems lead to holes or misaligned points in the final point cloud, which hinders the accurate and precise detection of building elements. On the other hand, laser scanning requires many acquisition points and takes significantly more time and manual effort in acquisition.

1.2.3. Occlusions

Finally, both techniques are challenged by occlusions for regions that are not visible during construction. The as-built 3D point cloud with n points holds all respective coordinates but also color information based on the feature's pixel color value in the initial image. The value n depends on many factors such as

- lighting conditions
- feature detection from different points of view
- surface textures
- amount and resolution of the images taken

A point cloud from one timestamp on one of our test construction sites can be seen in Figure 1 d). Besides scaffolding and formwork, various holes are visible in the point cloud that exist due to insufficient image quality for reconstruction or occlusions. The depicted point cloud matches the expected quality for an as-built acquisition and is incomplete due to changing visibility conditions from working equipment and similar items. However, it is not sufficient for reliable results in a purely geometric as-planned vs. as-built comparison as significant parts of the actual building are occluded. As seen in figure 2, another problem lies in elements that are occluded by temporary construction elements. In particular, scaffolding and formwork occlude the direct view on walls or slabs, making it harder for algorithms to detect the current state of construction progress.

Figure 2: Occluded construction elements in generated point cloud caused by scaffolding, formworks, existing elements and missing information during the reconstruction process

Currently available methods do not take these problems into account and make only limited use of BIM-related information such as type of construction and the general structure of a building.

1.3. Contributions

In this paper, the authors propose a number of inter-related methods to tackle the aforementioned problems. Specifically, this paper presents the following contributions:

- Known technological dependencies of construction sequences are used to enrich the model by precedence relationships, by applying formal graph theory. This allows the inference of the existence of elements, if they have not been directly detected.

- A method is presented that makes use of the knowledge of construction

methods and 4D data to adjust the detection thresholds (as-planned vs. as-built deviations allowed) according to their expected construction stage. This permits the detection of elements that are currently under construction and are, for example, covered by formwork.

- We introduce a method based on visibility analysis to identify elements that are detectable from the identified camera positions. Deep learning on projected element positions in the 2D plane of the gathered images for the initial SfM process allows the detection rates of built elements to be further enhanced.

The combined application of these methods helps to significantly improve the accuracy of construction progress monitoring, as documented by the case studies presented in this paper.

The details of the individual methods are described in Section 3.

2. Related Work

2.1. Scan vs. BIM

Progress monitoring has become a heavily researched topic in recent years. Omar and Nehdi [10] provide an overview of these developments and compare the individual approaches:

The as-built status of a construction site is usually captured by laser scanners or cameras using SfM methods. Laser scanning has the advantage that 3D point measuring is fast and very accurate (within the range of millimeters). However, the equipment is heavy and requires trained personnel. Additionally, the setup at the point of observation is time-consuming and,

depending on the size of the construction site, many observation points are required to scan the whole construction site. Photogrammetric approaches produce less accurate point clouds in comparison to laser scanning and require significant computing power for the reconstruction. However, this method is more flexible and easier in its application, as camera equipment is standard, low-cost, and widely used on UAVs. Other devices, such as Microsoft Kinect, combine multiple sensors and can also be used for progress monitoring [11].

The registration of the acquired point cloud and corresponding as-planned geometry is either performed manually or semi-automatically, e.g. by point-to-point matching through Iterative Closest Point (ICP) algorithms. Here, the algorithm minimizes the distance between the points of the laser scan and the BIM geometry [12].

Currently, three methods are deemed to be established in the comparison with the as-planned status:

1. comparison of points from the as-planned geometry with as-built point clouds. These methods compare point clouds that are acquired by laser scanners [13, 14] or SfM methods and derived point clouds from as-planned surfaces [15]. Point proximity metrics mainly do this following a data-alignment process.
2. Feature detection in the acquired images from the as-built state. Using feature detection algorithms to assess the progress of as-planned elements (as the construction site evolves in time) by comparing measurements with dynamic thresholds learned through a Support Vector Machine (SVM) classifier, construction elements are directly identified from the acquired images [16].

3. Matching the as-planned geometry surfaces directly with the as-built points. Here, relevant points from the point cloud are directly matched onto triangulated surfaces of the as-planned model after using octree-based checks for occupied regions [17].

The first approaches involving object detection in laser-scanned point clouds were published by Bosché and Haas [2]. Turkan et al. [14] extend this system and uses it for progress estimation. Kim et al. [18] detect specific component types using a supervised classification based on Lalonde features derived from the as-built point cloud. An object is regarded as detected if the type matches the type present in the model. As above, this method requires that the model is sampled into a point representation. Zhang and Arditi [19] introduce a measure for deciding four cases (object not in place, point cloud represents a full object or a partially completed object or a different object) based on the relationship of points within the boundaries of the object and the boundaries of the shrunk objects. However, the authors test their approach in a very simplified artificial environment, which is significantly less challenging than the processing of data acquired on real construction sites. In Mahami et al. [20], SfM and Multi-View Stereo (MVS) algorithms are coupled with coded targets to improve the photogrammetric process itself. Ibrahim et al. [21] use a single camera approach and compare images taken during a specified period, and rasterize them. Individual elements are identified for each use case. Most publications focus on identifying one particular type of element like, for example, columns or walls. Indoor monitoring has been researched by several groups. Asadi et al. [22] propose a new method to localize and align the camera position and building model in a real-time

scenario. Kropp et al. [23] tried to detect in-door construction elements based on similarities. Turkan et al. [24] present an approach for detecting elements under construction that uses threshold extensions for those elements. Han and Golparvar-Fard [25] published another attempt to solve the problem of elements under construction. The focus lies on visibility issues, e.g., assuming that an anchor bolt for a column must be present, despite being invisible, as the column on top of it requires the anchor bolt for structural reasons. Further research has been conducted in regard to multi-layered elements and the introduction of construction sequencing [26].

Another critical aspect of the as-planned vs. as-built comparison is dependencies. Technological dependencies determine which element is depending on another element, meaning that it cannot be built after the first element is finished. Precedence relationships [27] can define these dependencies. Szczesny et al. [28] discuss a storage solution for these dependencies. The approach with regard to progress monitoring is presented in Braun et al. [29]. Hamledari et al. [30] introduce an IFC-based schedule updating workflow that relies on detected construction elements.

In their outlook, Turkan et al. [24] state that further improvements to their work should include color analysis or even image-based methods. Thus, the authors propose incorporating these techniques, as well as the use of semantic data like construction methods, model analysis using technological dependencies, and image-based deep learning, to further enhance the detection of elements in an as-planned vs. as-built comparison.

2.2. Computer vision and deep learning

Rising computational power has enabled significant advances in machine learning in recent years. Deep learning [31] and especially Convolutional Neuronal Networks (CNN) provide solutions for training computers to learn patterns and apply them to previously unseen data. In this context, computer vision is a heavily researched topic that has received even more attention through recent advances driven by, for example, the needs of autonomous vehicles. Image analysis in the construction sector, on the other hand, is a rather new topic. Up to now, the main focus has been on defect detection (for example, cracks) in construction images [32]. Crack detection for asphalt roads has also been the subject of research [33]. Since one of the critical aspects of machine learning is the collection of large datasets, current approaches focus on data gathering. In the scope of automated progress monitoring, Han and Golparvar-Fard [34] published an approach for labeling image datasets based on the commercial service Amazon Turk. Braun and Borrmann [35] introduce a method for automated image labeling by fusing semantic and photogrammetric data.

Regarding the application of deep learning for construction progress tracking, Chi and Caldas [36] used initial versions of neural networks to detect construction machinery on images, and Kim et al. [15] used ML-based techniques for construction progress monitoring. They analyzed images by filtering them to remove noise and uninteresting elements, so as to focus the comparison on relevant construction processes. Hamledari et al. [37] applied CV approaches to indoor appliances like electrical outlets and insulation.

These approaches are currently mostly independent from the actual build-

ing model, as orientation and scale with respect to the digital twin are neglected or assumed to be given for the application of CV methods. The application of these methods, in combination with SfM-based orientation data, has not been the subject of research to date.

3. Concept

3.1. Objective

The main goal of this research is to improve the results of element detection from a point-cloud-based as-planned vs. as-built comparison by using additional information provided through the Structure-from-Motion process (images and camera positions), as well as the as-designed building information model (semantic data, geometric representation of elements, and position and dependencies of elements). The following concept presents the proposed solutions to tackle the mentioned challenges with several approaches, such as incorporating additional information on construction methods into the comparison algorithms.

3.2. Point of departure

The concept builds upon the body of knowledge of the research community as well as the previous research conducted by the authors. Thus, several steps in the process of automated progress monitoring are assumed to be given. Firstly, image acquisition for the generation of point clouds and camera position estimation is required. The authors provided several studies on image acquisition and proposed a UAV-based method as it is more flexible in comparison to fixed cameras [38]. Secondly, the point cloud and the

as-designed building information model must be aligned to one another (also known as registration). According to the well-documented state of the art, this is either performed via geodetic reference points that align the as-planned digital model with the point cloud on the measured geodetic position, via automated ICP methods (as mentioned earlier), or manually via point-to-point picking. The authors provide a detailed description of these approaches in [6] and [35]. In this paper, we significantly extend the state-of-the-art approach by using computer vision (CV) methods.

3.3. Concept overview

The concept presented in this paper relies on the exploitation of as-design building information models to improve the progress-detection process. We assume them to be available as IFC instance models. These models provide a geometric representation of all relevant building components, as well as the related semantic information (such as component type, material or the attribute "load-bearing") as well as 4D process data. The general idea is to enhance the purely geometric as-planned vs. as-built comparison from point-cloud to geometry level, with additional layers of information. Fig. 3 depicts the conceived processing chain. The highlighted process components provide new elements that are introduced in this paper. After defining the different sets of building elements required for the process in Sec. 3.4, these new elements are explained in detail in dedicated subsections.

The creation of the precedence relationship graph is discussed in Sec. 3.5. The following sections focus on schedule analysis (Sec. 3.6), and color detection (Sec. 3.7). The latter process step helps to identify whether an element is present or occluded by other structures. Finally, we introduce a method

that projects the 3D as-designed geometry provided by the building information model into the 2D plane, so as to apply image analysis techniques for element detection. Sec. 3.8 describes the projection process. Subsequently, Sec. 3.9 discusses the application of computer vision methods to detect the type of the element that is visible in the projected region of interest.

The combination of these individual processing steps results in a significant improvement in the accuracy of the overall automated progress detection method, as demonstrated through the case studies presented in Section 4.6.

3.4. Sets of elements and detection status

In the context of the research presented, the following sets of construction elements are defined in regard to as-built vs. as-planned comparison:

- E represents all elements of the current building
- $E_P(t)$ defines the amount of elements that should be present at the time t of observation according to the as-planned schedule
- $E_{GT}(t)$ defines the ground truth as all elements that are built at observation t
- $E_D(t)$ defines all elements that were detected during an observation at timestamp t
- $E_{ND}(t)$ defines all elements that were not detected during an observation at timestamp t
- $E_V(t)$ defines all elements that are visible from the corresponding points of view during observation at timestamp t

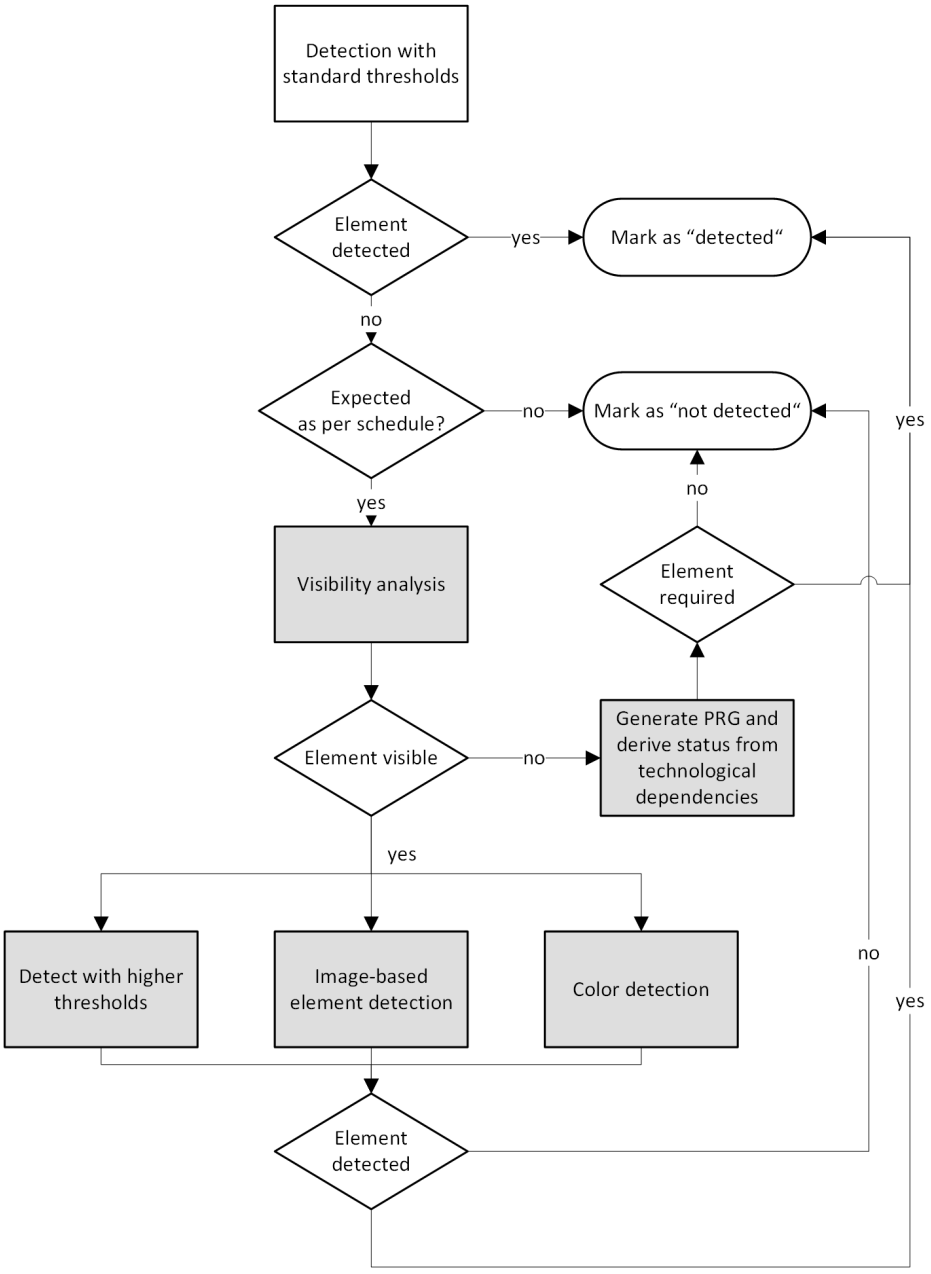


Figure 3: Concept for the enhancement of element detection. The highlighted process steps are introduced in this paper.

t defines the observation timestamp, at which the construction site has been monitored.

The following definitions hold true for all given sets at any timestamp t:

$$E = E_D(t) \cup E_{ND}(t) \tag{1}$$

$$E_D(t) \leqslant E_V(t) \leqslant E_{GT}(t) \leqslant E \tag{2}$$

According to these definitions, the set of $TruePositives$ is defined as

$$E_{TP}(t) = E_D(t) \cap E_{GT}(t) \tag{3}$$

while $FalsePositives$ are the counterpart:

$$E_{FP}(t) = E_D(t) \setminus E_{GT}(t) \tag{4}$$

The goal of this research is to verify as many existing construction elements as possible, so as to minimize the differences between these sets while keeping $E_{FP}(t)$ minimized. Mathematically speaking:

$$E_D(t) \longrightarrow E_{GT}(t) \tag{5}$$

It is not possible to define a relation between the planned elements $E_P(t)$ and the ground truth $E_{GT}(t)$ as the progress of the construction site depends on many external factors that cannot be formalized with the given data. The set of $E_P(t)$ can contain more elements than $E_{GT}(t)$ in the event of a delay on the construction site but also fewer elements in the event of faster construction times.

In addition to the mentioned sets, every construction element is classified individually for each of the following states: built (Ground Truth), detected, planned, encased in formwork, under construction.

These definitions are used in the described concepts.

3.5. Process sequencing and precedence relationships

As-built monitoring with SfM methods or laser scanning always captures one particular timestamp.

For automated handling of dependencies, a precedence relationship graph (PRG) is introduced [29]. The PRG formalizes technological dependencies between construction elements and is defined as a directed, acyclic graph (DAG) with each node representing one construction element [39]. Technological dependencies for load-bearing structures between two elements can be automatically detected when they have a particular spatial constellation that, in combination with the construction method applied, unambiguously defines their sequential order. For example, when conventional in-situ concreting methods are applied, a slab on top of a column can only be built after the column is finished. To generate this graph, the semantic as well as the geometric data from the digital model is used in combination with a knowledge base representing the construction methods. The geometric data is used to identify elements that are touching each other, and for sequencing them in their respective vertical order. Additionally, the semantic data is used to determine the construction method for an individual element, and to filter load-bearing elements. The generation of the initial precedence graph is performed as depicted in Algorithm 1. This method relies on a spatial query language, as introduced in Daum and Borrmann [40].

Algorithm 1 Pseudo code for the generation of an initial Precedence Relationship Graph

```
1: procedure GeneratePrecedenceRelationshipGraph
2:     E ← set of all construction elements
3:     for all E(LoadBearing) do
4:         for all ET do
5:             if Above(E(LoadBearing),ET) then
           AddDirectedEdge(E(LoadBearing),ET);
```

The initial precedence graph is completed manually in order to take project-specific boundary conditions and non-spatial precedence relationships into account.

The PRG is used to identify objects that are possibly under construction at the time of observation.

Using the introduced PRG, it is possible to identify elements that might be under construction and thus are considered for further investigation. The basic flowchart depicted in Figure 3 shows the implemented workflow for enhanced detection.

Based on the construction type and the erection method, different steps follow. As detailed above, walls and other vertically erected elements are considered for an extended threshold in order to identify possible formwork. Additionally, color matching helps to differentiate the material properties.

Moreover, the PRG allows for assumptions with regard to elements that are invisible due to occlusions, and thus not directly detectable. For example, load-bearing columns underneath a detected slab are expected to be built even if it is not possible to verify them via the point cloud.

3.6. Identified tasks during construction

Several tasks are required to construct in-situ concrete elements or similar elements. In concrete construction, formwork for in-situ concrete is the most common construction method. Several different methods are depicted in Figure 1 b) and d). All possible elements under construction are considered in order to detect formwork. In general, elements are counted as detected as soon as a certain amount of points per area $[Pts/m^2]$ with a distance of less than 2 cm are matched on the surface of the element [41]. If the expected elements are not detected, the threshold for the maximum distance can be adjusted to take into account the fact that the formwork with a thickness of around 0.20m might be currently in place. If this iteration brings positive results, the element can be marked as "under construction".

3.7. Color detection

In general, formwork for walls and columns consists of a wooden, smooth plate on the concrete side, and a steel structure for stability on the backside. This steel structure is often painted red, yellow or orange, and is distinct from the gray concrete. Formwork for slabs usually consists of elevated wooden plates that have the same color range as the steel structure mentioned. This color difference can be measured and may help to further improve the detection quality of formwork. The HSV (Hue-Saturation-Value) color space provided useful data for the color detection [42]. In contrast to the RBG color space, the HSV color space can describe color as perceived by humans but also saturation and brightness (value). Each value has a range from 0 to 1.

Comparing the color distribution of identified subsections of the point cloud can consequently help to achieve further verification of the existence of an element. The material color, as well as the type of construction, is retrieved from the building information model in order to gather color information. After identifying a gray color distribution for an expected concrete element, this data further confirms the existence of said element. In comparison, a mainly red or orange color distribution leads to the assumption that a formwork element is present, if the initial element has not been verified but is meant to be constructed with in-situ concrete.

3.8. Visibility analysis and projection of elements

Photogrammetry is based on estimating the position of all cameras that are used for the point cloud generation. Since the digital model of the construction sites is aligned to the point cloud during the comparison process, it is possible to project the 3D geometry of all elements into the respective 2D plane of a corresponding image [35]. Knowing the expected position of an element in image space enables highly accurate object-detection to be performed, using CV approaches.

More specifically, it is possible to perform a visibility detection by using the camera parameters to compute the projection of the model elements onto image space and of the process information, to define the set of construction elements that are supposed to be built. The building model coordinate system needs to be transformed into the camera coordinate system or vice versa in order to align both models. By applying this method, rendered images from all points of acquisition are generated that allow the determination of which elements are actually visible and can potentially be found in a gener-

ated point cloud. The resulting set of visible elements $E_V(t)$ enables greater detection accuracy.

The general approach for this method is explained in Braun and Bormann [35], though for a slightly different application scenario. For further clarification, the key steps are explained in this section.

In order to calculate the projection, the intrinsic camera matrix for the distorted images that projects 3D points in the camera coordinate frame to 2D pixel coordinates using the focal lengths (F_x, F_y) and the principal point (x_0, y_0) is required. Additionally, the skew coefficient s_k for the camera is required. It is zero if the image axes are perpendicular. The matrix K can be described as defined in equation 6.

$$K = \begin{bmatrix} F_x & s_k & x_0 \\ 0 & F_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

The translation of the camera is defined as:

$$T = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \tag{7}$$

Additionally, the rotation matrix for each image, as defined in equation 8 is needed.

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{8}$$

Both, translation and rotation can be described in one 3 x 4 matrix:

$$RT = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_1 \\ r_{21} & r_{22} & r_{23} & T_2 \\ r_{31} & r_{32} & r_{33} & T_3 \end{bmatrix} \tag{9}$$

Using the model coordinates of all triangulated construction elements, it is possible to calculate the projection of each element into the camera coordinate system and therefore overlay the model projection and the corresponding image taken from the point of observation with equation 10.

$$t = K * RT * p; \tag{10}$$

The resulting 2D coordinates that are rendered into the image are calculated by using the vector t and calculating the x and y coordinates by

$$x = t[0]/t[2] \tag{11}$$

and

$$y = t[1]/t[2] \tag{12}$$

With this projection, the model can be rendered from the camera's perspective for all images acquired during observation. After including the 4D temporal information from the as-planned model, this information can be fused, to render the model with the expected set of elements $E_P(t)$ from all estimated camera positions. The term "rendering" here refers to the creation of the 2D projection of the model according to the rendering pipeline established in computer graphics [43], but without applying advanced features such as reflections, light sources or shading. These rendered images are

analyzed for all visible elements $E_V(t)$ by applying the Painter's algorithm [44]. With knowledge of this set of elements, the set $E_D(t)$ can be checked for false positives, but also measured for accuracy regarding its true positive rate. This is done by excluding elements from set $E_P(t)$ or $E_GT(t)$ that are invisible from any camera position during acquisition.

3.9. Image-based object detection

To further enhance the detection of construction elements, we propose making use of the images taken in the course of the initial acquisition for the photogrammetric point cloud generation. By applying the previously described projection technique, all construction elements can be localized on any image taken during the acquisition. A sample is shown in Fig. 4: A column of interest is selected in the 3D view (marked red); detailed information about this element is shown in the lower right. Accordingly, a corresponding image that validated the existence of the selected element - and additionally the 3D to 2D projection described in Section 3.8 - is used to display the expected position of the element in this image.

As machine-learning methods have made significant advancements in recent years, tasks like image classification or even region detection on images are mow being used in various scenarios. For the task of progress monitoring, the authors propose the use of a Convolutional Neural Network (CNN) trained on construction elements and thus able to detect the type and instances of construction elements on the given images. In the case that an element is not detected and validated by the point cloud, the implemented workflow is followed as described in Fig. 5.

If the element is expected according to the up-to-date schedule and re-

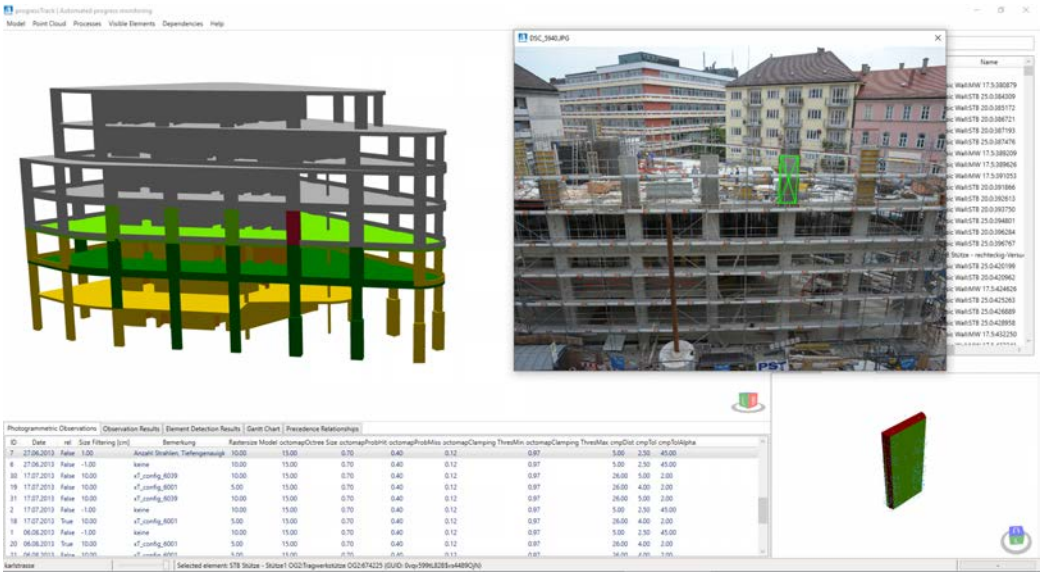


Figure 4: Projection of a selected 3D geometry into the 2D plane of a corresponding image

quires in-situ work, in a first step, the thresholds are increased as defined in Sec. 3.6. If this helps to validate the elements' existence, it is added to the set of detected elements $E_D(t)$. If not, the 2D projection, as mentioned in Sec. 3.8, is used to identify the region of interest in a suitable image.

Subsequently, the trained CNN [45] classifies the region according to the predefined states and thus contributes to a refined state detection. If, e.g., formwork is detected here, the element can be marked as "under construction".

In order to use a CNN for object-based region detection, the training of said network is required. For this purpose, 5,000 images were labeled with the categories *formwork*, *scaffolding*, *columns*, and *walls*. This resulted in 9,700 labeled formwork elements. The labeling procedure is depicted in Fig. 6. The data is converted into the COCO data format [46] and prepared for

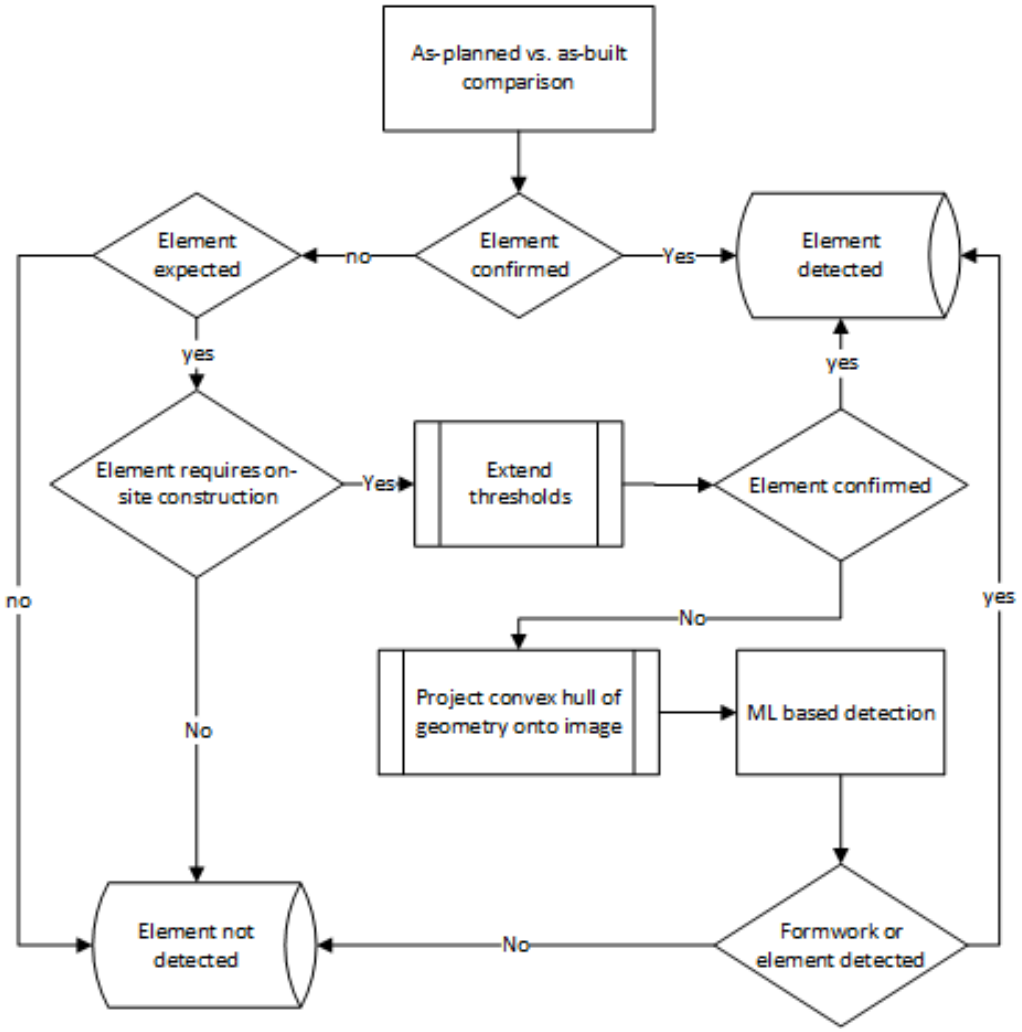


Figure 5: Occluded construction elements in generated point cloud caused by scaffolding, formworks, existing elements and missing information during the reconstruction process

462 training by augmenting the images to enlarge the training set even further.

Figure 6: Sample image of the labeling process. Displayed are the labeled formwork (blue) and column (green) elements. During this research, Labelbox [47] is used for labeling.

463 To sum up, all introduced methods make the overall process much more
464 robust compared to a purely geometry-based approach, and lead to a higher
465 detection accuracy.

466 4. Case Study

467 Several construction sites were monitored with different observation meth-
468 ods to validate the introduced concepts. The construction sites are all
469 German-based and cover a number of structural engineering buildings as well
470 as infrastructure (one bridge, one wastewater treatment plant). The main

construction method is in-situ concreting, this being the most common construction technique in Germany. Listed in Table 1 are the three construction sites that are used as case studies in this section.

Site	**Elements**	**Observations**	**Pictures taken**	**Duration**
Test Site A	671	6	1,805	5 months
Test Site B	943	9	2,350	10 months
Test Site C	2,229	23	3,144	5 months

Table 1: Test sites monitored during this case study

In this context, the authors published several papers presenting their approach and developed a software framework, which was introduced in Braun et al. [29] and shown in Figure 7. To visualize the comparison results and the detected elements, and to verify the used algorithms, all gathered data is stored in a database that is accessible via this software. The tool displays all geometric and semantic building element information as well as scheduling data that has been parsed from IFC instance models. The detected elements are highlighted for easy identification. Figure 7 shows the software interface with the example of one of the construction-site case studies used in this research. The building mainly consists of in-situ concrete elements that were cast using formwork on site. In the figure, one individual capturing event is selected, and all detected elements are highlighted. Green coloring represents elements that have been built and are correctly detected and confirmed through the point cloud. All yellow elements are built but were not confirmed through the point cloud.

There are several reasons why some of those elements may not be de-

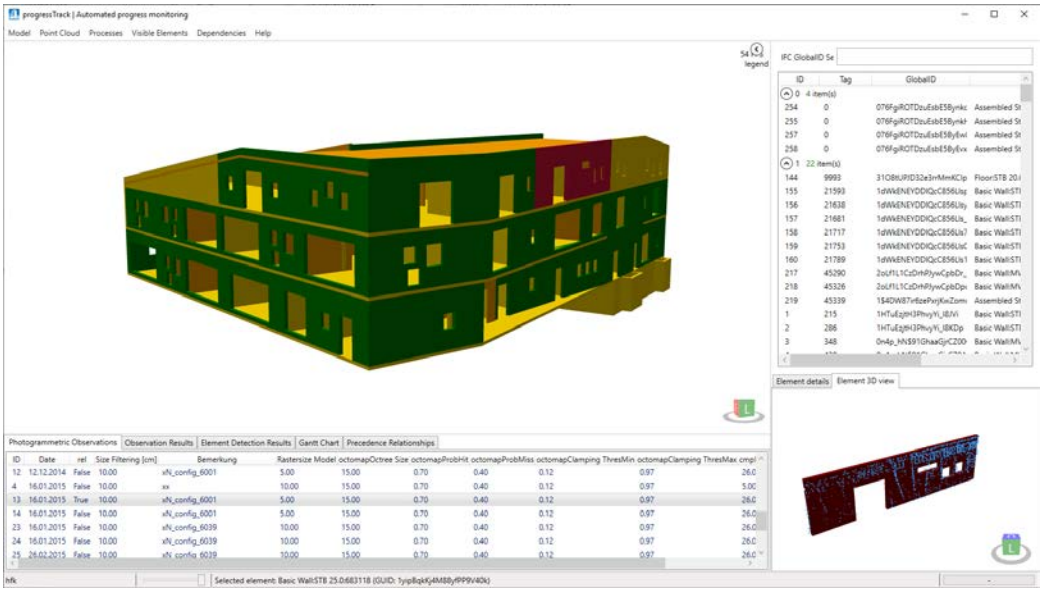


Figure 7: Screenshot of a developed tool for as-planned vs. as-built comparison. A specific observation is selected to visualize the detected construction elements at that time. Details of selected elements are shown in a separate viewer.

tected. The most prominent reason is the occlusions that occur on site. During construction, large amounts of temporary structures like scaffolds, construction tools, and construction machinery obstruct the view of the element surfaces. Limited acquisition positions further reduce the visible surfaces and hence the overall quality of the generated point clouds. Additionally, elements inside of the building are also occluded by other building elements for acquisitions outside of the building.

Another reason for weak detection rates is building elements that are currently under construction. As those elements count towards the overall progress, they must not be missed, and play a crucial role in defining the exact state in the current process. In general challenges exist for all construction methods, whose geometry under construction differs largely from the final element geometry which requires the use of temporary construction objects. This applies, e.g., to reinforced concrete and multi-layered walls. On the one hand, formwork which is used for concrete pouring, may obstruct the view of the element, making it impossible to be detected. On the other hand, the plane surface of formwork for a slab might be detected as the surface of the slab itself and thus would lead to a false positive. Due to these challenges, further enhancements to the comparison and detection algorithms are needed. Since the digital model contains information on construction methods, the authors propose using this knowledge in the overall detection process. By deducing the precedence relationships with a query language, assumptions regarding occluded elements can be made. Construction methods and derivation of expected elements lead to new as-planned vs. as-built comparison capabilities, such as extended thresholds and com-

puter vision methods to detect objects like formwork on the raw observation images, taken for the point cloud generation.

4.1. Precedence Relationship Graph

The PRG for all construction sites is generated by using a query language for Building Information Models (QL4BIM, Daum and Borrmann [40]). With the algorithm introduced in Sec. 3.5, any building information model that has sufficient semantic information can be analyzed, and technological dependencies are formalized by the introduced graph. Fig. 8 shows the PRG for one of the mentioned case studies. Each node represents one construction element; the directed edges show the corresponding dependency.

Based on the detected elements (marked in green and yellow), all dependent elements can be identified via this graph. Specifically, this graphs allows one to make assumptions regarding the construction elements that were either invisible during observation, or were not detected due to occlusions or other issues (as mentioned before). The elements marked in blue in Fig. 8 are identified as depending elements with this method.

Table 2 shows detailed enhancements for the introduced PRG. In particular, a significant amount of construction elements were identified as depending upon the detected elements. In this respect, these elements are logically required to be built despite the fact that they were not confirmed visually by the point cloud.

This information helps to obtain additional information for the as-planned vs. as-built comparison: if a slab is built, all load-bearing elements underneath it must have been built, even though they cannot be verified by any visual method.

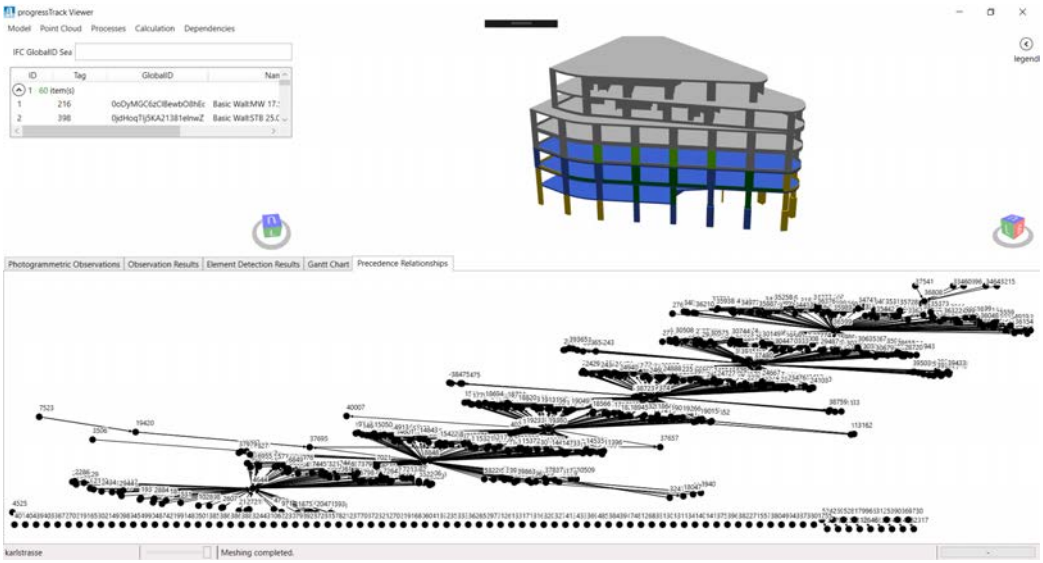


Figure 8: Generated precedence relationship graph for Test Site A. Elements marked blue were derived from the PRG in combination with the detected elements marked in green and yellow.

Date	$E_{GT}(t)$	$E_D(t)$	$\delta E_{PRG}(t)$
15.05.	89	37	20
12.06.	152	32	57
27.06.	184	59	54
17.07.	233	53	85
06.08.	277	95	102
04.09.	342	98	159

Table 2: Enhancing results by applying the introduced PRG for Case Study site A

4.2. Varying dimensions

Figure 9 depicts a part of a snippet of a point cloud, generated at one individual time-step during observation. It is overlaid with the corresponding 3D geometry and visualized in green, to symbolize the as-planned as well as the as-built status. Based on this example, the general workflow for elements under construction is shown. As depicted, the front wall is already finished, and the concrete surface is visible. The wall in the second row is currently under construction, and the formwork is present and registered in the point cloud.

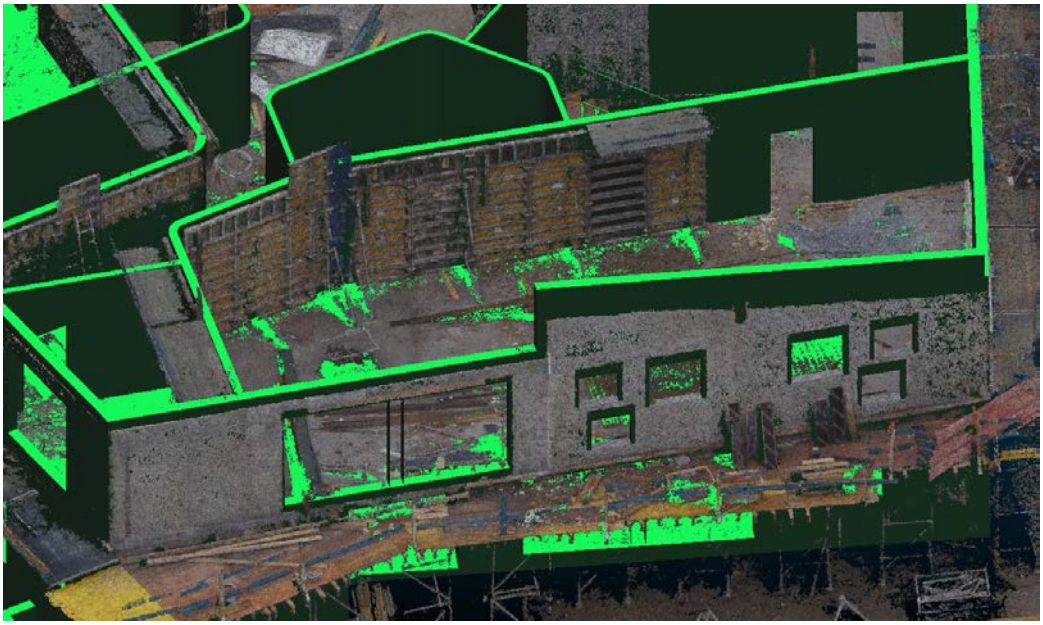


Figure 9: Point cloud of a finished, plain wall and formwork overlaid with the corresponding 3D geometry on Test Site B

During detection, it is expected that the first row of walls will be detected. Due to the threshold of max. 1 cm, the second row should not be detected due

to the formwork. Figure 10 a) shows the expected result, with an additional set threshold of 1000 $points/m^2$ (in green). Triangles marked in yellow have matching points but do not qualify for the set thresholds, while elements marked red have no qualifying points at all. The walls in the second row are expected to be in progress. As presented in the concept in Section 3, the detection is therefore carried out with a larger threshold. Based on this result, the accepted point-to-surface distance is increased to 10 cm, which leads to the results depicted in Figure 10 b).

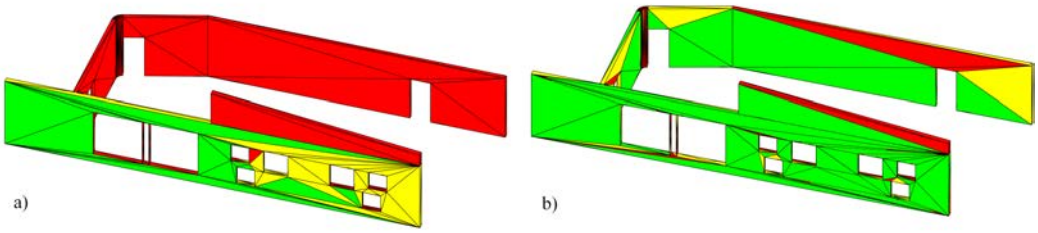
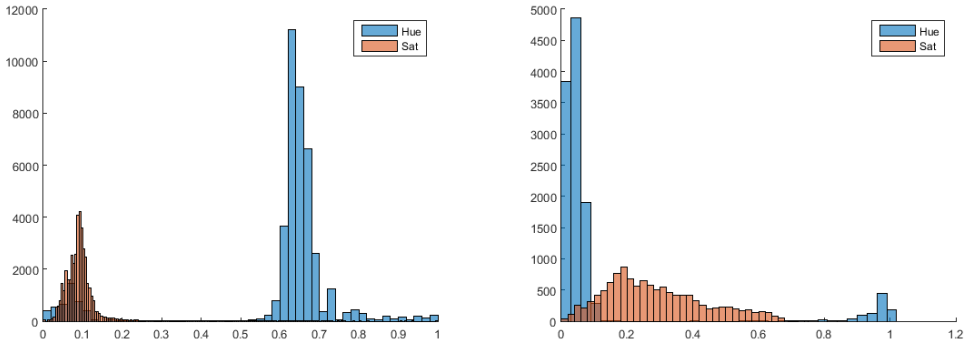


Figure 10: Triangles detected during the time-step shown in Figure 9. a) with $1cm$ gaps and $\rho > 1000pts/m^2$, b) with $10cm$ gaps and $\rho > 1000pts/m^2$

The increased threshold leads to the expected higher point density on the wall under construction, as the formwork is considered, too. According to the introduced workflow, the wall is now marked as "under construction", leading to a further detailed automated progress monitoring.

4.3. Color detection for formwork and reinforcement

As detailed in Section 3, taking colors into account can improve the detection of formwork or reinforcements due to their significantly varying colors, in comparison to the grey colors of the concrete. The color values of the different elements were compared to prove this statement. Figure 11 shows



(a) HSV color distribution for concrete (b) HSV color distribution for formwork

Figure 11: Distribution of frequency in the HSV color space shows clear deviations between concrete and formwork elements with the Hue value represented by blue bars and Saturation value represented by orange bars.

the calculated mean values for different elements under different lighting conditions.

In calculating the mean HSV values, all points relevant to an element are considered, along with the relevant color information. The results show that the brightness (value) varies largely, which is due to the lighting conditions itself. Therefore, this value has no further significance for this study. However, the hue values for formwork fall into the correct range for warm, red colors, whereas the concrete walls are based on "colder" colors. Additionally, the saturation differs by at least a factor 2.3. This color distribution analysis at a point-cloud level allows automated color interpretation to be carried out, and helps to identify differences between expected and actual color ranges based on material properties. The described process is used during the whole comparison to obtain a higher accuracy of information.

4.4. Visible elements

The visibility analysis is tested on several construction sites. Fig. 12 shows four samples from different observation times and construction sites. Each element has a unique color for identification purposes.

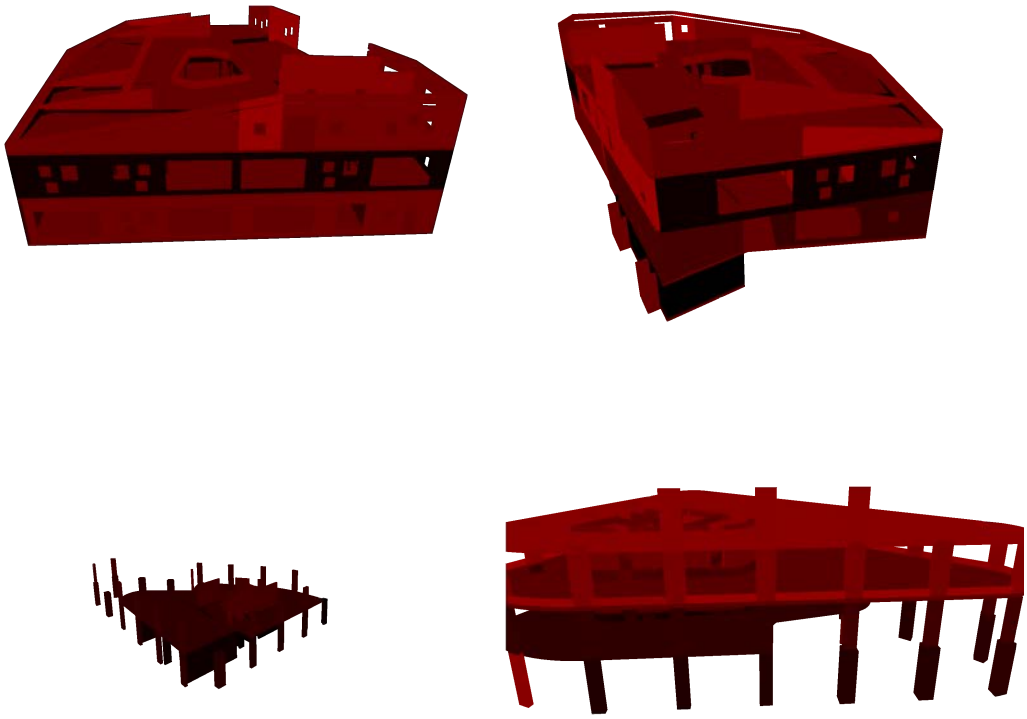


Figure 12: Visibility analysis with rendered geometry of set $E_P(t)$ for several construction sites and observations. All elements are rendered in different colors to distinguish them from each other.

Based on these results, all visible elements are identified and added to a corresponding set $E_V(t)$. This additional step does not detect any additional elements during the as-planned vs. as-built process, however it helps to set the detection results in a more accurate context. In detail, false positives

can be reduced by removing invisible elements. Additionally, the thresholds used for the comparison process can be validated in a more precise manner, as the invisible elements are not added to the set of not detected elements.

Table 3 shows this data for one of our case studies during the whole observation period.

Date	$E_{GT}(t)$	$E_D(t)$	$E_V(t)$	$\%_{Vis}$
15.05.	89	37	73	82.0 %
12.06.	152	32	122	80.3 %
27.06.	184	59	155	84.2 %
17.07.	233	53	214	91.8 %
06.08.	277	95	275	99.3 %
04.09.	342	98	325	95.0 %

Table 3: Visible elements based on the introduced algorithm for Test Site A.

4.5. Image-based object detection

For the image-based object detection described in Sec. 3.9 we trained a Mask R-CNN-based [45] neural network using a training set consisting of over 5,000 images from five different construction sites and 40 observations with 9,700 labeled formwork elements and around 5,000 labeled column elements. Depicted in Figure 14, the results for formwork and column elements are shown in an image that was not part of the training set. A common method to quantify the estimated result is the mean average precision that calculates as

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives} \tag{13}$$

In combination with the recall

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives} \tag{14}$$

the harmonized F_1 score can be calculated as:

$$F_1 score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{15}$$

An ideal network with perfect precision and recall values would achieve a F_1 score of 1. The trained network has a mean average precision (mAP) of 90.7% with an IoU (Intersection over Union) of 0.5 over all categories. With $TP = 11731$, $FP = 1099$ and $FN = 928$, the precision is at 0.914, the recall at 0.927, resulting in an $F_1 Score = 0.920$ proving the suitability of the implemented methods. Fig. 13 shows the corresponding precision-recall curve for the trained network.

It has been tested against previously unknown images from the internet and other construction sites.

The results of this image-based region detection are subsequently used for the as-planned vs. as-built comparison. As introduced in Figure 5, construction elements that have not been verified by the point cloud, are run through an additional workflow, in order to check for formwork elements. If the CNN verifies the existence of a formwork element, the corresponding concrete structure is labeled as "under construction", making the process estimation more accurate. After testing this approach on a real-world construction site, this additional step proved to be suitable for construction sites

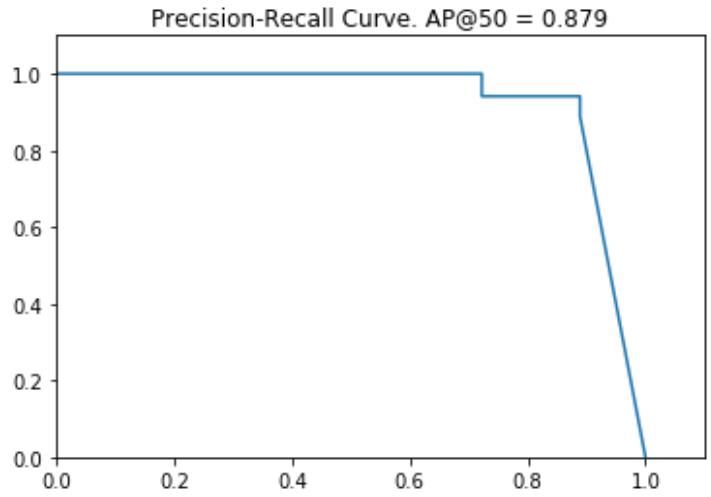


Figure 13: Precision-recall Curve for the trained network

Figure 14: Formwork and column elements detected by a trained CNN using Mask R-CNN on Test Site C

that use in-situ concreting as a manufacturing method. Table 4 shows the amount of detected formwork elements with the help of the trained network.

Date	15.05.	12.06.	27.06.	17.07.	06.08.	04.09.
Detected	9	11	8	2	0	8

Table 4: Detected formwork elements during the observations for Test Site A.

4.6. Results

After the evaluation of all steps, the methods are incorporated into the presented software framework. Table 5 shows the results for one of our case studies during the complete construction process. During the initial, point-cloud-based comparison, the following data was gathered:

Date	$E_P(t)$	$E_{GT}(t)$	$E_D(t)$	$E_{FP}(t)$	$A_D(t)[m^2]$	$A_{GT}(t)[m^2]$	$\%_A$
15.05.	60	89	37	2	1,162.76	1916.95	60.66 %
12.06.	133	152	32	11	1,326.95	3,557.74	37.3 %
27.06.	240	184	59	0	2,244.51	4,808.6	46.68 %
17.07.	348	233	53	5	4,147.65	6,261.07	66.25 %
06.08.	456	277	95	1	4,480.78	6,773.9	66.15 %
04.09.	569	342	98	1	4,763.63	9,197.7	51.79 %

Table 5: Resulting element sets for Test Site A

According to this data, the detection rates differ over a range of 37 % to 66 % correctly detected elements, based on the area surfaces. As mentioned above, these results largely depend on the point-cloud density and recon-

struction quality from the SfM process. For any construction planner, these results would be insufficient as a comprehensive progress-monitoring tool.

After applying the newly introduced methods to this initial as-planned vs. as-built comparison, these additional results were gathered as shown in Table 6 with detected, cast elements defined as $E_{FW}(t)$ and elements inferred by the PRG, in addition to the previously detected elements, as $\delta E_{PRG}(t)$.

This table summarizes the results of the previous sections.

Date	$E_V(t)$	$E_{FW}(t)$	$\delta E_{PRG}(t)$	$E_{D_{new}}(t)$	$A_D(t)[m^2]$	$A_V(t)[m^2]$	$\%_A$
15.05.	73	9	20	66	1509.42	1681.04	83.8 %
12.06.	122	11	57	100	2792.95	3284.76	85.0 %
27.06.	155	8	54	121	3975.51	4579.60	86.8 %
17.07.	214	2	85	140	4975.65	6059.13	82.1 %
06.08.	275	0	102	197	5780.78	6644.94	87.0 %
04.09.	325	8	159	265	7675.58	9021.86	85.1 %

Table 6: Enhanced results for the detection with the newly introduced methods for Test Site A.

As shown, the number of detected true positives is raised significantly by applying the introduced steps. The newly detected rates all lie in the range between 80% to 90% of the actually built elements. An improvement of more than 100% in detected elements in comparison to the pure point-cloud vs. geometry-based detection methods was achieved. To draw conclusions from the results, there is still potential for further improvements. However, the introduced methods were tested on real-world construction sites over the complete construction cycle, and not only on a limited test area which usually

constitutes a more controlled environment. Real-world data from construction sites always introduces many occlusions, and non-modeled elements that make it nearly impossible to detect all elements on a construction site.

5. Discussion and Outlook

5.1. Conclusion

Detailed progress monitoring is of utmost importance for efficient construction site management as it allows delays to be identified early, and for respective counter-measures to be taken. Matching the as-designed 4D building information model to point clouds provides a suitable basis for automating this process. The general approach of Scan-vs-BIM has been proposed and investigated by a number of researchers in recent years. In this paper, a number of methods are introduced that further improve the accuracy of the detection process of the as-planned vs. as-built comparisons. The common approach lies in fusing information generated by different techniques and from different sources, namely the images, the point cloud and the building information model. The formal description of the technological dependencies in the construction process in the form of a precedence relationship graph allows the inference of status information on components that are not directly detectable. Image-based color detection and a higher threshold for elements with possible formwork in place enable the correct identification of elements that are under construction at the time of capturing the site. As a core contribution, the paper presents how CNN-based object-detection methods are applied to the captured images to correctly detect elements that tend to be otherwise falsely classified. Significant synergies are created by training

the network with images that are automatically labeled, by applying Scan-vs-BIM techniques. The use of image-based object detection extends the reliability of the status-detection process significantly, due to the larger density of pixel-based information, in comparison with a pure point-cloud-based approach.

5.2. Limitations

It is crucial to note that the image data can only be used thanks to the photogrammetric process and the underlying camera pose estimation. Laser scanners usually do not provide this data and are therefore not suitable for this approach. Another limitation is the requirement for a well-aligned BIM. In our approach, this is achieved by markers on site. However, a minor manual step is required in order to find the exact orientation and scaling. Only after combining this data with the aligned building information model is it possible to gather additional information from the images in relation to the building model.

The described ML approach is limited to the provided training data. This data currently only includes construction sites in Germany, which might make the network biased and unsuitable for different regions that use different construction methods. The observed construction sites so far mainly used in-situ concreting and a small number of prefabricated elements.

5.3. Outlook

All introduced methods enhance the automated construction progress monitoring workflow. However, it is still the case that not all elements can

be detected. Better acquisition methods will play an essential role in solving these issues. Several research groups have proposed different acquisition methods to detect indoor elements, too. A combination of all these methods could help to improve element detection even further.

More comprehensive data sets for image-based ML are required to cover different construction methods and materials from other regions.

Acknowledgments

We would like to thank the Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences and Humanities (BAdW) for the support and provision of computing infrastructure essential to this publication. We would like to thank the German Research Foundation (DFG) for funding the initial project of progress monitoring.

Additionally, we would like to thank our students Bernhard Mueller and Sebastian Behnke for their support during the labeling and machine-learning training phase. The support of all construction companies, providing data for the research conducted in this field is much appreciated: Adidas, Leitner GmbH & Co Bauunternehmung KG, Kuehn Malvezzi Architects, Staatliches Bauamt Muenchen, Baugesellschaft Brunner+Co, BKL Baukranlogistik GmbH, Geiger Gruppe, Baureferat H5, Landeshauptstadt Muenchen, Baugesellschaft Mickan mbH & Co KG, h4a Architekten, Wenzel + Wenzel, Zueblin, and Stadtvermessungsamt Muenchen.

References

[1] B. Hardin, D. McCool, BIM and construction management : proven tools, methods, and workflows, Wiley, ISBN: . 9781118942765.

[2] F. Bosché, C. Haas, Automated retrieval of 3D CAD model objects in construction range images, Automation in Construction 17 (2008) 499–512. doi:10.1016/j.autcon.2007.09.001.

[3] M. Golparvar-fard, F. Pena-Mora, S. Savarese, D4AR - a 4 dimensional augmented reality model for automation construction progress monitoring data collection, processing and communication, Journal of Information Technology in Construction 14 (2009) 129–153. doi:https://www.itcon.org/paper/2009/13.

[4] J. J. Lin, K. K. Han, M. Golparvar-Fard, A Framework for Model-Driven Acquisition and Analytics of Visual Data Using UAVs for Automated Construction Progress Monitoring, in: Computing in Civil Engineering 2015, American Society of Civil Engineers, Reston, VA, 2015, pp. 156–164. URL: http://ascelibrary.org/doi/10.1061/9780784479247.020. doi:10.1061/9780784479247.020.

[5] C. Wu, Towards linear-time incremental structure from motion, in: Proceedings - 2013 International Conference on 3D Vision, 3DV 2013, IEEE, 2013, pp. 127–134. URL: http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=6599068. doi:10.1109/3DV.2013.25.

[6] A. Braun, S. Tuttas, U. Stilla, A. Borrmann, Incorporating knowledge on construction methods into automated progress monitoring techniques, in: 23rd International Workshop of the European Group for Intelligent Computing in Engineering, EG-ICE 2016, 2016, pp. 1–11. URL: https://publications.cms.bgu.tum.de/2016BraunEGICE.pdf.

[7] W. Huhnt, S. Richter, F. Enge, Modification Management for Construction Processes, Tsinghua Science and Technology 13 (2008) 185–191. doi:10.1016/S1007-0214(08)70147-0.

[8] J. Tulke, Kollaborative Terminplanung auf Basis von Bauwerksinformationsmodellen, Ph.D. thesis, Bauhaus Universität Weimar, 2010. URL: https://d-nb.info/1116284456/34.

[9] C. Kropp, C. Koch, M. König, Drywall State Detection in Image Data for Automatic Indoor Progress Monitoring, in: Computing in Civil and Building Engineering (2014), American Society of Civil Engineers, Reston, VA, 2014, pp. 347–354. URL: http://ascelibrary.org/doi/10.1061/9780784413616.044. doi:10.1061/9780784413616.044.

[10] T. Omar, M. L. Nehdi, Data acquisition technologies for construction progress tracking, Automation in Construction 70 (2016) 143–155. doi:10.1016/j.autcon.2016.06.016.

[11] Z. Pučko, N. Šuman, D. Rebolj, Automated continuous construction progress monitoring using multiple workplace real time

760 3D scans, Advanced Engineering Informatics 38 (2018) 27–40.
761 doi:10.1016/j.aei.2018.06.001.

762 [12] F. Bosché, Plane-based registration of construction laser scans with
763 3D/4D building models, Advanced Engineering Informatics 26 (2012)
764 90–102. doi:10.1016/j.aei.2011.08.009.

765 [13] F. Bosché, Automated recognition of 3D CAD model objects in laser
766 scans and calculation of as-built dimensions for dimensional compliance
767 control in construction, Advanced Engineering Informatics 24 (2010)
768 107–118. doi:10.1016/j.aei.2009.08.006.

769 [14] Y. Turkan, F. Bosché, C. T. Haas, R. Haas, Automated progress track-
770 ing using 4D schedule and 3D sensing technologies, Automation in Con-
771 struction 22 (2012) 414–421. doi:10.1016/j.autcon.2011.10.003.

772 [15] C. Kim, B. Kim, H. Kim, 4D CAD model updating using image
773 processing-based construction progress monitoring, Automation in Con-
774 struction 35 (2013) 44–52. doi:10.1016/j.autcon.2013.03.005.

775 [16] M. Golparvar-Fard, F. Peña-Mora, S. Savarese, Automated Progress
776 Monitoring Using Unordered Daily Construction Photographs and
777 IFC-Based Building Information Models, Journal of Computing in
778 Civil Engineering 29 (2015) 04014025. doi:10.1061/(ASCE)CP.1943-
779 5487.0000205.

780 [17] S. Tuttas, A. Braun, A. Borrmann, U. Stilla, Validation of Bim Com-
781 ponents By Photogrammetric Point Clouds for Construction Site Moni-
782 toring, ISPRS Annals of Photogrammetry, Remote Sensing and Spatial

Information Sciences II-3/W4 (2015) 231–237. doi:10.5194/isprsannals-II-3-W4-231-2015.

[18] C. C. Kim, H. Son, C. C. Kim, Fully automated registration of 3D data to a 3D CAD model for project progress monitoring, Automation in Construction 35 (2013) 587–594. doi:10.1016/j.autcon.2013.01.005.

[19] C. Zhang, D. Arditi, Automated progress control using laser scanning technology, Automation in Construction 36 (2013) 108–116. doi:10.1016/j.autcon.2013.08.012.

[20] H. Mahami, F. Nasirzadeh, A. Hosseininaveh Ahmadabadian, S. Nahavandi, Automated Progress Controlling and Monitoring Using Daily Site Images and Building Information Modelling, Buildings 9 (2019) 70. doi:10.3390/buildings9030070.

[21] Y. Ibrahim, T. Lukins, X. Zhang, E. Trucco, a. Kaka, Towards automated progress assessment of workpackage components in construction projects using computer vision, Advanced Engineering Informatics 23 (2009) 93–103. doi:10.1016/j.aei.2008.07.002.

[22] K. Asadi, H. Ramshankar, M. Noghabaei, K. Han, Real-Time Image Localization and Registration with BIM Using Perspective Alignment for Indoor Monitoring of Construction, Journal of Computing in Civil Engineering 33 (2019) 04019031. doi:10.1061/(ASCE)CP.1943-5487.0000847.

[23] C. Kropp, C. Koch, M. König, Interior construction state recognition with 4D BIM registered image sequences, Automation in Construction 86 (2018) 11–32. doi:10.1016/j.autcon.2017.10.027.

[24] Y. Turkan, F. Bosché, C. T. Haas, R. Haas, Tracking of secondary and temporary objects in structural concrete work, Construction Innovation 14 (2014) 145–167. doi:10.1108/CI-12-2012-0063.

[25] K. K. Han, M. Golparvar-Fard, Appearance-based material classification for monitoring of operation-level construction progress using 4D BIM and site photologs, Automation in Construction 53 (2015) 44–57. doi:10.1016/j.autcon.2015.02.007.

[26] K. K. Han, D. Cline, M. Golparvar-Fard, Formalized knowledge of construction sequencing for visual monitoring of work-in-progress via incomplete point clouds and low-LoD 4D BIMs, Advanced Engineering Informatics 29 (2015) 889–901. doi:10.1016/j.aei.2015.10.006.

[27] I. C. Wu, A. Borrmann, U. Beißert, M. König, E. Rank, Bridge construction schedule generation with pattern-based construction methods and constraint-based simulation, Advanced Engineering Informatics 24 (2010) 379–388. doi:10.1016/j.aei.2010.07.002.

[28] K. Szczesny, M. Hamm, M. König, Adjusted recombination operator for simulation-based construction schedule optimization, in: Proceedings of the 2012 Winter Simulation Conference, 2012, pp. 1219–1229. URL: https://dl.acm.org/citation.cfm?id=2429835.

[29] A. Braun, S. Tuttas, A. Borrmann, U. Stilla, Automated Progress Monitoring Based on Photogrammetric Point Clouds and Precedence Relationship Graphs, in: Proceedings of the 32nd International Symposium

on Automation and Robotics in Construction and Mining (ISARC 2015), 2017, pp. 274–280. doi:10.22260/isarc2015/0034.

[30] H. Hamledari, B. McCabe, S. Davari, A. Shahi, Automated Schedule and Progress Updating of IFC-Based 4D BIMs, Journal of Computing in Civil Engineering 31 (2017) 04017012. doi:10.1061/(ASCE)CP.1943-5487.0000660.

[31] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444. doi:10.1038/nature14539.

[32] B. Akinci, F. Boukamp, C. Gordon, D. Huber, C. Lyons, K. Park, A formalism for utilization of sensor systems and integrated project models for active construction quality control, Automation in Construction 15 (2006) 124–138. doi:10.1016/j.autcon.2005.01.008.

[33] H. Nhat-Duc, Q. L. Nguyen, V. D. Tran, Automatic recognition of asphalt pavement cracks using metaheuristic optimized edge detection algorithms and convolution neural network, Automation in Construction 94 (2018) 203–213. doi:10.1016/j.autcon.2018.07.008.

[34] K. Han, M. Golparvar-Fard, Crowdsourcing BIM-guided collection of construction material library from site photologs, Visualization in Engineering 5 (2017) 14. doi:10.1186/s40327-017-0052-3.

[35] A. Braun, A. Borrmann, Combining inverse photogrammetry and BIM for automated labeling of construction site images for machine learning, Automation in Construction 106 (2019) 102879. doi:10.1016/j.autcon.2019.102879.

[36] S. Chi, C. H. Caldas, Automated Object Identification Using Optical Video Cameras on Construction Sites, Computer-Aided Civil and Infrastructure Engineering 26 (2011) 368–380. doi:10.1111/j.1467-8667.2010.00690.x.

[37] H. Hamledari, B. McCabe, S. Davari, Automated computer vision-based detection of components of under-construction indoor partitions, Automation in Construction 74 (2017) 78–94. doi:10.1016/j.autcon.2016.11.009.

[38] S. Tuttas, A. Braun, A. Borrmann, U. Stilla, Acquisition and Consecutive Registration of Photogrammetric Point Clouds for Construction Progress Monitoring Using a 4D BIM, Photogrammetrie, Fernerkundung, Geoinformation 85 (2017). doi:10.1007/s41064-016-0002-z.

[39] G. W. Zobrist, J. V. Leonard, Progress in simulation, Intellect Books, 1992. ISBN: , 978-0893916527.

[40] S. Daum, A. Borrmann, Processing of Topological BIM Queries using Boundary Representation Based Methods, Advanced Engineering Informatics 28 (2014) 272–286. doi:10.1016/j.aei.2014.06.001.

[41] S. Tuttas, A. Braun, A. Borrmann, U. Stilla, Evaluation of acquisition strategies for image-based construction site monitoring, in: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, volume 41, ISPRS Congress, Prague, Czech Republic, 2016, pp. 733–740. doi:10.5194/isprsarchives-XLI-B5-733-2016.

[42] S. Sural, Gang Qian, S. Pramanik, Segmentation and histogram generation using the HSV color space for image retrieval, in: Proceedings. International Conference on Image Processing, volume 2, IEEE, 2003, pp. II–589–II–592. doi:10.1109/icip.2002.1040019.

[43] J. D. Foley, J. D. Foley, Computer graphics : principles and practice, Addison-Wesley, ISBN: . 0201121107.

[44] T. T. Elvins, A survey of algorithms for volume visualization, ACM SIGGRAPH Computer Graphics 26 (2005) 194–201. doi:10.1145/142413.142427.

[45] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2016, pp. 770–778. URL: http://ieeexplore.ieee.org/document/7780459/. doi:10.1109/CVPR.2016.90.

[46] M. Andriluka, L. Pishchulin, P. Gehler, B. Schiele, 2D Human Pose Estimation: New Benchmark and State of the Art Analysis, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2014, pp. 3686–3693. URL: http://arxiv.org/abs/1405.0312 http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6909866. doi:10.1109/CVPR.2014.471.

[47] Labelbox, training data workbench, 2020. URL: www.labelbox.com, accessed: 2020-01-15.