Scalable construction monitoring for an as-performed progress documentation across time

Fiona C. Collins[1*] , Fabian Pfitzner[1*] and Jonas Schlenger[1*]
[1]Chair of Computational Modeling and Simulation, TU Munich, Arcisstr. 21, 80333 Munich, Germany
E-mail(s): fiona.collins@tum.de, fabian.pfitzner@tum.de, jonas.schlenger@tum.de
* all authors have contributed equally

Abstract: Compared to other industry sectors, the construction sector's productivity is relatively low. By collecting data directly from the construction site, the main bottlenecks can be identified and support decision-makers in making well-informed decisions. This is by no means a simple task because of the complexity and unpredictability of construction sites and the challenges of monitoring heterogeneous on-site data. While other researchers focus on elaborating specific steps or use-cases of the construction monitoring processes, we present a holistic workflow for scalable shell construction monitoring that uses state-of-the-art data processing techniques. The result of the proposed workflow is a per-instance database for on-site progress across time. Such a database has many possibilities for application. This can range from giving an overview to construction managers, providing a backbone for sophisticated analysis and digital twins, to the validation of computer vision approaches.

Keywords: Construction Monitoring, Image Processing, Photogrammetric Point Clouds

1 Introduction

The planning phase preceding the construction of buildings has already taken considerable steps to implement the BIM methodology fully. Conversely, the digitization of the as-performed construction process still significantly lacks behind. The absence of a common and regularly synchronized data environment fundamentally goes in hand with the risk for undocumented deviations from the project intent, resulting in extensive delays and increasing costs [1].

A construction site is a highly dynamic, complex, and chaotic environment where many parties from various disciplines simultaneously work on the same or different building components. Monitoring such environments manually for documentation is very time-intensive and error-prone, especially for big construction sites. At the same time, the number of qualified on-site personnel in the construction area is too small to meet the increasing demands of politics in respect of housing construction

[2]. Automated methods and computer vision approaches promise to significantly benefit tracking on-site assembly of construction elements in closer detail. However, the research community is limited in developing algorithms of such kinds since extensively inter-linked monitoring datasets are largely absent [3]. Therefore, a holistic, detailed, and continuous dataset of the on-site as-performed construction processes, including as-built construction products, is highly requested.

This work presents the first step of an extensive multi-sourced data-collection method for automated shell construction monitoring. The base of this work is an affordable, ready-to-use camera setup for cranes that automates RGB image acquisition. Furthermore, a complementing data pipeline is specifically designed for construction site monitoring. The introduced acquisition approach is scalable, allowing it to be easily deployed on construction sites of various sizes. Using photogrammetric reconstruction, point alignment, and projection methods, the underlying 3D reference between the images and the BIM model, is established for selected construction phases, resulting in a per-instance database for direct data consumption. RDF databases have proven their extensibility and ability to link data across domains and life cycle phases in various construction-related use cases [4]. For this reason, we have chosen an RDF database to implement the per-instance database.

In the following background section 2, we dive into the state of the art of the leading technologies and concepts employed in this work and highlight our contribution. Afterward, the proposed workflow is presented in section 3. In section 4, the presented methodology is showcased on a construction site in Regensburg, Germany. To close off, we discuss the overall results in section 5 and conclude with a summary and future works in section 6.

2 Theoretical Background

To set the theoretical background of the paper, we introduce related works in the field of construction monitoring. Moreover, established algorithms to derive a 3D reconstruction from 2D images are explained, which set the starting point for the present paper. Finally, with the shortcomings of existing approaches, we state our contribution to the current state of the art.

2.1 Construction Monitoring

Previous research in construction monitoring has shown serious potential in documenting the as-performed construction process. Various monitoring systems have been used on construction sites targeting different purposes. Existing approaches focus on investigating construction sites, including safety management [5], vehicle and resource detection [6], and construction activities [3]. In addition, other studies focus on generating as-built models by continuously monitoring the construction site [7].

However, research approaches shown above [3], [5], [6] can barely be used to fulfill the need to generate large-scale datasets of construction sites. This is due to several points: First, the introduced methods manually and statically generate data. Here the data acquisition depends, e.g., on LiDAR scans, resulting in considerable efforts. Yet, a dynamic and automatic data acquisition method is needed to compare target vs. actual over time steadily. Second, less focus has been put on the acquisitions for as-planned vs. as-performed comparison. This research area is still in its infancy. Third, many studies only focused on monitoring one specific construction site, not proving their scalability. Undoubtedly, there is a need for a generalized scalable approach for generating extensive datasets across various construction sites. The link between the data and the as-planned information lies there while at the center of attention.

2.2 3D reconstruction for facilitated progress track

As opposed to conventional construction progress track such as with means of textual logs, graphical representations have become indispensable. The reconstruction of the 3D space gives additional meaning to the recorded images and notes and offers stakeholders a unique way of localizing recorded events. At the core of the linkage between recorded images and the as-planned BIM model is the extraction of 3D information from a 2D image - a topic that has received much attention from the computer vision community. Researchers have chosen different methods to achieve the linkage [6]. The directest way is achieved if the camera's extrinsic and intrinsic parameters are known. In this case, the 2D-pixel to 3D-point correspondence can be found. However, this method is susceptible to errors if outdoor cameras are moved by wind or rotated by a crane. Repeated calibration would be needed. Perspective-based methods use a set of key points, lines, and/or surfaces to register the image information to the BIM [8]. A common way to reconstruct 3D information from images is to match unique feature points previously defined by specialized algorithms such as Scale Invariant Feature Transform (SIFT). Matches are generated between similar feature points of overlapping views taken from different viewpoints. Finally, the depth and normal information are computed for every pixel with triangulation. The final 3D point cloud is achieved by fusing depth and normal information of multiple images. In recent years AI has been shown to estimate the needed camera parameters well enough to perform single image registration [9]. This method is, however, still suspected to show limitations in highly chaotic scenes such as construction sites.

2.3 Contribution

Our work makes an essential step toward an automated continuous event monitoring of construction sites and offers significant value by tracking all structural elements in the shell assembly phase. Using multidisciplinary technologies along the workflow allows us to present an end-to-end solution: from construction-site events to a historical database queryable for each structural element given in the planning model. Additionally, the provided result is use case agnostic, provides a sound basis for automated reasoning, e.g., progress tracking or construction safety, and was designed with the potential for extensibility.

3 Proposed Workflow / Methodology

In this section, we first put light on the continuous data-acquisition setting and hardware systems mounted on the construction site before we dive deeper into the post-processing steps.

3.1 Hardware setup

The camera monitoring system is designed for various building construction sites and can be deployed flexibly, considering the following aspects: First, the number of monitoring systems installed at the construction site is determined by the number of cranes. Second, the system is crane agnostic, which is usable for bottom and top-slewing cranes. Third, the critical requirement for installing the system is the power supply at the bottom of the crane, which makes this system easy to install since there are no existing network requirements. The key components of the setup are the router, the Virtual Private Network (VPN), the Power over Ethernet (PoE)-Switch, the cameras, the local server, and the remote server, illustrated in fig. 1. The outdoor-proof router enables a local network with an independent mobile internet connection. The router and the Raspberry Pi, configured with VPN services, build a gateway to the setup allowing remote configurations. Water-sensitive systems, the VPN, PoE, and the local server are sealed in a water-proof box. The PoE-Switch supplies all devices with internet and power.

On each crane, three outdoor-proof surveilling cameras are placed. The location and orientation of the cameras depend on the crane type and the construction project, covering large areas of the site. Cameras take pictures every thirty seconds and send them directly to the local FTP server,

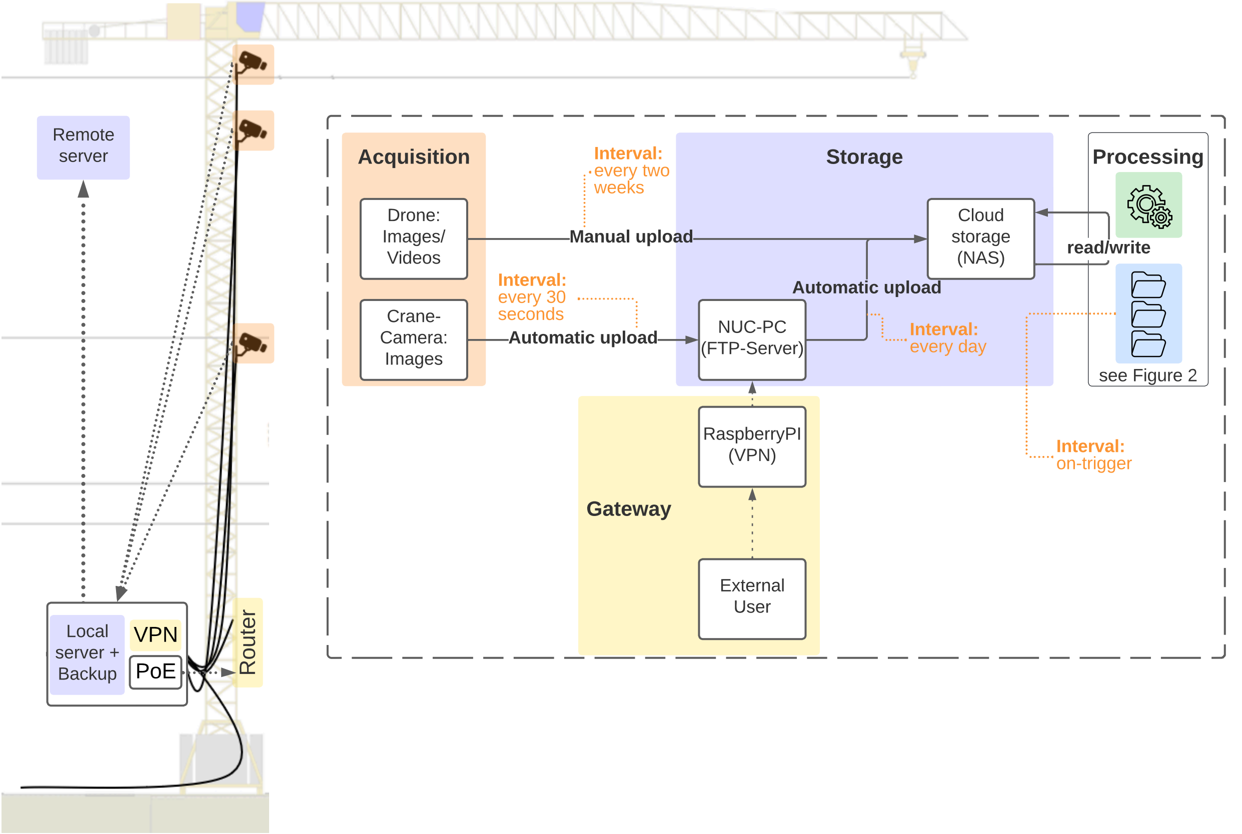


Figure 1: Hardware setup of the image acquisition

the NUC-PC. The continuous data acquisition enables steady documentation of the as-performed construction progress. The server processes the images by creating a backup and uploading the pictures to the remote server once a day. The remote server has extensive storage capabilities and is the system's endpoint. The processing of the images takes place on the remote server on-demand. The following sections will explain a technical workflow to treat the extensive data.

3.2 Data processing

This section presents all steps of the data processing pipeline: Starting from the input data to the 3D reconstruction, the point cloud alignment, and projection onto the BIM model to the final linked database. Figure 2 shows the overall technical workflow for processing the input data and structures this section.

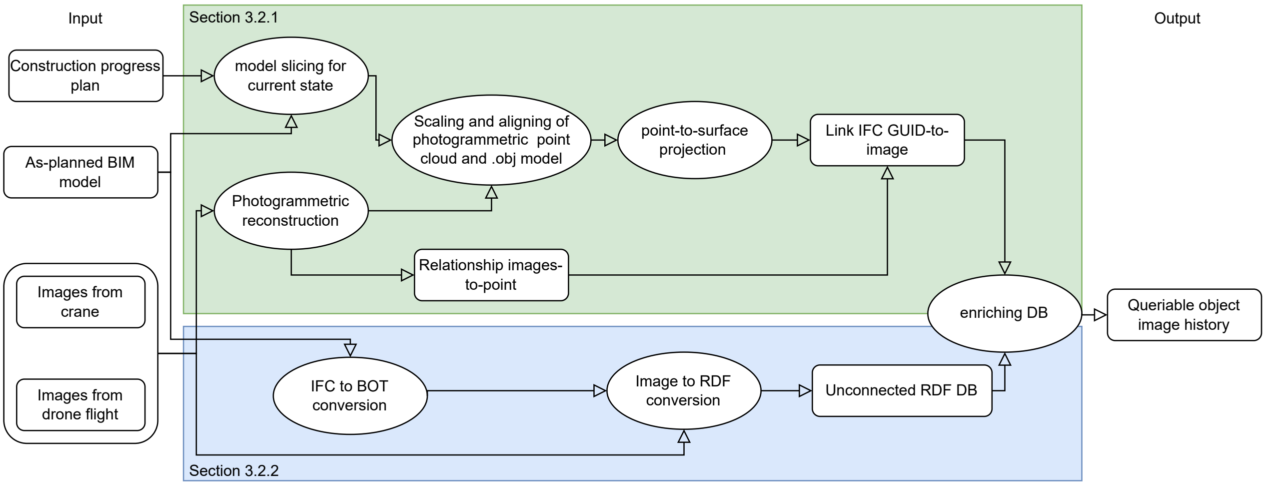


Figure 2: Technical workflow for data processing. Round shapes denote processes whereas rectangles denote states of data.

3.2.1 Linking images to BIM elements

A 3D reconstruction of the images taken from the construction site is needed to register the images to the 3D BIM model objects. To that end a photogrammetric reconstruction process is deployed using Colmap [10]–[12]. To minimize false matches due to weather effects (e.g. shadows) the images should ideally be chosen selectively, this however was set out of scope in this research. It is hypothesized that the construction progress within a short time range in the early morning or late afternoon is not considerable enough to impede SfM from matching feature points in pair images. The result is a 3D point cloud lacking absolute scale and orientation. Depending on the set-up, namely the number of cranes and cameras, the crane viewpoints alone might not cover enough view angles for the reconstruction to work well. In such cases, drone imagery is occasionally used to generate additional point references. For each reconstructed point, the initial overlapping and matched images used for its calculation are known. The point cloud is then aligned and scaled manually to overlay with the BIM model. Outlier removal and noise removal are common ways to reduce the scattering of photogrammetric point clouds, and both are applied here.

Once the as-is acquisition is located in the proximity of the corresponding as-designed, an iterative projection algorithm is used to find the closest IFC element to each point. The projection combined with the image-to-point relationship yields the final link between IFC GUID and image ID. The reconstruction process would be performed at a given interval according to the needed granularity of process monitoring. The database described in the next section is hence continuously updated.

3.2.2 RDF Per-Instance Database

An RDF database is created to uniformly access and query multi-sourced data from the design and construction phase. For the present use case, information from the BIM model is connected with the image data captured on the construction site. Since existing ontologies can describe the topology of buildings, images, and sensor data, the data modeling is limited to combining already defined ontologies. The applied ontologies are BOT, a well-established ontology to describe the topological structure of buildings [13], and Dublin Core Terms to describe pictures and their metadata [14].

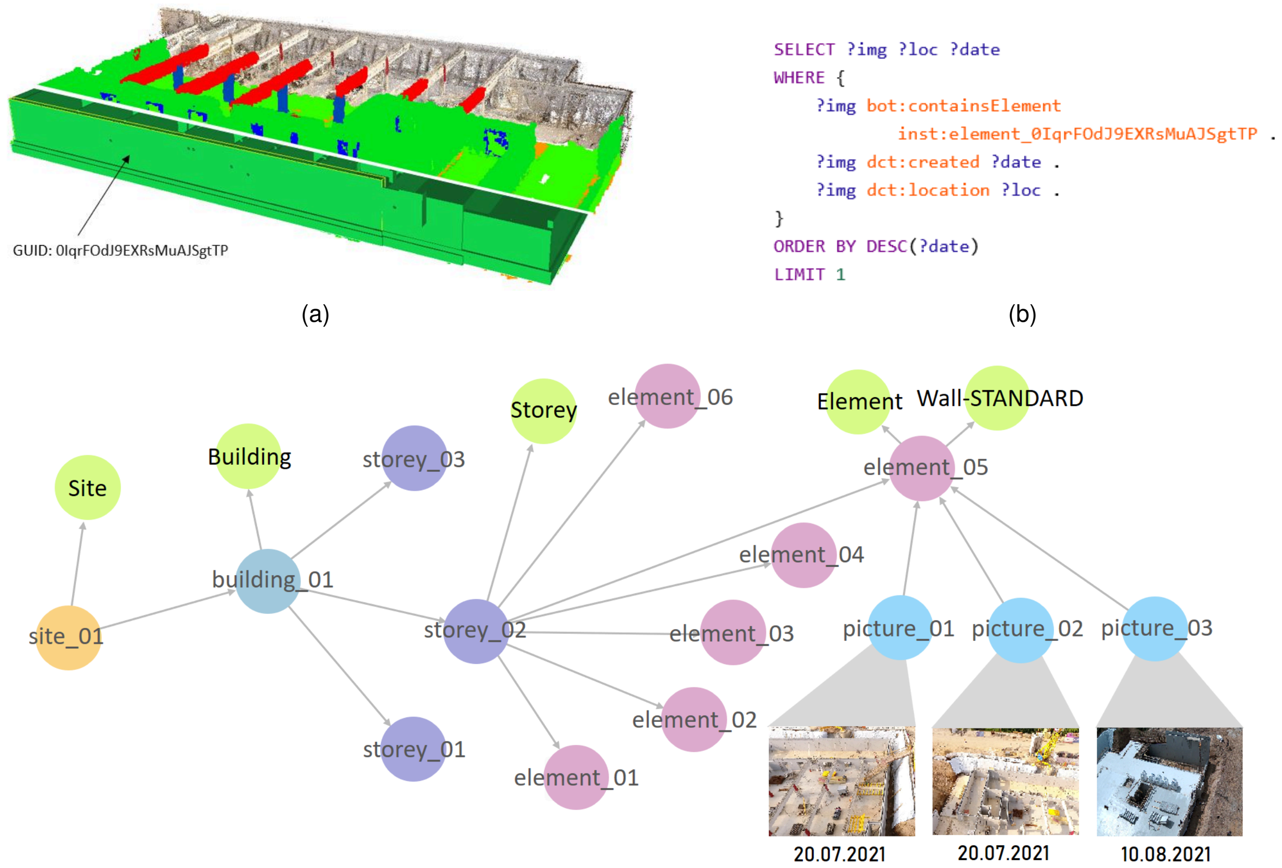
For the conversion of the IFC to RDF, the IFCtoLBD converter by Bonduel, Oraskari, Pauwels, *et al.* [15] is used. This translates the fundamental structure of the building with its stories, rooms, and building elements into its BOT representation. This structure serves as a starting point to connect data from the construction site to the corresponding building elements and spaces. Afterward, a self-implemented algorithm steps through the repository with all construction site images and translates their metadata to Dublin Core terminology. Since the pictures themselves can not be adequately represented in RDF format, only the file's location within the repository is referenced in the graph. Finally, the links between the pictures and the building elements that are visible on them, as obtained by the method described in section 3.2.1, are added explicitly to the RDF graph.

With the resulting RDF graph and SPARQL queries, extracting information related to a specific building element or picture instance, e.g., requesting the most recent picture of a building element, becomes a simple task. Overall, such an RDF database is a solid base for further analysis that can easily be enriched with additional data because of the flexible RDF structure.

4 On-site showcase

We conduct a first roll-out of the described method on a construction site in Regensburg, Germany with a footprint of 1200m^2. The carcass construction was performed with one bottom slewing crane from July to December 2021. The images were taken at an interval of 1/minute from 3 camera positions (1 on the cantilever and the two others at different heights from the tower). The diverse view angles and scene overlaps required for the optimal photogrammetric reconstruction are only partially satisfied in Regensburg since 1. the crane stands on one side of the construction site and 2. the crane movements are faster than what is mapped between the minute interval. Due to these limitations, the presented method is extended to include a series of drone images from the corresponding flight and acquisition dates to improve the overall result. The subset of crane images of all three cameras is combined with a few snapshots of the drone camera. Detailed analysis of how many images of either source were needed is out of scope for this contribution. The photogrammetric reconstruction results in point clouds accurate enough to perform the link between images and points, see figure fig. 3a. The deviation between the as-planned BIM model and the as-performed construction is assumed to be slight enough for the process to work. This reconstruction has been performed for two selected construction stages to showcase the usability: once the beams and walls in the underground storey are just completed (see fig. 3a) and in the other stage, the slab structure and stair elements are added.

The RDF database created from the described data is queriable with SPARQL query language, as fig. 3b shows. The query returns the most recent image linked to the selected BIM instance. The result of querying the database for the wall element highlighted in fig. 3a is shown in fig. 3c. It includes images of the same wall element, both sourced from the crane cameras and the drone flights from two separate dates. More complex queries are possible to receive more meaningful as-performed construction information.



(a)

```
SELECT ?img ?loc ?date
WHERE {
    ?img bot:containsElement
            inst:element_0IqrFOdJ9EXRsMuAJSgtTP .
    ?img dct:created ?date .
    ?img dct:location ?loc .
}
ORDER BY DESC(?date)
LIMIT 1
```

(b)

(c)

Figure 3: From point cloud to queryable database: a) BIM-to-cloud correspondence and projection for underground storey b) query that returns all elements linked to the specific BIM GUID c) excerpt from resulting RDF graph

5 Discussion

The developed camera system proved robust straightforward to install in the first roll-out and is already beeing tested on other construction sites. It allows us to generate extensive image data sets of construction sites at a low cost. Thereby, continuous and automated documentation of the as-performed progress is enabled. Regarding the data processing pipeline, a few limitations need additional attention. In this case, we addressed the limitations of too few view-angles and too little overlap in images with additional drone-sourced images. Still, we will further investigate this in future construction sites by equipping multiple cranes with the camera setup systems and increasing the frame rate to have more overlap between images. Alternatively, neighboring buildings or movable cameras offer additional placement and coverage possibilities. Given the setup's objective of monitoring shell construction only, it is acceptable to notice that the monitoring only covers the outside top-most elements. To go beyond this, additional indoor surveying must be performed with movable camera systems and registered with the outdoor images.

One short manual step is included in the pipeline, namely the point cloud to BIM model alignment and scaling. This is currently impeding the full scalability of the database updates and suggests further investigation. Furthermore, the computation times of the data processing are resource-intensive, especially for the photogrammetric reconstruction step, and suggest an analysis for alternative pipelines touching aspects highlighted in section 2. The quality of the photogrammetric reconstruction (e.g., amount of noisy points) directly influences the linking accuracy in the final database. Any little error in the image matching process will lead to erroneous links. A quantitative analysis of the error rate would require a manual detection of construction instances in images and lies beyond the scope of this work. Finally, it must be said that the images linked to the instances are, especially in early phases, still taken from far distance making the element appear only amongst many others. An additional mask could be laid on the images to show which pixels have been used to reconstruct the point in question.

6 Conclusion

With the proposed method, we achieve a continuous yet low-cost data acquisition of the shell construction processes by documenting the assembly steps taken from the crane's field of view. The image registration to the BIM objects using photogrammetric reconstruction, alignment, and point projection is successful; each BIM element is linked to the corresponding images if sufficient capturing overlap from sufficient view angles is given. We then apply SWT to store this information in an RDF database using established ontologies and allow for a quick query-based process investigation. The use case agnostic database might benefit managers for verifying safety compliance or verifying schedules or future research methods that can profit from the BIM links to enhance computer vision-based image interpretation.

The goal of the workflow presented in this article demonstrated viability as a whole and scalability in the data acquisition phase. Yet, the processing requires time-intensive intermediate steps and a manual one. In the future, we plan to investigate further and optimize the scalability and full automation of the data processing steps. More specifically, this encompasses more optimal camera placement and frame rates, an image classification prior to reconstruction to prevent unsharp or irrelevant images from causing noise in the point cloud, or an analytical comparison to other 2D-3D reconstruction methods. The publishing of our dataset is planned for a future point in time. Conclusively, our approach shows an autonomous way for future construction supervision using state of the art processing steps, enabling on-demand documentation for construction managers and researchers.

Acknowledgements

We thankfully acknowledge Innovation Management Bau for their financial support and for providing us access to multiple construction sites to conduct case studies. Moreover, the research described in this paper received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 958398, "BIM2TWIN: Optimal Construction Management & Production Control". Therefore, the authors also thankfully acknowledge the support of the European Commission in funding this project. Finally, we thank the chair of geodesy at OTH Regensburg for letting us use their Faro laser scanner and supporting us with their expertise on the construction site, and the MINT-Labs for allowing us to install cameras on their roof.

References

[1] R. Sacks, I. Brilakis, E. Pikas, H. S. Xie, and M. Girolami, “Construction with digital twin information systems”, *Data-Centric Engineering*, vol. 1, e14, Nov. 2020, ISSN: 2632-6736. DOI: 10.1017/dce.2020.16. [Online]. Available: https://www.cambridge.org/core/product/identifier/S2632673620000167/type/journal_article.

[2] G. Zika, T. Maier, A. Mönnig, *et al.*, “Die folgen der neuen klima-und wohnungsbaupolitik des koalitionsvertrags für wirtschaft und arbeitsmarkt”, 2022.

[3] G. Torabi, A. Hammad, and N. Bouguila, “Two-dimensional and three-dimensional cnn-based simultaneous detection and activity classification of construction workers”, *Journal of Computing in Civil Engineering*, vol. 36, 4 Jul. 2022, ISSN: 0887-3801. DOI: 10.1061/(ASCE)CP.1943-5487.0001024.

[4] P. Pauwels, S. Zhang, and Y. C. Lee, “Semantic web technologies in AEC industry: A literature overview”, *Automation in Construction*, vol. 73, pp. 145–165, 2017, ISSN: 09265805. DOI: 10.1016/j.autcon.2016.10.003.

[5] T. Cheng and J. Teizer, “Real-time resource location data collection and visualization technology for construction safety and activity monitoring applications”, *Automation in Construction*, vol. 34, pp. 3–15, 2013, ISSN: 09265805. DOI: 10.1016/j.autcon.2012.10.017.

[6] J. Xue, X. Hou, and Y. Zeng, “Review of image-based 3d reconstruction of building for automated construction progress monitoring”, *Applied Sciences (Switzerland)*, vol. 11, no. 17, 2021, ISSN: 20763417. DOI: 10.3390/app11177840.

[7] S. Tuttas, A. Braun, A. Borrmann, and U. Stilla, “Acquisition and consecutive registration of photogrammetric point clouds for construction progress monitoring using a 4d bim”, *PFG–journal of photogrammetry, remote sensing and geoinformation science*, vol. 85, no. 1, pp. 3–15, 2017.

[8] K. Asadi and K. Han, “Real-Time Image-to-BIM Registration Using Perspective Alignment for Automated Construction Monitoring”, no. March, pp. 388–397, 2018. DOI: 10.1061/9780784481264.038.

[9] D. Acharya, R. Tennakoon, S. Muthu, K. Khoshelham, R. Hoseinnezhad, and A. Bab-Hadiashar, “Single-image localisation using 3D models: Combining hierarchical edge maps and semantic segmentation for domain adaptation”, *Automation in Construction*, vol. 136, p. 104 152, 2022, ISSN: 09265805. DOI: 10.1016/j.autcon.2022.104152. [Online]. Available: https://doi.org/10.1016/j.autcon.2022.104152.

[10] J. L. Schönberger and J.-M. Frahm, “Structure-from-motion revisited”, in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[11] J. L. Schönberger, E. Zheng, M. Pollefeys, and J.-M. Frahm, “Pixelwise view selection for unstructured multi-view stereo”, in *European Conference on Computer Vision (ECCV)*, 2016.

[12] J. L. Schönberger, T. Price, T. Sattler, J.-M. Frahm, and M. Pollefeys, “A vote-and-verify strategy for fast spatial verification in image retrieval”, in *Asian Conference on Computer Vision (ACCV)*, 2016.

[13] M. H. Rasmussen, M. Lefrançois, G. F. Schneider, and P. Pauwels, “BOT: The building topology ontology of the W3C linked building data group”, *Semantic Web*, vol. 12, no. 1, pp. 143–161, 2020, ISSN: 22104968. DOI: 10.3233/SW-200385.

[14] K. Coyle and T. Baker, *Guidelines for dublin core application profiles*, 2009. [Online]. Available: https://www.dublincore.org/specifications/dublin-core/profile-guidelines/ (visited on 05/25/2022).

[15] M. Bonduel, J. Oraskari, P. Pauwels, M. Vergauwen, and R. Klein, “The IFC to linked building data converter - Current status”, in *6th Linked Data in Architecture and Construction Workshop*, vol. 2159, 2018, pp. 34–43.