Review of Image-Based 3D Reconstruction of Building for Automated Construction Progress Monitoring

Jingguo Xue [1,*] **, Xueliang Hou** [1] **and Ying Zeng** [2]

[1] School of Economic and Management, North China Electric Power University, Beijing 102206, China; houxueliang@ncepu.edu.cn
[2] State Grid Mianyang Power Supply Company, Mianyang 621000, China; zengying7@163.com
* Correspondence: jingguo1994@gmail.com

Abstract: With the spread of camera-equipped devices, massive images and videos are recorded on construction sites daily, and the ever-increasing volume of digital images has inspired scholars to visually capture the actual status of construction sites from them. Three-dimensional (3D) reconstruction is the key to connecting the Building Information Model and the project schedule to daily construction images, which enables managers to compare as-planned with as-built status and detect deviations and therefore monitor project progress. Many scholars have carried out extensive research and produced a variety of intricate methods. However, few studies comprehensively summarize the existing technologies and introduce the homogeneity and differences of these technologies. Researchers cannot clearly identify the relationship between various methods to solve the difficulties. Therefore, this paper focuses on the general technical path of various methods and sorts out a comprehensive research map, to provide reference for researchers in the selection of research methods and paths. This is followed by identifying gaps in knowledge and highlighting future research directions. Finally, key findings are summarized.

Keywords: construction progress monitoring; 3D reconstruction of building; daily construction image; building information model

1. Introduction

Schedule and cost have always been the focus of construction management. Early detection of actual or potential schedule delay or cost overrun provides opportunities for timely adjustments. This requires an automated, timely, and accurate progress-monitoring system to detect deviations between the planned process and the actual performance. In current practice, prevailing monitoring and management systems in the Architecture, Engineering, Construction (AEC) and Facilities Management (FM) industry are still dominated by traditional approaches, including manual paper-based collection and recoding of on-site activities [1,2]. These procedures are time-consuming, labor-intensive, and error-prone, which cannot be performed as frequently as required. Moreover, current methods may not be conducive to a clear and quick understanding of progress. Because the progress reports in text and graphic format are visually complex, they cannot intuitively reflect information related to space, so it often takes a while for managers to understand the status of progress, which affects the efficiency of information transmission [3].

Building Information Modeling (BIM) is an essential step to digital management of construction projects [4,5]. BIM creates a three-dimensional (3D) model of building that can be used to represent construction process (4D BIM) by linking activities of a schedule with corresponding building elements. It provides an opportunity to visually compare deviations between the planned process and the actual performance. Recently, several approaches and studies that address the comparison of as-built and BIM-based as-planned data have been presented. The as-built data came from barcoding, Radio-Frequency Identification (RFID), Ultra-Wideband (UWB), Geographic Information System

(GIS), Global Positioning System (GPS), laser scanners, image-capturing devices, and so forth. Among them, only laser scanners and image-capturing devices can realize the 3D reconstruction from the construction site to the digital model, which is a necessary part of automated construction progress monitoring in the future.

The goal of general 3D reconstruction is to infer the 3D geometry and structure of objects and scenes from one or multiple two-dimensional (2D) images [6]. In the AEC industry, 3D reconstruction of building is the key to connecting the BIM and the project schedule to daily construction images, which enables managers to compare as-planned with as-built status and detect deviations and therefore monitor project progress [7]. At present, 3D reconstruction methods of buildings are mainly divided into two categories: one is to directly generate 3D point clouds of the building through laser scanning, and the other is to take 2D images and then reconstruct a 3D model.

3D laser scanning technology, based on laser ranging principle, can quickly reconstruct the 3D model of the measured object by recording the 3D coordinates, reflectivity and texture of many dense points on the surface of the measured object [8]. This technology has unique advantages in efficiency and accuracy, and is not affected by illumination. However, there are also some shortcomings, such as high cost, high time consumption, and high technical requirements for operation [9–11].

In this article, the second method, image-based 3D reconstruction, is focused on. The image includes photographs, videos, and depth images. Among them, photographs and videos are RGB images, while depth images are RGB-D images. The spread of camera-equipped devices has promoted the explosion of image data. Massive images and videos were recorded on construction sites daily, and the ever-increasing volume of digital images had inspired scholars to visually capture actual status of construction sites from them. In comparison to other alternatives such as laser scanning, 3D reconstruction from images is at a fraction of cost [12], and the rich color and texture information retained in the image can also be used for semantic recognition and process reasoning.

Compared with the 3D reconstruction of general objects or the 3D reconstruction of building based on laser scanning, the image-based 3D reconstruction of building has different characteristics and faces special challenges. The research contents of the image-based 3D reconstruction of building can be roughly divided into six processes: image collection, 3D point cloud generation, image-to-BIM alignment, point cloud segmentation, point cloud semantic recognition, and progress reasoning. The main challenges or tasks of each process are as follows:

1. Although here are many ways to collect images including monocular cameras, binocular cameras, and multi-cameras, the challenges are similar. The intensity of light and shadows seriously affect image quality, and there are many dynamic and static occlusions in addition to self-occlusion at the construction site that prevent researchers from directly observing the building. These factors have brought huge obstacles to 3D reconstruction from images.
2. To generate point cloud from images, the feature points in different images need to be found first. Many algorithms have been studied, such as Scale-Invariant Feature Transform (SITF) [13] and Speeded-Up Robust Features (SURF) [14]. Then, these feature points need to be matched with each other to estimate the fundamental matrixes using algorithms such as Random Sample Consensus (RANSAC) [15]. When the images are taken by a moving camera, it is required to reconstruct the point clouds using Structure from Motion (SfM) [16].
3. There are various registrations in the process of 3D reconstruction of building, such as 2D–3D and/or 3D–3D registration among images, point clouds, and BIM models.
4. The point cloud generated from images is massy and complicated, which contains background, noise, obstacle, etc. Removing the redundant point cloud and keeping only the Region of Interest (RoI) are beneficial for simplifying data processing and improving calculational efficiency.

5. The point cloud usually only contains 3D coordinate information. To obtain a semantically rich model to infer progress, it is required to identify the type and state of building components represented by each point from the color and texture in RGB images, which is called semantic recognition/labeling of point clouds.

6. Process reasoning is the last step including geometry-based, appearance-based, and relationship-based reasoning.

To solve the above challenges and tasks, many scholars have carried out extensive research and produced a variety of intricate methods. In technical articles, however, the review part usually takes the related technologies and methods as the bedding for the follow-up discussion, and does not make a comprehensive analysis of the related technologies. At the same time, most of the review articles tend to focus on one perspective, such as point cloud [17,18], big data [12,19], data collection [2,20–22], algorithm [23], etc., which does not unify all the methods. Therefore, researchers cannot clearly identify the relationship between various methods to solve the difficulties.

This paper will sort out a comprehensive research map, and describe the relevant research results, to provide reference for researchers in the selection of research methods and paths. The goals of this article are three-fold: (1) to integrate the advanced image-based 3D reconstruction methods of buildings to form a research map; (2) to compare the differences among various methods and highlight the advantages and limitations of these methods; (3) to discuss the current challenges of image-based 3D reconstruction of building and explore feasible solutions. What should be noted is that the 3D reconstruction mentioned later is the image-based 3D reconstruction of building; the image refers to photographs, videos, and depth images; the buildings refer to civil infrastructures; and the 3D reconstruction refers to the reconstruction from reality rather than from Computer Aided Design (CAD) drawings.

The remainder of this paper is structured as follows. The first part briefly introduces the general process of the 3D reconstruction of building, and the representation of knowledge are described to limit or unify related concepts, and then six key steps of image-based 3D reconstruction of building are analyzed which covers the state-of-the-art. In the subsequent part, six important knowledge gaps for the image-based 3D reconstruction of building are explained in detail. The limitations and challenges are highlighted, and the future research directions are discussed. In the last part of the paper, key findings are summarized.

2. Methodology

This review focuses on the image-based 3D reconstruction in the field of construction progress monitoring, hoping to obtain a comprehensive technical path, which provides technical selection reference for researchers. To achieve this goal, the following work was carried out in this study.

1. Literature search and screening: This study searched the relevant research results since 2008 from Google Scholar, and the key words included image, photography, video, depth image, computer vision, three-dimensional reconstruction, construction progress monitoring, construction progress tracking, etc. Then, articles related to this topic were selected, and the papers indexed by Web of Science were focused on. Finally, a total of 66 articles were selected, as shown in Table 1.

2. Method classification: The knowledge and methods used in these papers are divided and classified into these six aspects: knowledge representation, image collection, and 3D point cloud generation, image-to-BIM alignment, point cloud segmentation, point cloud semantic recognition, and progress reasoning.

3. Methods comparative analysis: The methods of each aspect were classified and summarized, and the advantages and limitations of various methods were analyzed.

Table 1. Literature sources.

Publication	Number of Articles
Advanced Engineering Informatics	9
Automation in Construction	22
IEEE Conference	5
ISPRS Journal of Photogrammetry and Remote Sensing	5
Journal of Computing in Civil Engineering	4
Journal of Construction Engineering and Management	2
Journal of Information Technology in Construction	2
Sensors (Basel)	2
Others	15
Total	66

3. Technology Path of Image-Based 3D Reconstruction

This section presents a comprehensive synthesis of the state-of-the-art in image-based 3D reconstruction. By categorizing these existing studies, a research map is summarized. Figure 1 illustrates the research map for image-based 3D reconstruction starting from data collection to progress reasoning. The upper portion categorizes the as-planned models which is ready before construction, including geometry models, and relationships. Correspondingly, the bottom portion illustrates the as-built models that is collected on the construction site and reflects the actual construction status, including images and point clouds. Through the interaction among the as-planned models and the as-built models, combined with other technologies, the construction process is inferred from the geometry-based, relationship-based, and appearance-based information. The process can be divided into six steps: image collection, 3D point cloud generation, image-to-BIM alignment, point cloud segmentation, point cloud semantic recognition, and progress reasoning. Table A1 in the appendix shows literature related to these steps. In the subsequent part, the reconstruction process will be described in detail and at length and the advantages and possible obstacles of the state-of-the-art methods will be analyzed.
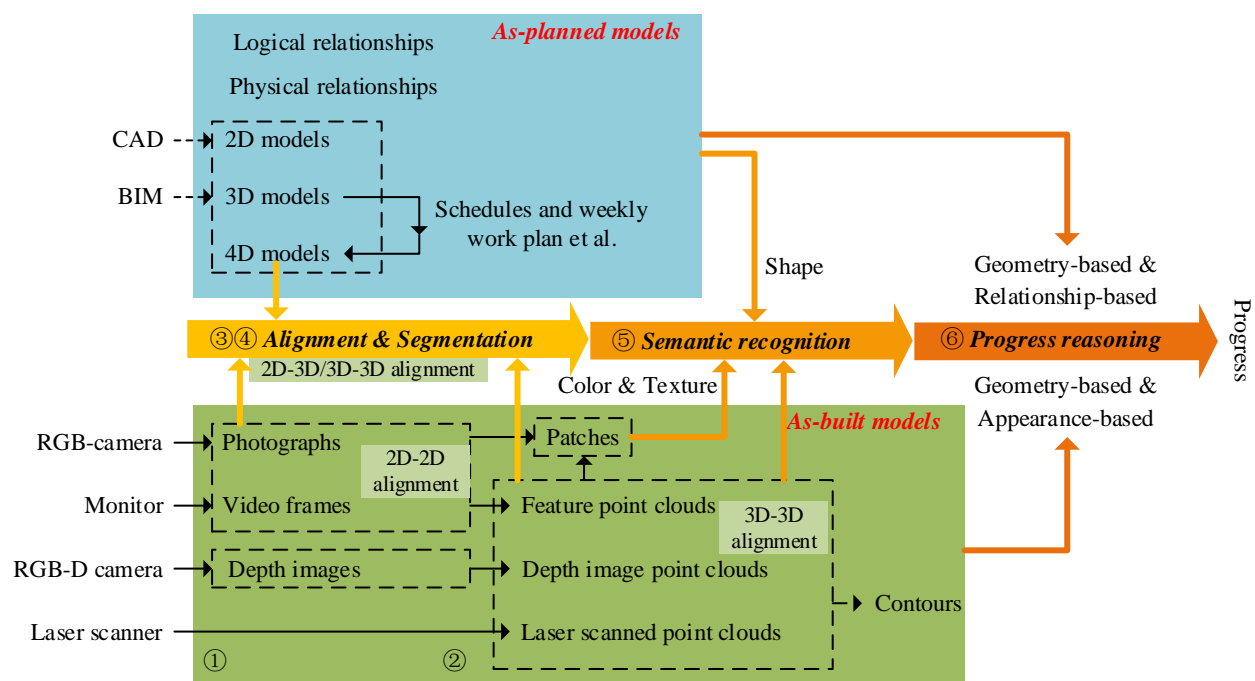
Figure 1. Research map for image-based 3D reconstruction.

3.1. Representation of Knowledge

In the field of visual 3D reconstruction, the representations of knowledge are diverse, such as 2D/3D/4D models, schedules, physical, and logical relationships, images and videos, point clouds, contours, patches, and so on. These representations of knowledge can be roughly categorized on three fronts:

1. Direct as-planned information: 2D/3D/4D models are widely known as the as-planned information which depict the planned process and final states, and the core purpose of these as-planned models, during the 3D reconstruction process, is to serve as a reference standard. Schedules and weekly work plan representing project execute process are usually combined with 3D models to form 4D models. Physical relationships represent the spatial connection between geometric primitives (including aggregation, topological and directional relationships) [24], while logical relationships represent the sequential relationship among building components due to procedural or technical requirements, similar to the construction sequence under the constraint of the activity-on-arrow network. Both physical and logical relationships can be used to assist decision-making [25].

2. Direct as-build information: Image is one of the most common as-build information including photographs, videos, and depth images. With the recent advances in smart devices and camera-equipped platforms, an exponential growth in the volume of images and videos that are recorded on construction sites [12,26]. Compared with ordinary photographs, depth images/RGB-D images generated by the range camera contain depth information, which makes them easy to generate as-build point clouds. Furthermore, laser scanned point cloud is also a common way to represent as-build models.

3. Derived information: Derived information comes from images or point clouds and provides support for the construction process reasoning. First, the point clouds derived from images or videos are a kind of derived information. Presently, taking real-time videos or time-lapse images and then aligning these sequential frames/images via feature detection, matching, and homography transformation to generate point clouds are common practices [12,27]. In addition, if point clouds are projected onto the plane which runs parallel to the floor/wall, contours of buildings can be extracted through the algorithm from Suzuki [28,29] to reason walls, doors, windows or other apertures [29]. Moreover, there is also a lot of useful information generated from images. For example, some researchers project 3D model elements onto image planes and the images are segmented into patches for the progress reasoning [30]. Furthermore, the image patches can be used for creating multiple discriminative material classification model and the Construction Material Library (CML) for the progress reasoning [31].

3.2. Image Collection and 3D Point Cloud Generation

3.2.1. Data Acquisition Device

In the AEC industry, many devices are used for image acquisition, including camera (monocular/binocular/camera array), smart devices (mobile phone/tablet/personal laptop), monitor, UAV with camera, laser scanner, depth camera (Kinect), satellite, etc. Main performance indexes of these devices are shown in Table 2.

The data generated by these devices can be divided into two categories: image and point cloud. The laser scanner for generating point cloud has the characteristics of high equipment cost, high technical requirements, limited texture information, etc., so that it is not accepted by most construction companies. Therefore, the image (including photograph and video) becomes an alternative way. Images can be collected by a variety of devices, and most of these devices have the characteristics of low cost, low technical requirements, portable, high-resolution, rich-texture information. This makes image-based 3D reconstruction a key technology for automated construction progress monitoring.

Table 2. Device performance comparison.

| Device	Raw Data Type	Cost	Technical Threshold	Portability	Resolution	Texture Representation	3D Surfaces Reconstruction
Camera	Photograph	Low	Low	High	High	Rich	-
Smart devices	Photograph	Low	Low	High	High	Rich	-
Monitor	Video	Low	Low	Low	Medium	Rich	-
UAV	Video	High	High	Low	Medium	Rich	-
Laser scanner	Point cloud	High	High	Low	High	Limited	Automatic
Depth camera	Photograph and point cloud	Medium	Low	High	High	Rich	Automatic
Satellite	Remote sensing images	High	Low	-	Low	Medium	-

3.2.2. Data Type

In different studies, the form of images is related to the acquisition equipment and affects the selection of subsequent methods. There are three main forms used by most researchers.

1. Time-lapse images/videos from fixed camera: Fixing the camera position means simplifying complicated registration process. As long as the camera coordinates and camera shooting direction are obtained, the images can be registered with the BIM model after simple rotation and scaling. Although this means a lack of flexibility in response to occlusions caused by changing structures, the benefits of always-on-demand images provide the possibility for fast and responsive assessment [32]. However, to reduce occlusion, it is necessary to increase the number of cameras shooting from multiple angles [33,34], as shown in Figure 2, which raises new questions—how to arrange multiple cameras and how to deal with data conflicts between cameras. In addition, the cameras need to be fixed on a stable object, which sometimes proves difficult. In addition, Golparvar-Fard et al. [33] found that small errors will significantly affect registration and minimize the allocated image area for each element, making the task of recognition much more challenging.

2. Unordered image sets: Unordered image sets can be taken from any location, so that almost all corners can be captured without occlusions. These images are usually taken by construction managers, owner representatives, contractors, and subcontractors and have capacity to enable complete visualization of a construction site [3]. However, developing computer vision and image processing techniques that can effectively operate on such imagery is a huge challenge [3]. Golparvar-Fard et al. [3,35] came up with a way—extract SIFT feature points from continuous images, match them to estimate the fundamental matrixes using the RANSAC algorithm, and use the SfM principle to generate point clouds, as shown in Figure 3. In this method, although the image sets can be unordered, the images in an image set are orderly, and a certain proportion of repeating regions among these images is needed to extract corresponding SIFT points. In addition, the user needs to initially register the as-planned and as-built models [35]. When there are many unordered image sets, it is necessary to manually record the camera position and external parameters, and each image set requires an initial registration that is quite troublesome. Moreover, to avoid occlusion and cover all observed objects, a large amount of overlap is necessary, which is almost impossible for manual acquisition. Some researchers use camera-equipped Unmanned Aerial Vehicles (UAVs) to professionally take images and document them [36,37], which allow for a wider range of views, especially from above, and the GPS coordinates and camera orientation are known in most cases. Even so, it is still very difficult and is tedious to find the exact views in BIM due to the inaccurate GPS coordinates especially in the vertical axis [12].

3. Depth images: Depth images, generated from range cameras/RGB-D cameras, contain not only RGB colors but also depth information, as shown in Figure 4. Similar

to the point cloud generated by laser scanners, the 3D point cloud model of the observed object can be generated directly from the depth images. The range camera has attracted the attention of many scholars because of its low cost and portability. However, due to the limited shooting range, it is only suitable for indoor shooting, not for large-scale image acquisition [38,39].
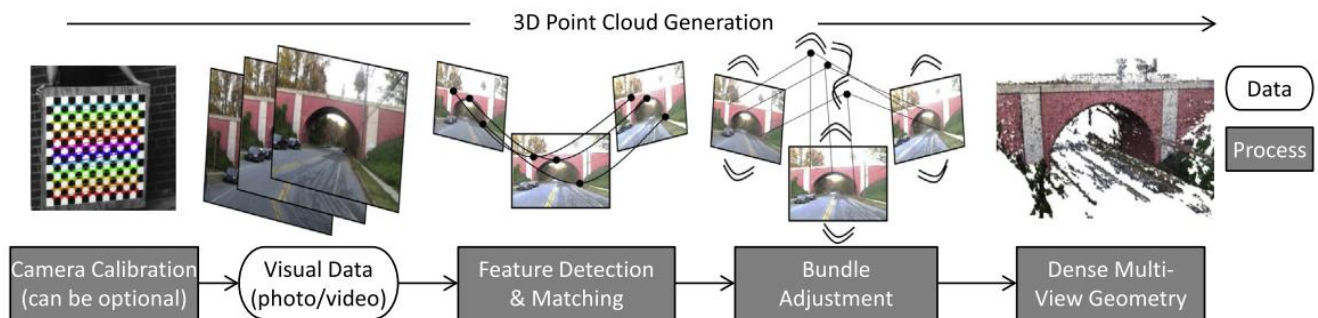
Figure 2. Schematic process for image-based 3D point cloud generation [17].
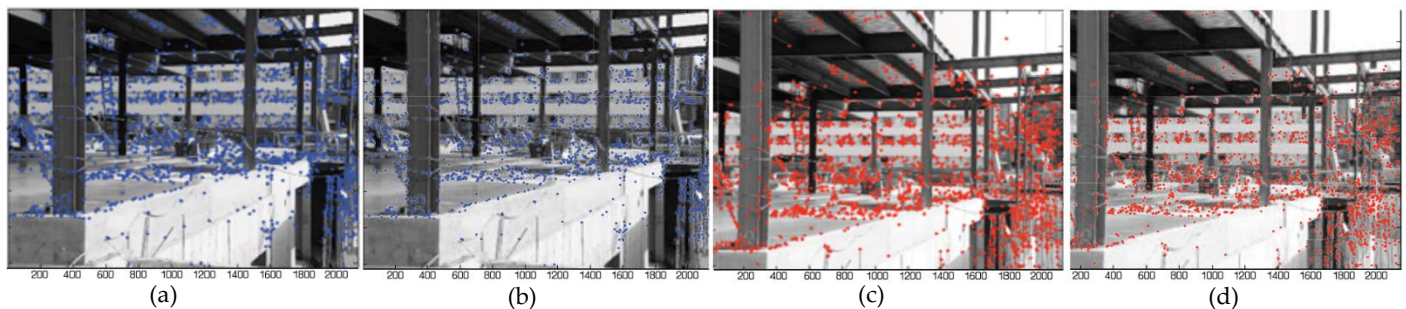
Figure 3. Left to Right: Visualizing keypoint matching between a pair of images shown as (**a**) Image1-Before RANSAC, (**b**) Image1-After RANSAC, (**c**) Image2-Before RANSAC, (**d**) Image2-After RANSAC [3].

Figure 4. A photo taken from the construction site and its depth map [40].

3.3. Image-to-BIM Alignment

The existing registration (or alignment) methods for 3D reconstruction can be categorized into two forms: one is the registration between homogeneous partial data to form a global model, including image–image (2D–2D) and point cloud-point cloud (3D–3D) alignment; the other is the registration between different types of data, such as image–BIM (2D–3D), point cloud–BIM (3D–3D), and/or image–point cloud–BIM (2D–3D–3D) alignment. Since these alignment processes start from the original data (images) and finally to BIM, the whole process will be called image-to-BIM alignment in this article.

To analyze the construction performance, an as-is condition needs to be compared to an as-planned condition [12]. The image-to-BIM alignment intends to make the acquired images comparable to the as-planned information contained in BIM [4]. There are four approaches proposed in recent years to support image-to-BIM alignment.

1. 3D–2D registration-based: Monitoring the construction process using fixed cameras without pan/tilt/zoom is one of the most convenient ways. Because once the user initially registers the as-planned and as-built models, the correspondence between the photograph and the virtual model would be set for all subsequent images [33]. Many scholars superimpose 3D visual models on images in Augmented Reality (AR) or Visual Reality (VR) environments [3,33,41]. Ideally, all visual models could be projected on the image plane and fully registered with the image. However, the outdoor camera is susceptible to environmental influences such as gravity and transverse winds, which can easily lead to the failure of automatic registration. Therefore, a set of key points with known positions in the photograph and the 3D visual environment is required to achieve more accurate registration [32].

2. Feature point-based: To avoid the problem of occlusion caused by fixing cameras, some scholars explore the methods of movable cameras. Golparvar-Fard et al. [3,35] studied a method of extracting SIFT feature points from unordered image sets. By identifying the common feature points of overlapping region, these images were registered with each other to generate feature point cloud. Then, the images and the virtual model were registered by aligning the feature point cloud with the 3D virtual model in 4D Augment Reality (D4AR) environment. They realized the image-to-BIM alignment through the registration of image-to-point cloud and point cloud-to-BIM. In addition, many automated methods have also been proposed to register point clouds with BIM models. Bueno et al. [42] presented a novel method (4-plane congruent set algorithm) for automatic registration of as-is 3D point clouds with 3D BIM models. Lei et al. [43] proposed a 3D patch registration approach based on Convolutional Neural Network (CNN) deep-learning algorithm for integrating sequential models in support of progress monitoring.

3. Depth image-based: Compared with the above method, the image-to-BIM alignment method based on the depth image simplifies the process of image registration and point cloud generation because the depth information is included in the depth image. In the research of Pučko et al. [38], workers captured all workplaces inside and outside of the building in real time and record partial point clouds, their locations, and time stamps by Kinect (helmet-mounted scanner). Then by manually picking the equivalent points, the partial point clouds were registered and merged into a complete 4D as-built point cloud of a building under construction. Finally, the image-to-BIM alignment was realized using a software developed at the University of Maribor [10]. Although the early process of image registration and point cloud generation has been simplified, the process of picking the equivalent points manually was not subtracted, which requires a lot of manual work. It is time-consuming, sometimes it must be repeated, the result is not precise and leads to limited usefulness.

4. Perspective-based: This method uses the relationship among points, lines, and surfaces in images to directly register the image with BIM. For example, Kropp et al. [4] and Asadi et al. [44] proposed a new method to register images with BIM using perspective alignment for indoor monitoring of construction, as shown in Figure 5. First, video frames were captured with a monocular camera system to create as-built data of the current construction status. Second, the first frame was registered initially with the BIM model by superimposing the wire frame model on it in an AR manner. Then the fine correspondence between the model and the as-built scene was calculated from line candidates extracted by scanning as-built images RoI. Although only the first frame needs manual alignment, each video needs manual processing, which obviously requires a lot of manual labor. Because each room needs a separate video, and they need to be shot daily. Similarly, Fernandez-Labrador et al. [45] propose a

novel procedure for 3D layout recovery of indoor scenes from single 360° panoramic images. The proposed method combined geometric reasoning and deep learning to generate a pruned set of lines belonging to the main structure of the room, from which they extracted candidate corners and generated layout hypotheses. These alignment methods register the as-built image directly with the as-planned model without the assistance of point clouds. However, these technologies only have been applied in the indoor decoration stage with little occlusion, which may not be applicable to the outdoor scenes.
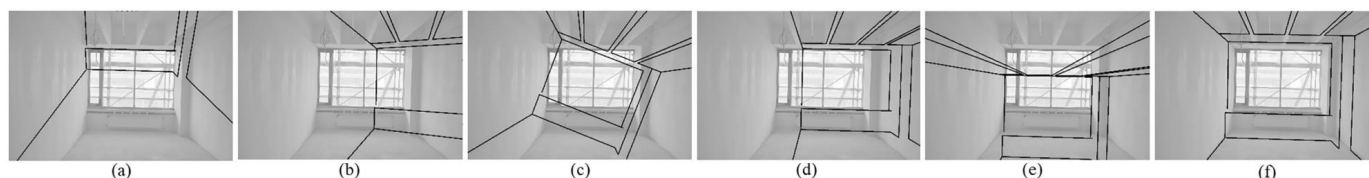
Figure 5. Projections of input model lines with divergent camera poses by rotation (**a**–**c**) and translation (**d**–**f**) [4].

3.4. Point Cloud Segmentation

No matter what method is used to generate point clouds, there will be a lot of noise, background, and obstacles (equipment, materials, personnel, tools, protective measures, garbage, etc.). The messy and redundant point clouds not only waste computer resources but also affect judgment. Therefore, it is necessary to segment the point cloud to delete the redundant points outside the RoI. There are various computational methodologies proposed to conduct point cloud segmentation. Wang et al. [46] divided them into six categories: clustering-based, edge-based, region-based, graph-based, model fitting-based, and hybrid. The advantages and disadvantages are shown in Table 3.

Table 3. Summary of data segmentation methodologies for point cloud data.

Segmentation Methodologies	Advantages	Disadvantages	Ref
Clustering-based	Easy to understand and implement	Accuracy problem: sensitive to the noise in data and is influenced by the definition of neighbor	[47,48]
Edge-based	Fast segmentation	Accuracy problem: sensitive to noise and uneven density of point clouds	[49,50]
Region-based	More accurate to noise	Over or under segmentation and accuracy of determining boundaries	[51]
Graph-based	Better on complex point cloud data with uneven density or noise	Cannot process in real time, and training or other system is required to assist process	[50,52]
Model fitting-based Hough transform	Fast and robust with outliers	Slower and more sensitive to segmentation parameters	[53]
RANSAC	Fast and robust with outliers, can process a large amount of data	Accuracy when processing different point cloud sources	[54]
Hybrid	Take advantage of multiple approaches more accurate	Contain all disadvantages of selected approaches	[55]

In addition, many scholars used the geometric primitives of BIM models to segment point clouds [50]. After aligned with BIM models, the point cloud was naturally divided into different regions, which is great for geometry-based reasoning. The specific analysis will be introduced in Section 3.6.

3.5. Point Cloud Semantic Recognition

The point cloud generated by laser scanning is a group of indistinguishable points that only contains the 3D coordinate of points. The points representing various components are glued together. However, the point clouds generated from RGB or RGB-D image can contain color, texture, and other information. They can be delivered to the point cloud

through the mapping relationship between the point cloud and the image to help identify which primitive a point belongs to.

Many scholars mark semantic information for point clouds in various ways, which is called semantic recognition/labeling. The key to semantic recognition is to establish semantic mapping. In general, practice, the point cloud is segmented into small homogeneous 3D patches, and then the features (including color, position, height, compactness, linearity, planarity, angle with the ground, etc.) of each patch are extracted to classify these patches to form semantic point cloud. Antonello et al. [56] proposed a multi-view frame fusion technique to enhance the semantic labeling results with 3D entangled forests and built semantic maps on RGB-D point cloud. The point cloud was over-segmented into homogeneous 3D patches and a feature vector of length 18 was calculated for each patch. Five binary tests defining the entangled features were used to describe complex geometrical relationship between segments in a neighborhood. Posada et al. [57] presented a purely semantic mapping framework which operates solely with omnidirectional images. The free space was found from the omni-image with a binary floor/obstacle classifier. In addition, a place category classifier was used to label the navigation relevant categories: room, corridor, doorway, and open room. Adán et al. [58] divided the semantic modeling process into five semantic levels, including (1) automatic data acquisition of the building's as-is state, (2) simple geometric building model, (3) recognition and labeling of primary structural elements (SEs) of the building, (4) recognition of openings within SEs of the building, and (5) recognition of small building service components on SEs, as shown in Figure 6. The integrated system they proposed can automatically reconstruct large scenes at a high level of detail and provide detailed as-is semantic models of building. Dimitrov [59] presented an image-based material classification method for semantically rich as-built 3D modeling, and a CML was formulated to train and test the proposed method, as show in Figure 7. Although their method was only used on images, it is feasible to graft this technology into point cloud semantic recognition.
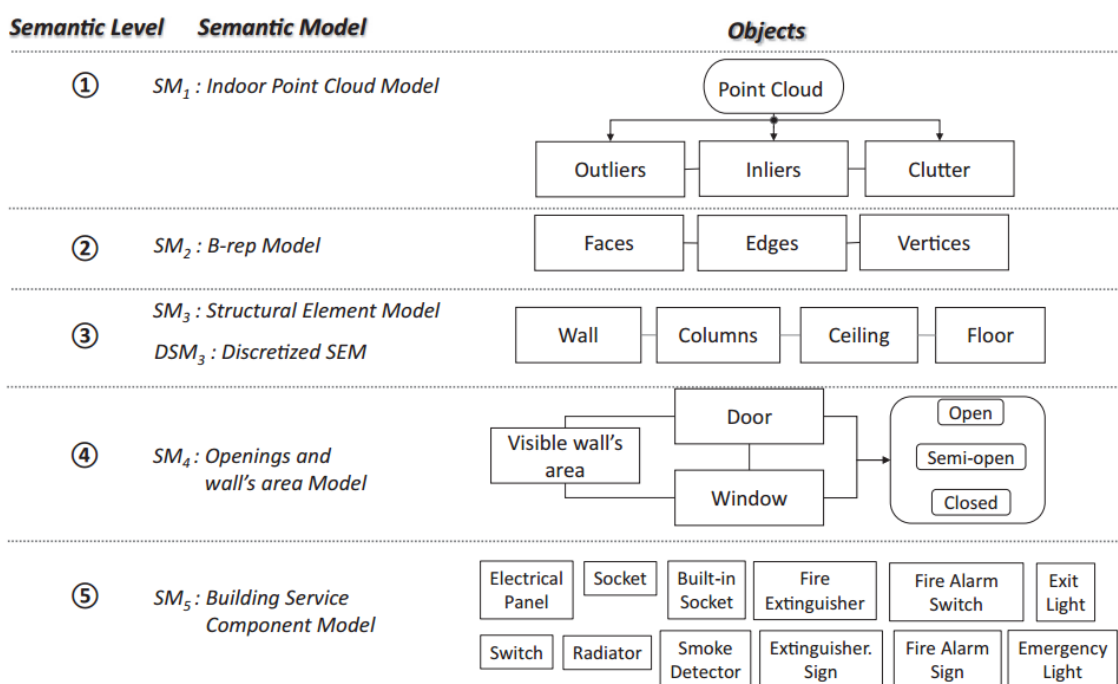

**Figure 6.** Levels of semantic 3D modeling [58].

Figure 7. The construction material library [59].

3.6. Progress Reasoning

Progress reasoning is a key approach that compares the as-panned model with the as-built model and detects deviation between them. In the last few years, many methods for progress reasoning have been proposed. These reasoning processes are based on geometry, appearance, relationship, and so on. In this paper, these methods are classified into four categories:

1. Based on the 3D space occupancy by the point clouds: Braun et al. [25] split the BIM element surface into 2D raster cells and verified the progress information by the number of points extracted for each raster cell within a certain distance before and behind the BIM element surface. Omar et al. [1] created internal and external surface planes for BIM model and measured the true column heights by the point cloud between the external and internal surface boundaries. Golparvar-Fard et al. [35] traversed and labeled for expected progress visibility and a machine-learning scheme built upon a Bayesian probabilistic model was proposed that automatically detects physical progress.

2. Based on the 2D plane projection of the point clouds: Rebolj et al. [10] and Pučko et al. [38] projected a BIM element to three orthogonal planes and rasterized them within a regular grid. Then, they projected the points in the element's proximity onto the same grids and the area of grid-cells containing projected points is considered to be a covered area. Finally, they identified the existing elements by assessing the percentage of elements' surface being covered by the point cloud. Volk et al. [29] projected the point clouds onto a plane which runs parallel to the floor generating a heat map, from which a closed loop providing the room's floor plan were construct, as shown in Figure 8. On this basis, 3D points were projected onto the walls creating an image per wall to extract contours which were characterized into windows, doors, or other apertures.

3. Based on the image changes of 3D–2D projection area: Kim et al. [60] applied 3D CAD-based image mask filters to identify the construction progress of a cable-stayed bridge on background with little noise, which may not be appropriate for complex environments. Zhu and Brilakis [61] identified the segmented image region using machine-learning techniques to determine whether the region was composed of concrete or not. The concrete identified by this method is a whole area, not refined to the component. For this defect, Ibrahim et al. [32] segmented the image into a set of discrete component masks and analyzed the texture or color changes of specific regions of interest related to each component to infer the timings of significant events. Unfortunately, most of these changes were related to spurious lighting and other variable conditions, such as equipment, or scaffolding being moved. Then Han and Golparvar-Fard [30] proposed a new appearance-based material classification method for monitoring construction progress deviations at the operational-level. They used pre-trained multiclass material classifier to recognize the texture of the region of inter-

est, rather than only based on the change of color. Afterward, Han et al. [62] combined the geometry-based and appearance-based reasoning methods for detecting construction progress, which had the potential to provide more frequent progress measures.

4. Based on the relationships of geometric primitives: Sometimes occlusions are inevitable. It is a wise choice to use auxiliary information to reason progress, because it can greatly reduce the duplication of effort in the data collection phase and the ambiguity of recognition results. The auxiliary information includes physical relationships (aggregation, topological and directional relationships) and logical relationships between objects or geometric primitives. Nuchter and Hertzberg [63] represented the knowledge model of the spatial relationships with a semantic net. Nguyen et al. [64] automatically derive topological relationships between solid objects or geometric primitives with a 3D solid CAD model. Braun et al. [25] attributed these relationships to technological dependencies and represented these dependencies with graphs (nodes for building elements and edges for dependencies). However, there is controversy about the use of ancillary information. For example, Ibrahim et al. [32] pointed out this approach would not be totally reliable, since the only way to truly gain confidence that a component is finished is to visually verify it. They suggested a combination of multiple sources of image to increase the overall reliability.
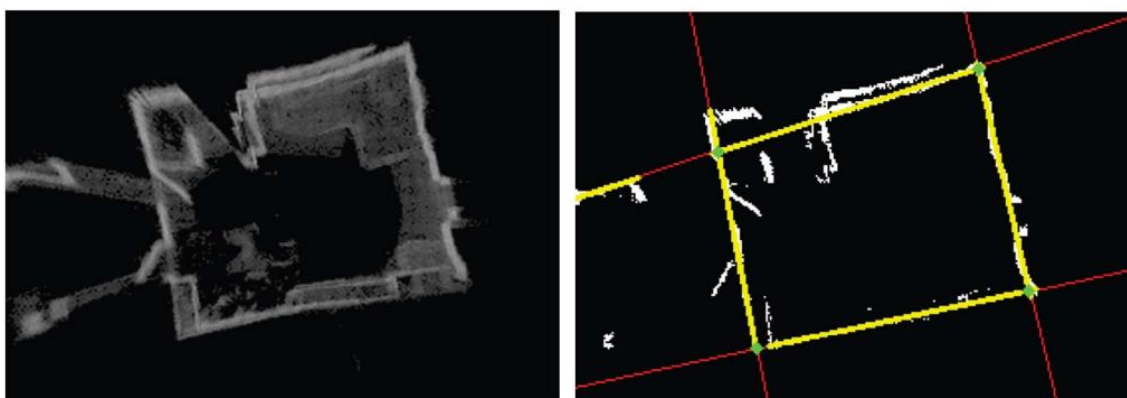
Figure 8. Results of wall detection. Visualization of a heat map provided by projecting the point cloud onto the floor (lightened up to improve visibility) (left). The analyzed heat map (right) [29].

4. Knowledge Gaps and Challenges

Literature shows that image-based 3D reconstruction techniques for project monitoring are still under development, and there remain research gaps that need to be addressed for image-based modeling techniques to become standard practices [7]. Some of these gaps are highlighted in the following paragraphs.

4.1. Occlusions and Limited Visibility

In the implementation of 3D construction, occlusions are inevitable and the most challenging issues that must be addressed. Occlusion is defined as any blockage of the camera vision by a physical object [65], which results in incomplete data and challenge reasoning under limited visibility [30]. Occlusion can be classified into two main categories based on its source, static occlusions which are self-occlusions caused by progress itself (e.g., a facade blocking the observation of elements in the interior) or occlusions caused by temporary structures (e.g., scaffolding or temporary tenting), and dynamic occlusions which is a result of movable objects (e.g., laborers, machines, etc.) [35], as shown in Figure 9.
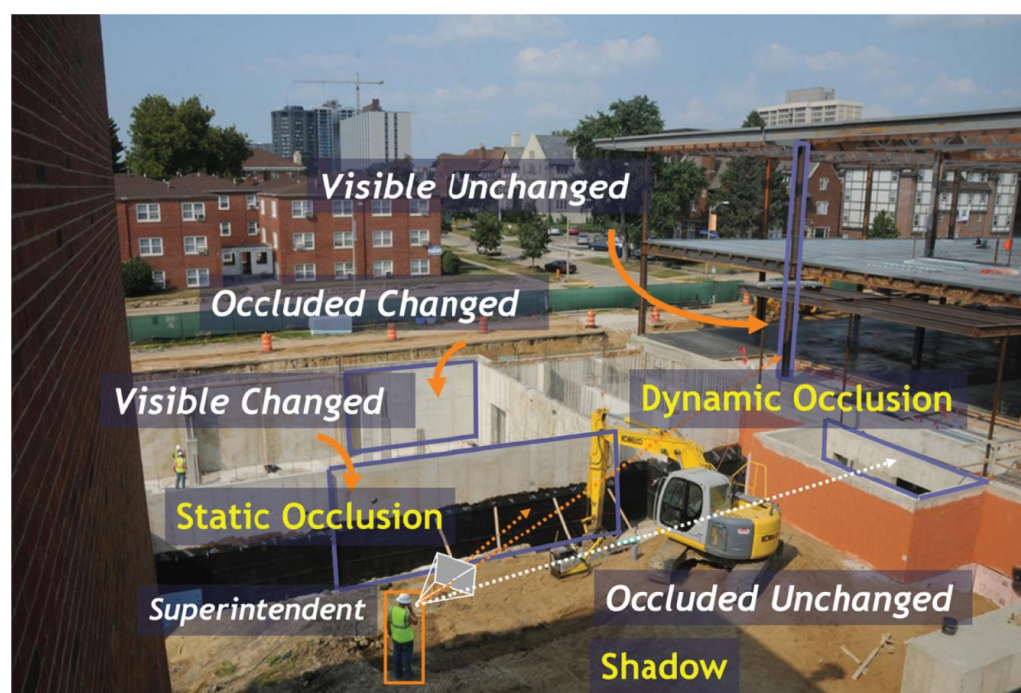
Figure 9. Occlusions [35].

To reduce dynamic occlusion, Omar et al. [1] decided to capture site photos after the duty time (i.e., after 5:00 p.m.). The selected time significantly reduced the dynamic occlusions for captured photos because the site was shut down and there are no active laborers or machines.

Compared with the dynamic occlusions, static occlusions are unavoidable. In particular, time-lapsed images or videos from fixed camera only show what is within range and field-of-view of the camera [3]. Golparvar-Fard et al. [33] described two different scenarios on horizontal and vertical occlusions and the challenges of visualizing progress only on a single view. To reduce the occlusion, a network of multiple cameras was used to realize the coverage of the whole building [34]. Golparvar-Fard et al. [33] suggested finding the optimum location of a network of cameras to make sure all the elements could be monitored [33]. However, it still cannot be used to track progress inside the building after the building envelope is placed.

This motivated scholar to use an unordered set of progress imagery that is taken from various viewpoints to tackle the occlusion issue [3,33]. These images and videos collected by digital cameras and smartphones were usually taken by field personnel, including construction managers, owner representatives, contractors, and subcontractors. Although they have the capacity to enable complete visualization of a construction site, these images are typically uncalibrated and their locations and orientations are unknown, which makes it very hard to accurately localize them with BIM [12].

In addition to the above direct avoidance of occlusion, scholars have also proposed some indirect methods to avoid the effects of occlusion. One approach is to use prior knowledge, for example, the projections of BIM models to define the RoI to guide the process of identifying [12,32]. In this case, a lot of occlusions can be ignored because the result can be obtained as long as a certain proportion of the target area meets the recognition requirements. In addition, in recent studies, advanced deep-learning technology has been used to identify construction components in images, which can deal with partial occlusion [66]. These methods can only be used for partial occlusion rather than full or almost all occlusions. In the more negative case, the semantic net describing the spatial and logical relationships between objects of geometric primitives can be used [63]. Objects that are easily recognizable can be detected first, and then more challenging structures can

be inferred using the information in semantic net [24]. In addition, the semantic net can also provide an effective way to verify the recognition results.

4.2. Lighting and Shadow Conditions

The camera is an optical sensor, so the image is extremely sensitive to light intensity. The quality of images collected under different lighting conditions varies greatly. Poor lighting results in blurry pictures, and the point cloud is inaccurate and noisy. Especially in the construction site, the similarity of the surface texture of many materials and adverse light conditions makes the appearance-based reasoning difficult. Furthermore, the constantly changing shadows during the day add a lot of messy lines to the image and make the same material show a completely different appearance. Various illumination, shadows, weather, and site conditions make it difficult to perform consistent image analysis on such imagery [3], as shown in Figures 10 and 11. In the face of this situation, using laser scanning is a more ideal solution, although there may be some problems such as cost and technical requirements.

Figure 10. Different weather conditions during a construction project: (**a**) fog; (**b**) rainy; and (**c**) snow [33].

Figure 11. Effect of shadow on a single working day: (**a**) 3:00 PM; (**b**) 4:00 PM; (**c**) 6:00 PM [33].

The above-mentioned various attempts to solve the occlusion problem were also applied to solve the problem of light intensity and shadow. Therefore, they will not be repeated here.

4.3. Indoor 3D Reconstruction

Continuous monitoring of the construction process is necessary, both indoors and outdoors. Compared with outdoor, indoor construction monitoring contains more contents. First, many elements need to be arranged (e.g., pipeline and cable installation, surface decoration, and fire protection) which makes detailed progress monitoring challenging [67]. Second, many construction activities occur indoors. All these cover a significant portion of the whole project and the delays associated with them can result in costly consequences and re-scheduling of the project [68]. Third, when the work moves indoors, the need for situation awareness and monitoring increases because of many trades involved including

site managers, framers, insulation installers, electricians, drywall installers, plasterers, painters, and laborers [69].

However, the interior construction sites are complicated, congested, and frequently changing. Especially in civil buildings, the available operating space is extremely limited. Therefore, some 3D reconstruction techniques are neither directly applicable indoors nor validated for interior sites [4], although most of them are applied on outdoor construction sites. Of course, many scholars have also explored the construction progress monitoring under the indoor scene, providing interesting directions, such as the results of Antonello et al. [56], Fernandez-Labrador et al. [45] and Volk et al. [29].

4.4. Non-Automated Image-to-BIM Registration

Schedule deviation is derived from comparing the as-planned model and the as-built model, and this process is based on their registration. In recent years, several research contributions have been presented that address the registration of the images/point clouds and the corresponding BIM models. It is expected to realize fully automatic construction process monitoring with fast information feedback. However, most practices either depend on manual intervention for the registration or work automatically under severe constraints [4]. According to the type of data collected, the challenges faced by automatic registration are analyzed below.

1. Time-lapse images or videos from fixed camera: Since the cameras are fixed, it is convenient to manually register each camera only once. However, the scenes captured by this method are so limited that it is only suitable for shooting large-scale scenes. Therefore, it is necessary to equip multiple high-resolution cameras at the construction site. Even so, they are still severely affected by lighting conditions and there are still many unavoidable occlusions [33].
2. Unordered images or videos: To avoid occlusion, some scholars proposed to free the camera and use unordered images that collected by field personal. If the camera is not calibrated, and the position and orientation are unknown, registration is almost impossible. Many researchers took video clips to generate partial point clouds, and then integrate different parts to form a global 3D model. However, each part needs to record the initial state of the camera and be registered. Usually, many clips need to be taken to cover all the details of a building, which is very troublesome [35,59].
3. Image sequences taken with UAVs: Images sequences are usually taken by camera-equipped UAVs and come with GPS coordinates and camera orientation which can used to align the point clouds and BIM. Ideally, only the starting position needs to be registered and all architectural details could be captured at once. However, the effect was not good in practice due to the inaccurate GPS coordinates especial in the vertical axis. The longer the flight path means the greater the accumulated deviation, and it is difficult to build accurate point clouds.

In fact, based on existing technology, these challenges can be attributed to finding a balance between automation and accuracy. Because the higher the degree of automation, the less chance of manual parameters to correct errors. Therefore, in-depth exploration of algorithms is needed in the future to find reasonable solutions.

4.5. Troubles of Point Cloud

The current 3D reconstruction process mainly relies on the point cloud, but there are some inherent shortcomings in point cloud-based 3D reconstruction. First, the point cloud-based method requires extensive computing resources to process huge amounts of data. It is time-consuming to remove all points of the backgrounds and the objects of no interest [50,62,70]. Second, there is no guarantee on the completeness of point clouds, and sufficient overlaps among images are required to cover all areas of interest [12,71]. In large-scale projects, the area of point clouds collected at one time is limited. If the area of one-time acquisition is too large, the accuracy of the point clouds would be relatively low affecting the effect of 3D reconstruction, while reducing the acquisition area would lead

to the soared acquisition cost. Third, the point clouds also have problems such as high noise, difficulty in segmentation and registration [46]. Therefore, are there other ways to reconstruct the as-built model without point clouds? Like image recognition. This is still an open challenge that needs further exploration.

4.6. Disputes about Prior Information

As mentioned above, point clouds and BIM models can be used to infer geometric changes and construction progress, and image information such as color and texture can be used to make the inference more accurate. However, due to the complex structure of the building nested in large and small spaces, self-occlusion is inevitable. Therefore, it is obviously unrealistic to obtain information of all the building components only through images.

In many studies, prior information, such as logical and physical relationships between objects or geometric primitives, is used to assist reasoning or reduce the ambiguity of recognition results. Usually, such relationships are represented by semantic net [24,63], as shown in Figure 12. However, it must be acknowledged that this information will not be totally reliable, since the only way to truly gain confidence that a component is finished is to visually verify it [32]. If the prior information plays a major role, most of the results can be inferred from them, as shown in Figure 13. These results may not be consistent with reality violating the original intention of monitoring. Therefore, how to reasonably use prior information is a question that needs to be explored.
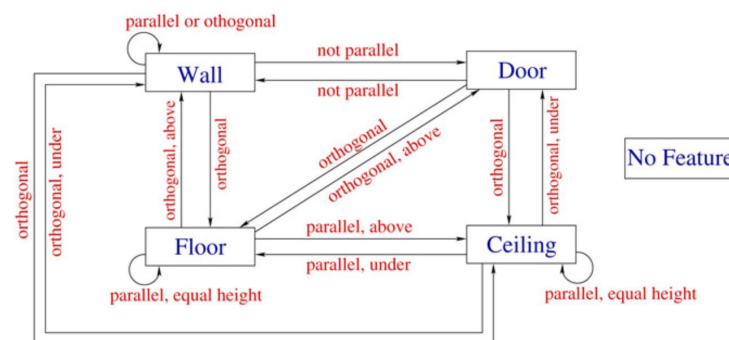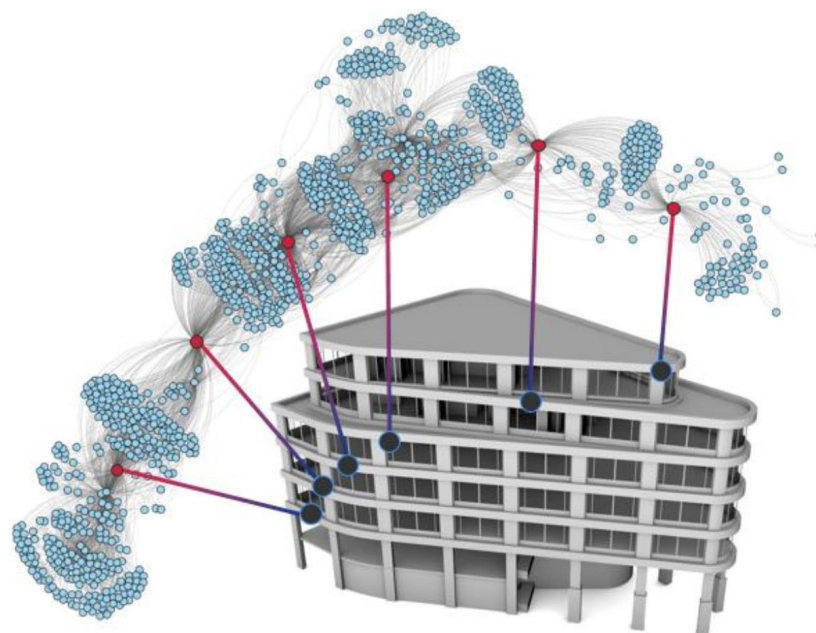

Figure 12. Semantic net [63].

Figure 13. Precedence relationship graph [25].

5. Research Findings

Automated construction progress monitoring can reduce schedule delay, enhance information visualization, and assist decision-making. In the past few years, the construction team has used various simulation tools to track construction progress. Unfortunately, while some paperless construction planning and tracking tools are available today, many construction companies do not use them. Because of the threat of cost, time, or complexity, construction companies around the world are putting digital and mobile strategies on the back burner and sticking to their old technologies. Compared with other 3D reconstruction methods, image-based 3D reconstruction seems to be a more critical and feasible technology, despite there are some challenges. Here are summaries of the uniqueness of this technology:

1. Image has more advantages than other forms of data. There are various types of automatic acquisition technologies, which can be roughly divided into Enhanced IT technologies, Geospatial technologies, Imaging technologies, and Augmented reality [20]. However, the imaging technologies, especially photography and video shooting, have the advantages of intuitive, rich information, accurate, low cost, and low technical requirements, which is congenitally advantageous to be accepted by construction companies.

2. Easy and cheap access to massive image data. The acquisition of reliable data is supported by the development of hardware, including camera, monitor, storage device, smartphone, UAV, etc. Daily images can not only be collected systematically, but also recorded by workers on site, due to the diffusion of devices with built-in cameras. Abundant and sufficient data means that enough information can be extracted in theory, while information extraction is up to the software.

3. Rapid development of software technology. The booming new image processing technologies, especially the ones based on deep learning, have been fully and deeply applied in biomedicine, aerospace, transportation, public security, and other industries. However, in the field of AEC, the research and application of these technologies are still in its infancy.

4. Most of the research is based on the point cloud, rather than the image itself. The methods without point cloud are still worth exploring, for example, VR-based registration and object detection-based reconstruction.

5. New technologies that can be combined with image-based 3D reconstruction have emerged. With the vigorous development of hardware, software, algorithms, and data in computers and related industries, various new technologies have emerged. These technologies have made huge breakthroughs and are sought after by researchers. Many scholars have begun to integrate these emerging technologies with existing 3D reconstruction technologies and have achieved amazing results.

6. Discussion

6.1. Contribution

To help researchers clarify the context of related technologies and clearly identify the relationship between various methods, this paper presented a comprehensive research map of the current practices about image-based 3D reconstruction. Following this, the fourth part of the paper focused on a critical synthesis of the main knowledge gaps and challenges in the 3D reconstruction process. Finally, main findings were summarized. In this process, the contribution of this paper is mainly divided into the following two aspects:

1. A more comprehensive technology roadmap is created. Reading and summarizing the previous work, the authors find that the relevant literature only focuses on point cloud-based methods or perspective-based methods, which are applied to outdoor and indoor monitoring, respectively. Few people analyze them together in one article. This paper breaks the barriers between them and obtains a comprehensive technology roadmap. The technical roadmap is new and covers a wider range of methods,

which provides a reference for the integration of indoor and outdoor construction progress monitoring.

2. The knowledge gap ignored by most scholars is highlighted. In the fourth part, the main knowledge gaps and challenges in the process of 3D reconstruction are analyzed, and the solutions are indicated. In addition to the problems concerned by scholars, this part also points out the problems of point cloud and prior knowledge, which are less concerned by scholars.

6.2. Practical Guidance

Through the analysis of related technologies, it is found that the image-based method is the development trend of construction progress monitoring in the future. The image-based method includes multiple branches and processes. In practice, different methods can be integrated, such as image-based modeling, perspective-based method, time-delay photography, target detection, and so on. For different scenarios, appropriate methods can be flexibly selected in terms of equipment, data form, registration method, schedule reasoning method, etc. For example, in the concrete pouring site of high-rise buildings, the latest construction progress is blocked by templates, scaffolds, and protective nets. Both point cloud-based and perspective-based methods fail in such a chaotic scene. Therefore, the target detection technology can be used to infer the construction progress through the context information in the image. In short, in the actual construction process, flexible technology combination should be adopted.

6.3. BIM Technology

Among the various methods discussed, BIM Technology is widely used. This is because BIM provides a visual digital model of the building, which enables the collected data to have a carrier and compare with each other. BIM also has good simulation performance, can carry out 3D visual simulation of design, construction, and other solutions, find problems in the simulation, and solve the problem in the planning stage.

In the digital management of construction projects, BIM is an important step. A major advantage of BIM is the comprehensive collection, linking, and provision of data for different planning, construction, and operation tasks. In the context of construction management, it is very common to apply 4D building model by connecting schedule activities with corresponding building elements. Based on 4D building model, construction sequence can be analyzed, and progress monitoring can be supported.

6.4. Future Development Trends

6.4.1. Combination with VR and AR

In recent years, virtual technology has not only achieved success in the game industry, but also promoted the development of other fields. The virtual technology is a computer simulation system that can use a computer to generate a simulation environment to immerse users in the environment. In the field of AEC, VR technology can provide realistic location and condition of structure element for remote construction monitoring [40,41], while AR technology can superimpose virtual BIM models on the real world to achieve a sensory experience beyond reality. For example, the electromechanical equipment models that need to be installed in the future can be projected into the screen to guide the on-site construction and check whether the construction progress is consistent with the BIM design at any time; information on construction procedures, issues, and attributes can be projected to the front by recognizing the scene and gestures of the wearer; the pipeline can be projected to the ground and walls for precise excavation in the renovation project.

In the 3D reconstruction scene, VR and AR also have application value. They can be used to align as-planned models with as-built models and observe which areas have not been reconstructed on the actual construction site. For example, Rahimian et al. [40] proposed framework for integration of BIM and interactive game-like immersive VR interfaces, which empowered project managers and stockholders with an advanced decision-making

tool. Golparvar-Fard et al. combined daily images and 3D/4D models to create D4AR models [3,33,35,72,73]. In addition, they aligned the as-built point clouds with the BIM model for automated progress deviation measurement [35]. Similarly, in a virtual environment, other forms of data can be integrated with BIM, such as images, laser point clouds, RFID, etc.

For progress monitoring, the combination of real-time 3D reconstruction and VR can make managers in the office as if they were in the construction site, which provides a new way for remote management of progress, quality, and safety, especially for the inspection of dangerous areas.

6.4.2. Combination with Deep Learning

Deep learning is booming recently and has basically replaced previous related technologies. An extraordinary breakthrough has been made in image classification, face recognition, speech recognition, and so on. At the same time, these technologies have also been introduced into the AEC industry, such as face recognition [74], workforce and equipment tracking [75,76], helmet identification [77,78], defect detection [72,73].

In particular, there have been studies on 3D reconstruction of building using deep learning in the past few years. They have focused on material recognition and classification [79,80], point cloud segmentation [48,50,81], automatic registration of point cloud and BIM [43], camera pose regression [82], structural component recognition [66] and so forth. The topic, however, is still in its infancy and further developments are yet to be expected [6]. First, the success of deep learning depends on the availability of data sets, but there are currently no large-scale data sets available in the field of AEC, especially the labeled data sets. To obtain more accurate results, a variety of images are required especially for those scenes that are difficult to recognize. However, the images in a project are similar, so how to combine many projects in the AEC industry to obtain a comprehensive and rich image database for training needs to be further explored. Second, the images in other industries are so different from those in the AEC industry, and thus the general feature-based object detection algorithms are not well suited for construction engineering structural component detection because of the complex spatial structural relationships (such as adjacency, aggregation, and hole inclusion) between various components in large buildings and the insignificant differences in features such as texture and color. Third, most of the state-of-the-art techniques deal with images that contain a single object, but there are a lot of occlusions, shadows, and messy backgrounds in construction site images. In this scenario, how to design a suitable point cloud segmentation strategy combined with deep learning is a hot research topic.

6.4.3. Combination with Big Data

In a construction project, there is a huge variety of available data. For example, monitoring records and various images taken at the construction site; files, records, data, and models generated in the early stage; information that can be further collected, such as the movement tracks of workers and equipment; data of other projects; and so on. However, the reality is that a large amount of useful information was collected and then discarded due to the inability to process the data promptly. How to use these abundant data at the same time to make the best judgment is a challenge.

Big data technology is a technology that quickly obtains valuable information from various types of data. Many new technologies have emerged in the field of big data, and they have become powerful weapons for big data collection, storage, processing, and presentation.

In the field of construction process monitoring, the data presents the following characteristics. First, there is a wide variety of information. In addition to imaging and laser scanning, many other technologies have also been applied for construction progress monitoring, such as barcoding, RFID, UWB, GIS, and GPS. Rich information will help reflect the actual state of the construction site in more detail. Second, images and videos occupy

more memory than type data, with smaller density of useful information. It is very difficult to store these data reasonably and extract useful information from them. Third, there are obstacles in data sharing between different enterprises and different projects. In addition, it is attractive to summarize the general rules used in other projects from these daily images. With these characteristics, the stage is set for big data technology.

Currently, due to the complexity of image processing, the application of big data in the field of AEC is still very weak. However, it is certain that employing big data could move the state of the art in the domain of construction progress monitoring to the next level [19]. In addition, big data analytics will enable massive data to be processed in time to reflect daily changes and update the BIM model and construction schedule accordingly.

7. Concluding Remarks

Construction site images, as instant records of the state of the construction site, contain rich information, which makes them natural materials for automatic construction process monitoring. On the one hand, the popularity of built-in camera equipment makes it feasible to obtain massive free images from the construction site. On the other hand, advanced software and hardware technologies provide powerful tools for extracting useful information from daily images. These make the image-based 3D reconstruction more easily accepted by the market, and it will be the main direction of future development.

At present, various image-based technologies are isolated from each other, such as image-based modeling, perspective-based method, time-delay photography, and so on. In practice, due to the particularity of the scene, the progress monitoring of a project may re-quire a combination of multiple methods. However, few scholars have broken and reorganized the relevant methods, which makes many methods unable to be fully grafted and used. Therefore, this paper combines the relevant technologies and methods into a comprehensive technology path. In this process, various technologies are separated and then integrated, which makes various technologies connect with each other and provides guidance for technology selection. This is very important for both researchers and engineering practitioners.

In addition, in the AEC field, although many methods and technologies have been proposed, the research on image-based construction progress monitoring is still in its infancy. The technology applied to practice is still very few. The knowledge gap and challenges still need to be further explored, such as occlusion, light problems, integration of emerging technologies, etc.

Appendix A

Table A1. Literature category.

Process	Type	References
As-built data	Photographs	[1,3,12,17,25,27,28,30,32,35,45,59–62,72,73]
	Video frames	[4,12,17,26,27,33,41,62,74]
	Feature point clouds from image	[1,3,12,25,26,30,35,62]
	Laser scanned/depth image point clouds	[17,24,27,29,38,62]
	Patches	[12,30,32,60,62]
	Contours	[29]
As-planned data	3D models	All references
	4D models	[4,12,30,32,38,60,62]
	Logical/physical relationships	[24,25]
Alignment	3D–2D registration-based	[3,32,33,41]
	Feature point-based	[3,35,42,43,62]
	Depth image-based	[10,38]
	Perspective-based	[4,44,45]
Point cloud segmentation	Clustering-based	[47,48]
	Edge-based	[49,50]
	Region-based	[51]
	Graph-based	[50,52]
	Model fitting-based	[53,54]
	Hybrid	[55]
Point cloud semantic recognition	Point cloud semantic recognition	[56–59]
Progress reasoning	Based on the 3D space occupancy by the point clouds	[1,25,35]
	Based on the 2D plane projection of the point clouds	[10,29,38]
	Based on the image changes of 3D–2D projection area	[30,32,60–62]
	Based on the relationships of geometric primitives	[25,32,63,64]

References

1. Omar, H.; Mahdjoubi, L.; Kheder, G. Towards an automated photogrammetry-based approach for monitoring and controlling construction site activities. *Comput. Ind.* **2018**, *98*, 172–182. [CrossRef]
2. Salehi, S.A.; Yitmen, I. Modeling and analysis of the impact of BIM-based field data capturing technologies on automated construction progress monitoring. *Int. J. Civ. Eng.* **2018**, *16*, 1669–1685. [CrossRef]
3. Golparvar-Fard, M.; Peña-Mora, F.; Savarese, S. Application of D4AR–a 4-dimensional augmented reality model for automating construction progress monitoring data collection, processing and communication. *J. Inf. Technol. Constr.* **2009**, *14*, 129–153.
4. Kropp, C.; Koch, C.; König, M. Interior construction state recognition with 4D BIM registered image sequences. *Autom. Constr.* **2018**, *86*, 11–32. [CrossRef]
5. Czerniawski, T.; Leite, F. Automated digital modeling of existing buildings: A review of visual object recognition methods. *Autom. Constr.* **2020**, *113*, 103131. [CrossRef]
6. Han, X.-F.; Laga, H.; Bennamoun, M. Image-based 3D object reconstruction: State-of-the-art and trends in the deep learning era. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 1578–1604. [CrossRef] [PubMed]
7. Rankohi, S.; Waugh, L. Image-based modeling approaches for projects status comparison. In Proceedings of the CSCE 2014 General Conference, Halifax, NS, Canada, 28–31 May 2014; pp. 1–10.
8. Chu, C.-C.; Nandhakumar, N.; Aggarwal, J. Image segmentation using laser radar data. *Pattern Recognit.* **1990**, *23*, 569–581. [CrossRef]
9. Bechtold, S.; Höfle, B. Helios: A multi-purpose lidar simulation framework for research, planning and training of laser scanning operations with airborne, ground-based mobile and stationary platforms. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *3*, 161–168. [CrossRef]
10. Rebolj, D.; Pučko, Z.; Babič, N.Č.; Bizjak, M.; Mongus, D. Point cloud quality requirements for scan-vs-BIM based automated construction progress monitoring. *Autom. Constr.* **2017**, *84*, 323–334. [CrossRef]
11. Hong, S.; Park, I.; Lee, J.; Lim, K.; Choi, Y.; Sohn, H.-G. Utilization of a terrestrial laser scanner for the calibration of mobile mapping systems. *Sensors* **2017**, *17*, 474. [CrossRef]
12. Han, K.K.; Golparvar-Fard, M. Potential of big visual data and building information modeling for construction performance analytics: An exploratory study. *Autom. Constr.* **2017**, *73*, 184–198. [CrossRef]

13. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]
14. Bay, H.; Tuytelaars, T.; Van Gool, L. Surf: Speeded up robust features. In *Computer Vision—ECCV 2006, Proceedings of the 9th European Conference on Computer Vision, Graz, Austria, 7–13 May 2006*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 404–417. [CrossRef]
15. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395. [CrossRef]
16. Ullman, S. The interpretation of structure from motion. *Proc. R. Soc. London Ser. B Boil. Sci.* **1979**, *203*, 405–426. [CrossRef]
17. Fathi, H.; Dai, F.; Lourakis, M. Automated as-built 3D reconstruction of civil infrastructure using computer vision: Achievements, opportunities, and challenges. *Adv. Eng. Inform.* **2015**, *29*, 149–161. [CrossRef]
18. Pătrăucean, V.; Armeni, I.; Nahangi, M.; Yeung, J.; Brilakis, I.; Haas, C. State of research in automatic as-built modelling. *Adv. Eng. Inform.* **2015**, *29*, 162–171. [CrossRef]
19. Bilal, M.; Oyedele, L.O.; Qadir, J.; Munir, K.; Ajayi, S.O.; Akinade, O.; Owolabi, H.A.; Alaka, H.A.; Pasha, M. Big data in the construction industry: A review of present status, opportunities, and future trends. *Adv. Eng. Inform.* **2016**, *30*, 500–521. [CrossRef]
20. Omar, T.; Nehdi, M.L. Data acquisition technologies for construction progress tracking. *Autom. Constr.* **2016**, *70*, 143–155. [CrossRef]
21. El-Omari, S.; Moselhi, O. Data acquisition from construction sites for tracking purposes. *Eng. Constr. Arch. Manag.* **2009**, *16*, 490–503. [CrossRef]
22. Zhu, Z.; Brilakis, I. Comparison of optical sensor-based spatial data collection techniques for civil infrastructure modeling. *J. Comput. Civ. Eng.* **2009**, *23*, 170–177. [CrossRef]
23. Ekanayake, B.; Wong, J.K.-W.; Fini, A.A.F.; Smith, P. Computer vision-based interior construction progress monitoring: A literature review and future research directions. *Autom. Constr.* **2021**, *127*, 103705. [CrossRef]
24. Tang, P.; Huber, D.; Akinci, B.; Lipman, R.; Lytle, A. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Autom. Constr.* **2010**, *19*, 829–843. [CrossRef]
25. Braun, A.; Tuttas, S.; Borrmann, A.; Stilla, U. A concept for automated construction progress monitoring using BIM-based geometric constraints and photogrammetric point clouds. *J. Inf. Technol. Constr.* **2015**, *20*, 68–79.
26. Brilakis, I.; Fathi, H.; Rashidi, A. Progressive 3D reconstruction of infrastructure with videogrammetry. *Autom. Constr.* **2011**, *20*, 884–895. [CrossRef]
27. El-Omari, S.; Moselhi, O. Integrating automated data acquisition technologies for progress reporting of construction projects. *Autom. Constr.* **2011**, *20*, 699–705. [CrossRef]
28. Koch, C.; Paal, S.; Rashidi, A.; Zhu, Z.; König, M.; Brilakis, I. Achievements and challenges in machine vision-based inspection of large concrete structures. *Adv. Struct. Eng.* **2014**, *17*, 303–318. [CrossRef]
29. Volk, R.; Luu, T.H.; Mueller-Roemer, J.S.; Sevilmis, N.; Schultmann, F. Deconstruction project planning of existing buildings based on automated acquisition and reconstruction of building information. *Autom. Constr.* **2018**, *91*, 226–245. [CrossRef]
30. Han, K.; Golparvar-Fard, M. Appearance-based material classification for monitoring of operation-level construction progress using 4D BIM and site photologs. *Autom. Constr.* **2015**, *53*, 44–57. [CrossRef]
31. Xiao, D. Application and development of BIM Technology in spatial structure. Master's Thesis, Shanghai Jiaotong University, Shanghai, China, 2015.
32. Ibrahim, Y.; Lukins, T.; Zhang, X.; Trucco, E.; Kaka, A. Towards automated progress assessment of workpackage components in construction projects using computer vision. *Adv. Eng. Inform.* **2009**, *23*, 93–103. [CrossRef]
33. Golparvar-Fard, M.; Peña-Mora, F.; Arboleda, C.A.; Lee, S. Visualization of construction progress monitoring with 4D simulation model overlaid on time-lapsed photographs. *J. Comput. Civ. Eng.* **2009**, *23*, 391–404. [CrossRef]
34. Leung, S.-W.; Mak, S.; Lee, B.L. Using a real-time integrated communication system to monitor the progress and quality of construction works. *Autom. Constr.* **2008**, *17*, 749–757. [CrossRef]
35. Golparvar-Fard, M.; Peña-Mora, F.; Savarese, S. Automated progress monitoring using unordered daily construction photographs and IFC-based building information models. *J. Comput. Civ. Eng.* **2015**, *29*, 04014025. [CrossRef]
36. Ham, Y.; Han, K.K.; Lin, J.J.; Golparvar-Fard, M. Visual monitoring of civil infrastructure systems via camera-equipped unmanned aerial vehicles (UAVs): A review of related works. *Vis. Eng.* **2016**, *4*, 1. [CrossRef]
37. Álvares, J.S.; Costa, D.B. Literature review on visual construction progress monitoring using unmanned aerial vehicles. In Proceedings of the 26th Annual Conference of the International Group for Lean Construction, IGLC, Chennai, India, 16–22 July 2018; pp. 669–680.
38. Pučko, Z.; Šuman, N.; Rebolj, D. Automated continuous construction progress monitoring using multiple workplace real time 3D scans. *Adv. Eng. Inform.* **2018**, *38*, 27–40. [CrossRef]
39. Czerniawski, T.; Leite, F. 3D facilities: Annotated 3D reconstructions of building facilities. In *Transactions on Petri Nets and Other Models of Concurrency XV*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 10863, pp. 186–200.
40. Pour Rahimian, F.; Seyedzadeh, S.; Oliver, S.; Rodriguez, S.; Dawood, N. On-demand monitoring of construction projects through a game-like hybrid application of BIM and machine learning. *Autom. Constr.* **2020**, *110*, 103012. [CrossRef]
41. Kim, H.; Kano, N. Comparison of construction photograph and VR image in construction progress. *Autom. Constr.* **2008**, *17*, 137–143. [CrossRef]

42. Bueno, M.; Bosché, F.; González-Jorge, H.; Martínez-Sánchez, J.; Arias, P. 4-Plane congruent sets for automatic registration of as-is 3D point clouds with 3D BIM models. *Autom. Constr.* **2018**, *89*, 120–134. [CrossRef]

43. Lei, L.; Zhou, Y.; Luo, H.; Love, P.E. A CNN-based 3D patch registration approach for integrating sequential models in support of progress monitoring. *Adv. Eng. Inform.* **2019**, *41*, 100923. [CrossRef]

44. Asadi, K.; Ramshankar, H.; Noghabaei, M.; Han, K. Real-time image localization and registration with BIM Using perspective alignment for indoor monitoring of construction. *J. Comput. Civ. Eng.* **2019**, *33*, 04019031. [CrossRef]

45. Fernandez-Labrador, C.; Perez-Yus, A.; Lopez-Nicolas, G.; Guerrero, J.J. Layouts from panoramic images with geometry and deep learning. *IEEE Robot. Autom. Lett.* **2018**, *3*, 3153–3160. [CrossRef]

46. Wang, Q.; Tan, Y.; Mei, Z. Computational methods of acquisition and processing of 3D point cloud data for construction applications. *Arch. Comput. Methods Eng.* **2019**, *27*, 479–499. [CrossRef]

47. Lu, X.; Yao, J.; Tu, J.; Li, K.; Li, L.; Liu, Y. Pairwise linkage for point cloud segmentation. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *3*, 201–208. [CrossRef]

48. Zeng, S.; Chen, J.; Cho, Y.K. User exemplar-based building element retrieval from raw point clouds using deep point-level features. *Autom. Constr.* **2020**, *114*, 103159. [CrossRef]

49. Wani, M.A.; Arabnia, H.R. Parallel edge-region-based segmentation algorithm targeted at reconfigurable multiring network. *J. Supercomput.* **2003**, *25*, 43–62. [CrossRef]

50. Chen, J.; Kira, Z.; Cho, Y.K. Deep learning approach to point cloud scene understanding for automated scan to 3D reconstruction. *J. Comput. Civ. Eng.* **2019**, *33*, 04019027. [CrossRef]

51. Vo, A.-V.; Truong-Hong, L.; Laefer, D.F.; Bertolotto, M. Octree-based region growing for point cloud segmentation. *ISPRS J. Photogramm. Remote Sens.* **2015**, *104*, 88–100. [CrossRef]

52. Strom, J.; Richardson, A.; Olson, E. Graph-based segmentation for colored 3D laser point clouds. In Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 18–22 October 2010; pp. 2131–2136.

53. Maas, H.-G.; Vosselman, G. Two algorithms for extracting building models from raw laser altimetry data. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 153–163. [CrossRef]

54. Chen, D.; Zhang, L.; Mathiopoulos, P.T.; Huang, X. A methodology for automated segmentation and reconstruction of urban 3-D buildings from ALS point clouds. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4199–4217. [CrossRef]

55. Vieira, M.; Shimada, K. Surface mesh segmentation and smooth surface extraction through region growing. *Comput. Aided Geom. Des.* **2005**, *22*, 771–792. [CrossRef]

56. Antonello, M.; Wolf, D.; Prankl, J.; Ghidoni, S.; Menegatti, E.; Vincze, M. Multi-view 3D entangled forest for semantic segmentation and mapping. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Los Alamitos, CA, USA, 21–26 May 2018; pp. 1855–1862.

57. Posada, L.F.; Velasquez-Lopez, A.; Hoffmann, F.; Bertram, T. Semantic mapping with omnidirectional vision. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Los Alamitos, CA, USA, 21–26 May 2018; pp. 1901–1907.

58. Adán, A.; Quintana, B.; Prieto, S.; Bosché, F. An autonomous robotic platform for automatic extraction of detailed semantic models of buildings. *Autom. Constr.* **2020**, *109*, 102963. [CrossRef]

59. Dimitrov, A.; Golparvar-Fard, M. Vision-based material recognition for automated monitoring of construction progress and generating building information modeling from unordered site image collections. *Adv. Eng. Inform.* **2014**, *28*, 37–49. [CrossRef]

60. Kim, C.; Kim, B.; Kim, H. 4D CAD model updating using image processing-based construction progress monitoring. *Autom. Constr.* **2013**, *35*, 44–52. [CrossRef]

61. Zhu, Z.; Brilakis, I. Parameter optimization for automated concrete detection in image data. *Autom. Constr.* **2010**, *19*, 944–953. [CrossRef]

62. Han, K.; DeGol, J.; Golparvar-Fard, M. Geometry- and appearance-based reasoning of construction progress monitoring. *J. Constr. Eng. Manag.* **2018**, *144*, 04017110. [CrossRef]

63. Nuechter, A.; Hertzberg, J. Towards semantic maps for mobile robots. *Robot. Auton. Syst.* **2008**, *56*, 915–926. [CrossRef]

64. Nguyen, T.-H.; Oloufa, A.A.; Nassar, K. Algorithms for automated deduction of topological information. *Autom. Constr.* **2005**, *14*, 59–70. [CrossRef]

65. Chen, C.; Yang, B. Dynamic occlusion detection and inpainting of in situ captured terrestrial laser scanning point clouds sequence. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 90–107. [CrossRef]

66. Hou, X.; Zeng, Y.; Xue, J. Detecting structural components of building engineering based on deep-learning method. *J. Constr. Eng. Manag.* **2020**, *146*, 04019097. [CrossRef]

67. Roh, S.; Aziz, Z.; Peña-Mora, F. An object-based 3D walk-through model for interior construction progress monitoring. *Autom. Constr.* **2011**, *20*, 66–75. [CrossRef]

68. Kropp, C.; Koch, C.; Brilakis, I.; Knig, M. A framework for automated delay prediction of finishing works using video data and BIM-based construction simulation. In Proceedings of the 14th International Conference on Computing in Civil and Building Engineering (ICCCBE), Moscow, Russia, 27–29 June 2012.

69. Hamledari, H.; McCabe, B.; Davari, S. Automated computer vision-based detection of components of under-construction indoor partitions. *Autom. Constr.* **2017**, *74*, 78–94. [CrossRef]

70. Bortoluzzi, B.; Efremov, I.; Medina, C.; Sobieraj, D.; McArthur, J. Automating the creation of building information models for existing buildings. *Autom. Constr.* **2019**, *105*, 102838. [CrossRef]

71. Li, Y.; Li, W.; Tang, S.; Darwish, W.; Hu, Y.; Chen, W. Automatic indoor as-built building information models generation by using low-cost rgb-d sensors. *Sensors* **2020**, *20*, 293. [CrossRef]

72. Galantucci, R.A.; Fatiguso, F. Advanced damage detection techniques in historical buildings using digital photogrammetry and 3D surface anlysis. *J. Cult. Herit.* **2019**, *36*, 51–62. [CrossRef]

73. Koch, C.; Georgieva, K.; Kasireddy, V.; Akinci, B.; Fieguth, P. A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Adv. Eng. Inform.* **2015**, *29*, 196–210. [CrossRef]

74. Fang, Q.; Li, H.; Luo, X.; Ding, L.; Rose, T.; An, W.; Yu, Y. A deep learning-based method for detecting non-certified work on construction sites. *Adv. Eng. Inform.* **2018**, *35*, 56–68. [CrossRef]

75. Almasi, M. An investigation on deep learning applications for 3D reconstruction of human movements. *Invent. J. Res. Technol. Eng. Manag.* **2020**, *4*, 1–8.

76. Zhu, Z.; Ren, X.; Chen, Z. Integrated detection and tracking of workforce and equipment from construction jobsite videos. *Autom. Constr.* **2017**, *81*, 161–171. [CrossRef]

77. Wu, H.; Zhao, J. An intelligent vision-based approach for helmet identification for work safety. *Comput. Ind.* **2018**, *100*, 267–277. [CrossRef]

78. Fang, Q.; Li, H.; Luo, X.; Ding, L.; Luo, H.; Rose, T.M.; An, W. Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Autom. Constr.* **2018**, *85*, 1–9. [CrossRef]

79. Deng, H.; Hong, H.; Luo, D.; Deng, Y.; Su, C. Automatic indoor construction process monitoring for tiles based on BIM and computer vision. *J. Constr. Eng. Manag.* **2020**, *146*, 04019095. [CrossRef]

80. Zhao, C.; Sun, L.; Stolkin, R. A fully end-to-end deep learning approach for real-time simultaneous 3D reconstruction and material recognition. In Proceedings of the 2017 18th International Conference on Advanced Robotics (ICAR), Hong Kong, China, 15 April 2017; pp. 75–82.

81. Landrieu, L.; Simonovsky, M. Large-scale point cloud semantic segmentation with superpoint graphs. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Institute of Electrical and Electronics Engineers (IEEE), Salt Lake, UT, USA, 18–22 June 2018; pp. 4558–4567.

82. Acharya, D.; Khoshelham, K.; Winter, S. BIM-PoseNet: Indoor camera localisation using a 3D indoor model and deep learning from synthetic images. *ISPRS J. Photogramm. Remote Sens.* **2019**, *150*, 245–258. [CrossRef]