Article

An Application Oriented Scan-to-BIM Framework

Qian Wang [1],*, Jingjing Guo [1] and Min-Koo Kim [2]

[1] Department of Building, School of Design and Environment, National University of Singapore, Singapore 117566, Singapore; bdggj@nus.edu.sg
[2] Department of Building and Real Estate, The Hong Kong Polytechnic University, Hong Kong 999077, China; minkoo.kim@polyu.edu.hk
* Correspondence: bdgwang@nus.edu.sg

check for
updates

Abstract: Building information modelling (BIM) has been adopted in the construction industry. The success of BIM implementation relies on the accurate building information stored in BIM models. However, building information in BIM models can be inaccurate, out-of-date, or missing in real-world projects. 3D laser scanning has been leveraged to capture the accurate as-is conditions of buildings and create as-is BIM models of buildings; this is known as the scan-to-BIM process. Although industry practitioners and researchers have implemented and studied the scan-to-BIM process, there is no framework that systematically defines and discusses the key steps and considerations in the process. This study proposes an application-oriented framework for scan-to-BIM, which describes the four major steps of a scan-to-BIM process and their relationships. The framework is oriented towards the specific BIM application to be implemented using the created as-is BIM, and includes four steps: (1) identification of information requirements, (2) determination of required scan data quality, (3) scan data acquisition, and (4) as-is BIM reconstruction. Two illustrative examples are provided to demonstrate the feasibility of the proposed scan-to-BIM framework. Furthermore, future research directions within the scan-to-BIM framework are suggested.

Keywords: 3D laser scanning; building information modelling (BIM); scan-to-BIM; information requirements; scan data quality; scanning parameters; as-is BIM reconstruction

1. Introduction

Building information modelling (BIM) has been popularly adopted in the architecture, engineering, construction and facilities management (AEC/FM) industries in recent decades. BIM serves as a knowledge repository that generates and manages building information throughout a facility lifecycle, from inception to demolition. The rich building information available in BIM models and their 3D visualization ability enable improved design, construction, and operation and maintenance (O&M) of buildings. Accurate building information in BIM models is the key factor to successful BIM implementation. However, building information stored in BIM models is often inaccurate, out-of-date, or missing. For buildings in the design and construction phases, BIM models often contain the as-designed building information rather than the updated as-is information because the as-designed BIM models are not updated when design changes are made. Similarly, for buildings in the O&M phase, the BIM models may contain the as-designed or as-built building information rather than the as-is information if the BIM models are not updated when addition and alteration works are carried out. In addition to inaccurate and out-of-date building information, most of the existing buildings do not have any BIM model as the popularity of BIM only began in the recent decade. It is impossible to implement BIM applications for existing buildings in the O&M phase when the BIM model is missing.

To resolve the problems arising from inaccurate, out-of-date, or missing building information, reality capture technologies such as 3D laser scanning have been widely adopted in the AEC/FM industry. 3D laser scanning is also known as light detection and ranging (LiDAR), which measures the distance to a target by emitting a laser beam and detecting the reflected laser signal. As the scanner emits laser beams and measures distances in different directions with the aid of a rotating mirror, eventually a complete point cloud containing distance measurements of the whole scene is obtained. Compared to traditional surveying tools, 3D laser scanning is advantageous in its high measurement accuracy (millimeter level) and fast measurement speed (hundred thousand points per second) [1]. As the obtained laser scan data represent the as-is geometries of existing facilities, as-is BIM models of the facilities can be created based on the laser scan data.

The process of transferring laser scan data into BIM models is known as scan-to-BIM [2]. The common practice in the industry is to manually feed laser scan data into BIM authoring tools (e.g., Autodesk Revit, Bentley, Graphisoft ArchiCAD) and then create BIM models with reference to the laser scan data. However, the manual process is time-consuming and error prone especially for large-scale projects with numerous building elements and complex geometries. Hence, researchers have developed semi-automated or automated scan-to-BIM techniques to replace the manual approach [3–6]. These developed techniques are focused on the automation of object recognition and geometric modelling of building components from laser scan data. Detailed reviews and discussions of the relevant techniques for automated scan-to-BIM techniques can be found in [1,7,8].

Although industry practitioners have adopted scan-to-BIM to generate BIM models, the scan-to-BIM process is mostly carried out based on the experiences of engineers. The scan-to-BIM process has not been well planned such that the BIM model is created with the minimum time and effort and is able to fulfill the requirements of the intended BIM application. On the other hand, researchers have been studying the different steps within the scan-to-BIM process to accelerate and optimize the scan-to-BIM process. Nevertheless, there have been limited efforts in systematically defining and discussing the key steps and considerations in the scan-to-BIM process as a framework for both research and practical implementation.

To tackle these limitations, this study aims to provide an application oriented framework for scan-to-BIM. The proposed framework initiates the scan-to-BIM process by determining the BIM application to be implemented after creating the as-is BIM model. Then, the information requirements for the BIM application are identified and the required scan data quality for satisfying the information requirements is determined. Afterwards, the scan data are obtained with pre-defined scanning parameters that fulfill the required scan data quality. Lastly, as-is BIM reconstruction is carried out to create an as-is BIM containing all the required information. The proposed framework is expected to provide a guideline to the industry for better planning and execution of the scan-to-BIM process such that the as-is BIM is created with the minimum time and effort and is able to fulfill the needs of the intended BIM application. Additionally, this study suggests future research directions within the scan-to-BIM framework for researchers in this field. The scientific contributions of this study include (1) proposing a holistic application-oriented scan-to-BIM framework, (2) establishing relationships between different steps in the scan-to-BIM process and proposing methodological frameworks to implement the process based on the intended BIM application, and (3) suggesting future research directions within the scan-to-BIM framework.

2. Applications of As-Is BIM

The as-is BIM models generated from scan-to-BIM have been adopted for various applications and have potential for a wider range of applications in the future. According to different project phases (i.e., design, construction, O&M), the current and potential applications of as-is BIM are summarized in Table 1 and illustrated in the following subsections.

Table 1. Applications of as-is building information modelling (BIM) in different project phases.

Phase	Application	References
Design	Design with 3D site model	Aydin [9]
Construction	QA/QC	Wang et al. [10], Bosché and Guenet [11], Tang et al. [12], Bosché [13]
	Construction progress tracking	Kim et al. [14], Son and Kim [15], Turkan et al. [16]
	Virtual installation and assembly	Nahangi et al. [17], Rausch et al. [18]
	Safety management	Zhang et al. [19], Wang et al. [20]
	Digital reproduction	Xu et al. [21]
O&M	Documentation	Yastikli [22], Rüther et al. [23]
	Building performance analysis (e.g., energy performance analysis, accessibility diagnosis, structural analysis)	Ham and Golparvar-Fard [24], Balado et al. [25], Riveiro et al. [26]
	FM functionalities (e.g., operation and maintenance, space management, renovation planning and execution, emergency management, and personnel training and development)	Becerik-Gerber et al. [27]

2.1. Design Phase

In the design phase, the as-is BIM models of the construction site terrain and the surrounding buildings and environment enable the designers to better understand the site conditions and to make better decisions about the building design. For example, Aydin [9] leveraged laser scan data of existing buildings to evaluate a building façade design with regard to the visual harmony, aesthetics, design compatibility, architectural features, etc.

2.2. Construction Phase

In the construction phase, the as-is BIM models of on-site construction works or prefabricated components are widely adopted for quality assessment/quality control (QA/QC). The as-is BIM is often compared to the as-designed BIM to identify any discrepancy between them, which is further compared against the tolerance values specified in relevant codes and regulations. Research has been reported on different types of QA/QC, including dimensional discrepancy [10], surface flatness and distortion [11], surface spalling defects [12], and positioning error [13].

Progress tracking is another popular application of as-is BIM in the construction phase [14–16]. By scanning the on-site construction works at a certain project milestone, the as-is BIM models are created and compared to the as-designed 4D BIM models containing project schedule information. Therefore, the actual project schedule is compared to the planned schedule in the BIM environment to track the construction progress and identify construction works that are behind schedule.

Virtual installation and assembly with accurate as-is BIM models of construction assemblies becomes possible with scan-to-BIM [17,18]. For prefabricated building components, scan-to-BIM can be adopted to create as-is BIM models to simulate the installation and assembly process in a virtual environment. With virtual installation and assembly, any potential difficulty or problem can be identified before the actual installation and assembly, such as incorrect component sizes or clashes between the prefabricated component and the other construction works, leading to construction time and cost savings.

As-is BIM also has a potential to enhance construction safety management. Zhang et al. [19] developed a rule-based engine for automated safety checking that could identify the safety measures needed for preventing fall-related accidents before the actual construction. Wang et al. [20] analysed the laser scan data of excavated pits to identify and localize fall hazards and then applied the previously developed rule checking system to suggest the installation of safety equipment. Although the previous studies did not

implement scan-to-BIM, it is a potential research direction to generate as-is BIM models for buildings and construction sites with scan-to-BIM for the identification of safety hazards and safety measure planning.

Furthermore, scan-to-BIM can also be used for digital reproduction of construction components [21]. The generated as-is BIM models of a component can be fed into a construction 3D printer such as a concrete printer to reproduce the component with high accuracy. Reproduction with scan-to-BIM is particularly useful for complex-geometry building components without any blueprint or 3D model, for example, as components of historical buildings.

2.3. O&M Phase

Documentation of building geometries and textures is the most fundamental application of scan-to-BIM in the O&M phase. It is popularly adopted for the preservation of historical heritage such as the Fatih Mosque in Istanbul [22] and the Wonderwerk Cave in South Africa [23]. BIM-based documentation enables better preservation of historical heritage and facilitates their maintenance and repair.

The as-is BIM from scan-to-BIM has also been adopted for building performance analysis including energy performance analysis [24], accessibility diagnosis [25], and structural analysis [26]. Scan-to-BIM provides accurate as-is BIM models such that BIM-based building performance analysis can be extensively carried out to study and improve building performance as regards its energy consumption, accessibility, and structural reliability. In the future, it is expected that scan-to-BIM can be applied for many other analyses and simulations such as emergency evacuation simulation, spatial program validation, daylight simulation, etc.

As-is BIM models of existing buildings are also beneficial to multiple FM functionalities due to the superior 3D visualization ability and the rich and well-organized building information in BIM. Potential applications of BIM-enabled FM include O&M, space management, renovation planning and execution, emergency management, and personnel training and development [27]. For example, the O&M of building components can be improved with a BIM-based process because BIM enables efficient localization of building components and real-time access to the relevant facility data, thereby reducing the cost and time for maintenance work [27]. To facilitate the implementation of BIM-enabled FM for existing buildings without BIM models, scan-to-BIM has great potential to be adopted for generating as-is BIM models for these existing buildings [7].

3. Proposed Scan-to-BIM Framework

As illustrated in Figure 1, the proposed scan-to-BIM framework is oriented towards a BIM application that is implemented using the generated as-is BIM. To ensure that the generated as-is BIM satisfies the requirements of the intended BIM application, the scan-to-BIM process is performed in four steps including (1) identification of information requirements, (2) determination of required scan data quality, (3) scan data acquisition, and (4) as-is BIM reconstruction. The four steps and their relationships are further explained in the following subsections.
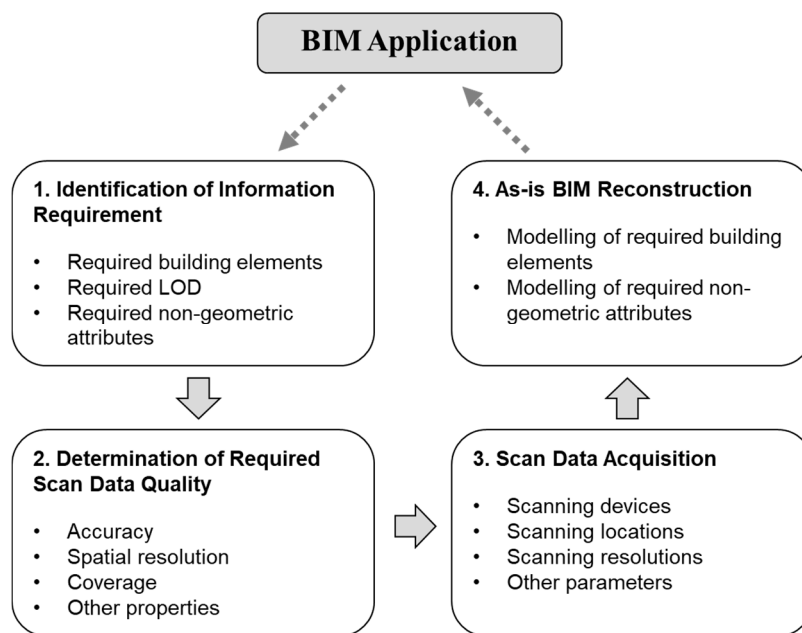
Figure 1. The proposed application oriented scan-to-BIM framework.

3.1. Identification of Information Requirements

Different BIM applications require different sets of information to be stored in the as-is BIM models. Collecting too much information is a waste of time and effort while collecting insufficient information will render a BIM model that is useless for the intended application [1]. Therefore, it is necessary to identify the required information for each specific BIM application before collecting scan data. The identification of the information requirements consists of three aspects: (1) required building elements, (2) required level of detail (LOD), and (3) required non-geometric attributes.

3.1.1. Required Building Elements

The required building elements refer to the building elements that should be modelled in the as-is BIM. A list of required building elements should be identified using a certain building element classification system. The OmniClass Construction Classification System (known as OmniClass or OCCS) [28] is one example, which has been popularly adopted in the construction industry. The OmniClass consists of 15 tables representing different types of construction information and it has been adopted in some research efforts. For example, Table 21 (representing construction Elements) of OmniClass was adopted to define information exchange by Messner et al. [29] and Table 23 (representing construction Products) was adopted to define facility assets to be tracked through BIM by Pishdad-Bozorgi et al. [30]. Besides OmniClass, some other well-known classification systems include the MasterFormat [31] from the U.S. and Canada and the UniClass from the U.K. [32].

3.1.2. Required LOD

For each building element that needs to be modelled, the required modelling LOD of the element should be defined. LOD refers to "level of detail" in this paper, although this term is often used equally with another term "level of development" in the construction industry. The term "level of development" describes the progressive development of the BIM model from the conceptual design stage onwards until the completion of construction. On the other hand, the term "level of detail" is rather to describe the degree to which the element model is detailed and specific. The two terms are mostly utilized in the development of BIM models for new buildings, in which the two terms describe the same model development process but from different perspectives. Various model LOD specifications have been proposed and adopted in the industry and a typical LOD specification consists of five levels including

conceptual (LOD 100), approximate geometry (LOD 200), precise geometry (LOD 300), fabrication (LOD 400) and as-built (LOD 500) [33,34]. Table 2 shows the detailed descriptions of the typical five levels of model LOD specification.

Table 2. The typical five levels of model level of detail (LOD) [33,34].

LOD	Description	Model
100	Conceptual	A symbol or other generic representation
200	Approximate geometry	Approximate quantity, size, shape, location, and orientation
300	Precise geometry	Specific quantity, size, shape, location, and orientation
400	Fabrication	Specific size, shape, location, quantity, and orientation with detailing, fabrication, assembly, and installation information
500	As-built	Field verified size, shape, location, quantity, and orientation

Because this study is focused on the creation of as-is BIM for existing buildings, the existing LOD specifications cannot be directly adopted in this study. One major issue is that all the as-is BIM models should be LOD 500 due to the "as-built" nature according to the existing LOD specifications. In addition, the existing LOD specifications do not clearly define the quantitative criterion for each level. t The distinction between approximate geometry and precise geometry is rather vague and subjective. Therefore, in the proposed scan-to-BIM framework, the model LOD consists of three levels, namely "conceptual", "approximate geometry", and "precise geometry". Wherever applicable, the required modelling accuracy of the element quantity, size, shape, location, and orientation should be defined as supplementary to the required LOD. For example, the required LOD for a wall can be "precise geometry" with 10 mm modelling accuracy regarding its size, shape, and location.

3.1.3. Required Non-Geometric Attributes

While the model LOD is focused on the required geometric attributes of an element, the required non-geometric attributes should also be defined for each element to fulfil the needs of the intended BIM application. For example, for BIM-based spatial reasoning for the diagnosis of the water distribution system of a building, the spatial relationships between building elements are essential [1]. These spatial relationships include aggregation relationships, topological relationships, and directional relationships. For BIM-based building energy simulation, the materials of building elements and thermal properties of building materials are necessary. Some other non-geometric attributes include surface properties (e.g., colours and reflectivity) of building elements, mechanical properties of structural elements, acoustic properties of walls, technical specification and warranty information of equipment, etc. It is worth noting that the scan-to-BIM process can collect only some of the above-mentioned non-geometric attributes such as the spatial relationships between building elements, materials of building elements, and surface properties of building elements. However, it is still useful to identify all the required non-geometric attributes in this step so that the requirements are well documented and other data sources may be deployed in a later stage to provide the required attributes.

3.1.4. Sample of Identified Information Requirements

Figure 2 shows a sample result of the identification of information requirements. A set of required building elements are listed in the table with their respective OmniClass ID. For each building element, the required LOD with a specific modelling accuracy requirement is provided. Some elements require a modelling accuracy of 20 mm in size, shape, and location while some other elements require a lower accuracy of 50 mm. The required non-geometric attributes are also specified for some building elements. The floor, roof, and stairs need their respective material information, and exterior walls need the fire rating information.

OmniClass ID	Element	Required LOD (Accuracy)	Non-geometric Attributes
21-02-10-10	Floor Construction	Approximate geometry (50 mm for size, shape, and location)	Material
21-02-10-20	Roof Construction	Approximate geometry (50 mm for size, shape, and location)	Material
21-02-10-80	Stairs	Approximate geometry (50 mm for size, shape, and location)	Material
21-02-20-10	Exterior Walls	Precise geometry (20 mm for size, shape, and location)	Fire rating
21-02-20-20	Exterior Windows	Precise geometry (20 mm for size, shape, and location)	
21-02-20-50	Exterior Doors and Grilles	Precise geometry (20 mm for size, shape, and location)	

Figure 2. Sample result of the identification of information requirements.

3.2. Determination of Required Scan Data Quality

3.2.1. Parameters for Scan Data Quality

After the identification of the required information, the required scan data quality that can fulfil these information requirements should be determined. The parameters for the scan data or point cloud data quality have been discussed in previous research and industry guidelines [35]. For example, the U.S. General Services Administration (GSA) BIM Guide for 3D Imaging defined four different levels of detail for point cloud data. Each level of detail corresponds to a specific criteria on the allowable dimensional deviation and the dimensions of the smallest recognizable feature [36]. Dai et al. [37] assessed point cloud data quality using three parameters: accuracy, completeness, and density. Here, the completeness was measured as the percentage of coverage of point cloud data, which was calculated as the ratio between the number of covered regions and the total number of regions. In addition, Rebolj et al. [35] adopted the minimum local density, the minimum local accuracy, and level of scatter as parameters for measuring the quality of point cloud data.

In the proposed scan-to-BIM framework, four parameters were adopted to measure scan data quality including (1) accuracy, (2) spatial resolution, (3) coverage, and (4) other properties, as illustrated in Figure 3. The accuracy of the scan data refers to the ranging accuracy of each laser scan data point. As shown in Figure 3a, the ranging error e of a scan data point is measured as the distance between the correct measurement (point A) and the actual measurement (point B). The spatial resolution of scan data is measured as the distance between two adjacent laser scan points. For terrestrial laser scan data, the spatial resolution should be calculated at two perpendicular directions (i.e., horizontal and vertical) as s_h and s_v, respectively, as shown in Figure 3b. For mobile laser scan data, other methods should be deployed to calculate the spatial resolution such as by calculating the average distance from one scan point to its nearest eight scan points. The coverage of laser scan data is measured as the ratio between the area covered by laser scan data and the total area of interest, i.e., $A_{covered}/A_{total}$ shown in Figure 3c. Note that only if the area is covered by laser scan data with the required ranging accuracy and spatial resolution, can the area be counted as being covered. This is because the fulfilment of coverage is meaningless when the scan data do not have sufficient ranging accuracy or spatial resolution. To calculate the coverage, the total area of interest may be divided into many small regions with horizontal and vertical grids, and a small region is deemed as covered if there exists at least one scan point within this region. If multiple scans are conducted, a small region is deemed as covered if at least one scan provides data points with the required ranging accuracy and spatial resolution at this region. The grid size should be close to the required scan data spatial resolution. The parameter of coverage is particularly important when occlusions exist among building elements so that the area of

interest cannot be fully covered by the scan data. Last, other properties of scan data refer to properties other than the XYZ coordinates, such as the reflectivity value and red, green, blue (RGB) colour values, as shown in Figure 3d.
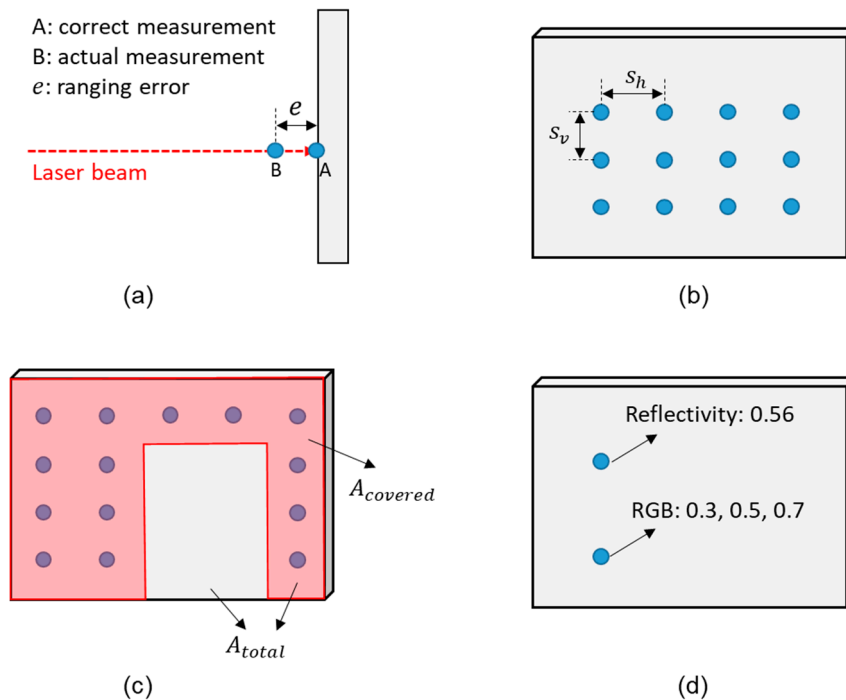

Figure 3. Four parameters to measure scan data quality: (**a**) accuracy is measured as the ranging accuracy of each scan point, (**b**) spatial resolution is measured as the distance between two adjacent scan points, (**c**) coverage is measured as the ratio between covered area and total area, and (**d**) other properties such as reflectivity value and red, green, blue (RGB) colour values.

3.2.2. Determination of Required Scan Data Quality

The required scan data quality is mainly determined by the information requirements and their relationships are depicted in Figure 4. The required accuracy of laser scan data is determined by the required model LOD and modelling accuracy. Assuming that the range measurement of a scan point has an error of 5 mm, it will result in an error of at least 5 mm for the location of the modelled element. Therefore, the accuracy of scan data must be equal to or higher than the required modelling accuracy. The required spatial resolution of laser scan data is also determined by the required model LOD and modelling accuracy. With a scan data resolution of s, the resulting size of the modelled element can have an error of more than $2s$ (s for each edge) due to the problem of "edge loss" [38]. Even though the proposed edge loss compensation model in Tang et al. [38] is implemented, the size error is still up to s ($s/2$ for each edge). As a result, the spatial resolution of scan data must be equal to or higher than the required modelling accuracy. The required coverage of scan data depends on the building elements and their required LOD. For elements with simple shapes such as planar walls, even though a part of the element is not covered by laser scan data, its shape, size, and location can be inferred from the covered part. In this case, the required coverage can be relatively low. On the other hand, for elements with irregular shapes and numerous details such as MEP elements, a higher coverage is required to achieve the desired model LOD. In general, when the required model LOD is higher, the required coverage should be higher to capture more details of the building elements. Lastly, the other properties of scan data mainly depend on the required non-geometric attributes. For example, when the RGB colours of an element are identified as required non-geometric attributes, the laser scan data should contain such information accordingly.
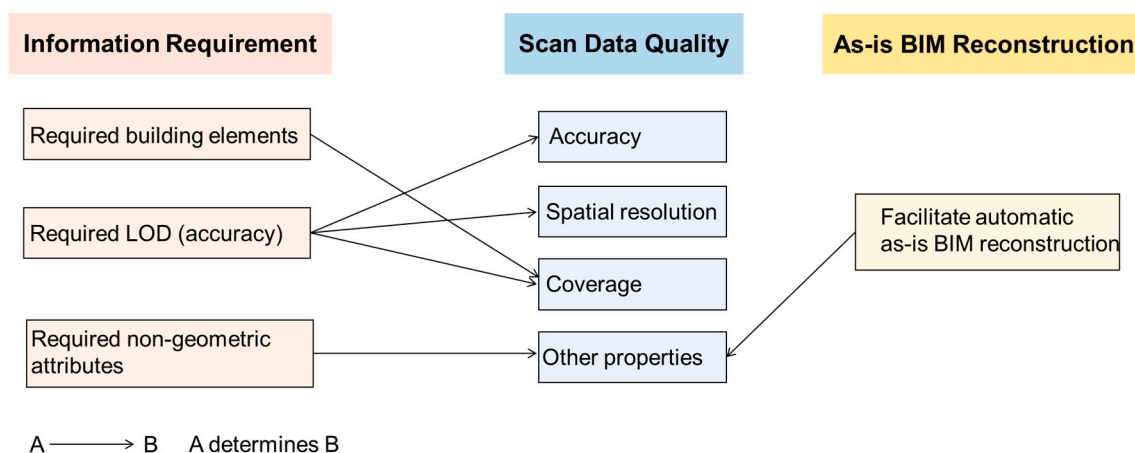
Information Requirement　　　**Scan Data Quality**　　　**As-is BIM Reconstruction**

Required building elements

Required LOD (accuracy)

Required non-geometric attributes

Accuracy

Spatial resolution

Coverage

Other properties

Facilitate automatic as-is BIM reconstruction

A ⟶ B　A determines B

Figure 4. Determination of required scan data quality based on information requirements and needs for automatic as-is BIM reconstruction.

The required scan data quality should also facilitate as-is BIM reconstruction, especially when semi-automatic or automatic as-is BIM reconstruction techniques are adopted. For example, RGB colour values of scan data are able to improve the accuracy of object recognition and classification from laser scan data. In this case, even though the required non-geometric attributes do not include RGB colours, it is preferred to require RGB colours for the scan data.

3.3. Scan Data Acquisition

3.3.1. Determination of Scanning Parameters

Before scan data acquisition, the scanning parameters including scanning devices, scanning locations, angular resolutions, and other parameters are determined. The scanning parameters are determined to satisfy the required scan data quality as illustrated in Figure 5. However, it is worth noting that the scanner may be given and even the scanning locations are limited in most practical cases.
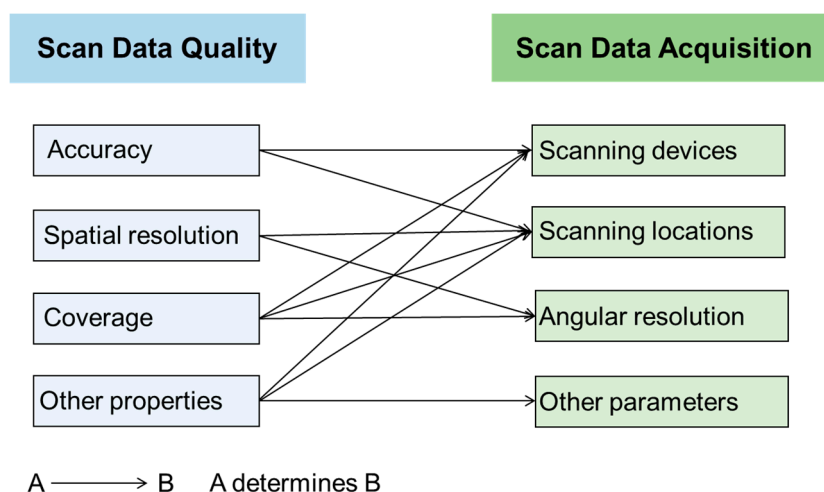
Scan Data Quality　　　　　　　**Scan Data Acquisition**

Accuracy

Spatial resolution

Coverage

Other properties

Scanning devices

Scanning locations

Angular resolution

Other parameters

A ⟶ B　A determines B

Figure 5. Determination of scanning parameters based on required scan data quality.

The accuracy of scan data is related to the scanning devices and scanning locations. Different scanning devices have different ranging accuracy depending on the hardware performance. For terrestrial laser scanners, the ranging accuracy is usually between 1 mm and 10 mm. For the same laser scanner, the accuracy is also affected by the scanning locations, especially the distance L from the scanner to the target object and incident angle α of the laser beam with respect to the normal direction

of the target object, as shown in Figure 6. For long distances, the longer the distance *L* and the larger the incident angle *α*, the lower the ranging accuracy [39]. On the other hand, too short distances (e.g., below ~0.5 m) cannot be properly measured as well, which can cause problems in narrow spaces such as factory-like environments. Many other factors also affect the ranging accuracy including the surface reflectivity of the target object, the surface smoothness of the target object, atmosphere humidity and temperature, and ambient light [37]. Due to the large number of factors that influence ranging accuracy, simplification is usually needed for practical consideration. For example, Biswas et al. [40] considered two major factors, distance *L* and incident angle *α*, and set up a maximum *L* and a maximum *α* as the criteria for the required scan data accuracy. Wang et al. [4] simply set up a maximum incident angle *α* as the criterion considering that the adopted laser scanner always had a high accuracy within short scanning distances.
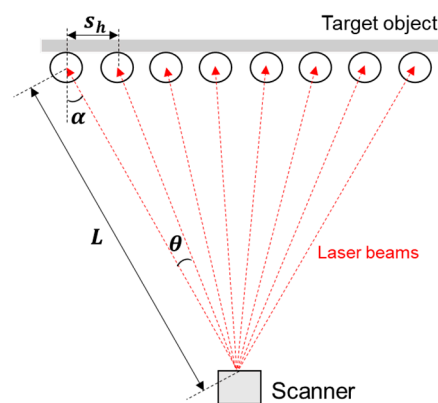

Figure 6. Geometric relationship between scanning parameters and scan data quality.

The spatial resolution of scan data is related to the scanning location and angular resolution *θ*. As shown in Figure 6, a terrestrial laser scanner emits laser beams in different directions with an angle increment of *θ*. Hence, the horizontal spatial resolution is obtained as:

$$s_h = \theta L / \cos\alpha \qquad (1)$$

The vertical spatial resolution s_v can be calculated similarly with the incident angle in a vertical direction.

The coverage of scan data is related to scanning devices, locations and angular resolutions. The covered area must be covered by scan data with the required accuracy and spatial resolution, which are fundamentally determined by scanning devices, locations, and angular resolutions.

Last, other properties of scan data determine scanning devices, scanning locations, and other scanning parameters. In order to capture other properties, the scanning devices must include the necessary sensors such as a camera to capture RGB colours. Furthermore, the scanning locations must enable a line of sight from the scanner to the target to capture such properties. Lastly, the users often have the options to enable or disable a sensor by setting scanning parameters.

3.3.2. Optimization of Scanning Parameters

In practice, simply having more scanning locations with shorter distances, lower incident angles and higher angular resolutions can easily satisfy the required scan data accuracy, spatial resolution, and coverage. However, such a scanning process is time-consuming and generates redundant scan data. Therefore, studies have been conducted to find the optimal scanning parameters with the shortest scanning time or the minimum number of scans [4,40–43] while also satisfying the required scan data quality. Since there are an infinite number of possible combinations of scanning parameters, it is impractical to search for the optimal scanning parameters using a brute-force

approach. Hence, more efficient algorithms have been proposed in previous studies. For example, the "divide-and-conquer" method was proposed by Zhang et al. [43], which first clusters target points that need to be scanned and then finds the scanning locations for each cluster individually. The greedy algorithm was adopted by Wang et al. [4], which always chooses the scanning location that brings the most new scan points with the required accuracy and spatial resolution.

3.4. As-Is BIM Reconstruction

After scan data acquisition, the as-is BIM is reconstructed from the scan data. The reconstructed as-is BIM must address all information requirements identified in the first step. Therefore, the as-is BIM reconstruction includes (1) modelling of required building elements with required LOD, and (2) modelling of required non-geometric attributes. To replace the manual as-is BIM reconstruction process, automated as-is BIM reconstruction techniques have been reported by researchers and detailed reviews of these techniques can be found in [1,44,45]. In the following subsections, automated as-is BIM reconstruction techniques are discussed with respect to the scan-to-BIM framework.

3.4.1. Modelling of Required Building Elements

Automated modelling of required building elements includes two major tasks, namely, geometric modelling and object recognition. Geometric modelling aims to model the geometry of a building element from laser scan data, such as detecting planes from laser scan data or modelling parametric surfaces. On the other hand, object recognition aims to label scan data into object categories such as wall, floor, roof, etc. It is worth noting that these two tasks are not necessarily carried out sequentially. The strategies and algorithms to model building elements vary substantially for different scenarios. The three most common scenarios are described and their respective modelling approaches are recommended in the following.

Scenario 1: An as-designed BIM model exists. In this scenario, the scan-vs.-BIM approach can be applied in the following three steps [13,46–48]. First, the scan data are aligned with the as-designed BIM model of the building. This can be done by manually selecting at least three pairs of corresponding points or by a two-step registration process including a coarse registration and a fine registration. In addition, the geodetic coordinate system applied during construction and scanning can also be used for matching the scan datasets. Second, each laser scan point is matched to an element in the as-designed BIM based on geometric and/or semantic features. For example, Zhang and Arditi [46] found laser scan points within a certain distance from the as-designed column, and treated them as belonging to the column. Bosche and Haas [47] created a simulated as-designed point for each as-built point based on the BIM model and matched each as-built point to the corresponding as-designed point based on the range difference between them. Third, after scan points are matched to as-designed elements, the as-is BIM for each element can be created based on geometric modelling methods. For example, an as-is wall element can be modelled by finding the least-squares fitting plane of all scan points matched to this wall.

Scenario 2: Building elements to be modelled have simple geometries and distinct geometric features. The recommended approach to this scenario is the hard-coded knowledge-based approach [5,49–51]. A simple example is to model walls, roof, and floor within a room. As all these building elements have planar surfaces, this approach first detects planes from laser scan data. Then, the detected planes are labelled into different object categories based on hard-coded prior knowledge including walls being vertical, roof being above and horizontal, and floor being below and horizontal. For the hard-coded knowledge-based approach, it is important to find distinct features for different object categories. Four categories of knowledge are commonly used including size, position, orientation, and topology [5]. Size measures the dimensions of an element, such as a door with a certain width or height. Position refers to the relative position of an element, such as a floor being below walls. Orientation is focused on the principal direction or normal vector of an element, such as

a floor being horizontal. In addition, topology looks at the topological relationship between elements, such as a wall intersecting a floor.

Scenario 3: The geometries of building elements are complex and a large amount of labelled dataset is available. This scenario can be handled using a supervised learning-based approach [3,14,52–54]. Large amounts of laser scan data that are already labelled into object categories are used to train classification algorithms such as support vector machines (SVM) and random forest (RF). Then, the trained classifier is able to classify new laser scan points into different object categories. The features used for classification include both geometric features such as shape descriptors and semantic features such as RGB colours. The supervised learning-based approach can be applied to individual scan points or scan data segments. For the former case, each individual laser scan point is classified into an object category and afterwards scan points belonging to the same object category are modelled as building elements using geometric modelling methods [52]. For the latter case, laser scan data are first segmented into multiple segments by geometric modelling methods or clustering techniques, and each scan data segment is then classified into an object category [3].

3.4.2. Modelling of Required Non-geometric Attributes

The modelling of non-geometric attributes from laser scan data mainly includes the modelling of spatial relationships between building elements and modelling of materials and surface properties of building elements. The surface properties of building elements including reflectivity and colour values are associated with each laser scan point. Hence, little processing is required to model these properties into the BIM model. On the other hand, the spatial relationships and materials cannot be directly obtained from laser scan data, thus requiring additional data processing. For relationship modelling, research is lacking on relationship modelling from laser scan data, although Nguyen et al. [55] deducted topological relationships between building components from 3D CAD models and Son et al. [56] leveraged topological relationships for as-is BIM reconstruction. For material identification, studies are reported to classify airborne laser scan data into ground, vegetation, and building roof using geometric and colour information [57,58]. A recent study attempted to classify scan data into structural concrete, light-weight concrete, and clay brick based on reflectivity and RGB colour values [59].

4. Illustrative Examples

4.1. Classroom Case Study

4.1.1. Identification of Information Requirements

This illustrative example uses the scan-to-BIM framework to create an as-is BIM model of a classroom for emergency evacuation simulation. As shown in Figure 7, the classroom is rectangular with a size of around 13 m (length) × 6 m (width) × 2.5 m (height). The classroom contains four doors at four corners respectively with an identical size of around 2 m (height) × 1 m (width). There are 17 identical rectangular desks in the classroom and each desk has a height of 0.75 m and a top surface size of around 1.5 m (length) × 0.7 m (width). For evacuation simulations, the geometric information of the floor (class of floor construction in OmniClass), ceiling (class of suspended ceiling construction in OmniClass), walls (class of interior partitions in OmniClass), doors (class of interior doors in OmniClass), and desks (class of furniture in OmniClass) is identified as the information requirements for as-is BIM, as illustrated in Figure 8. The floor, ceiling, and walls compose the profile of the classroom, and doors and desks serve as exits and obstacles in evacuation simulations. The geometries of the above-mentioned elements should be precisely modelled with 20 mm accuracy regarding their size, shape, and location. Note that only the top surfaces of desks need to be modelled because the locations of table legs can be inferred from the top surfaces. In addition, the swing direction of doors is also essential information to be modelled in the as-is BIM.

Figure 7. The classroom used for the illustrative example.

OmniClass ID	Element	Required LOD (Accuracy)	Non-geometric Attributes
21-02-10-10	Floor Construction (floor)	Precise geometry (20 mm for size, shape, and location)	
21-03-10-70	Suspended Ceiling Construction (ceiling)	Precise geometry (20 mm for size, shape, and location)	
21-03-10-10	Interior Partitions (walls)	Precise geometry (20 mm for size, shape, and location)	
21-03-10-30	Interior Doors (doors)	Precise geometry (20 mm for size, shape, and location)	Swing direction
21-05-20-50-30	Furniture (desks)	Precise geometry (20 mm for size, shape, and location) (only top surface needed)	

Figure 8. Information requirements for the as-is BIM of the classroom.

4.1.2. Determination of Required Scan Data Quality

Based on the information requirements, the required scan data quality is determined as shown in Table 3. Since the modelled elements should have 20 mm accuracy in size, shape, and location, the accuracy and spatial resolution of scan data for these elements must be equal to or less than 20 mm. As the floor, ceiling, and walls are planar and intersect with each other, it is possible to infer their size, shape, and location even though they have a low coverage. Hence, the required coverage for floor, ceiling, and walls is set to 20%. For doors, as the size of doors is known and the location of doors is partially known (i.e., doors must be in walls and must intersect with the floor), the only unknown information is the horizontal location of a door. It is found that each door has a glass panel, which can be easily recognized from laser scan data for door detection. As a result, the coverage requirement for doors is to ensure the glass panels are fully covered. For desks, because the shape and size of the top surface of a desk is known, at least two desk edges should be covered to determine the horizontal location of a desk. Lastly, scan data should include RGB colours to facilitate automated as-is BIM reconstruction.

Table 3. Required scan data quality for the classroom.

Scan Data Parameter	Requirement
Accuracy	≤20 mm
Spatial resolution	≤20 mm
Coverage	≥20% for floor, ceiling, and walls; the glass panel of each door must be covered; at least two edges of each desk must be covered
Other properties	RGB colours

4.1.3. Scan Data Acquisition

　　Based on the required scan data quality, scanning parameters including scanning devices, locations, resolutions, and other parameters are determined, as shown in Table 4. A commercial terrestrial laser scanner (FARO Focus S70) with 1 mm ranging accuracy within 25 m is adopted to fulfil the scan data accuracy requirement. Regarding scanning locations, because the classroom is small and symmetric, the simplest setting is considered, which is to conduct only one scan at the centre of the classroom as shown in Figure 9. The height of the laser scanner is set to 2 m to reduce the incident angles of laser beams for both desks and ceiling. Furthermore, scan data with different angular resolutions were simulated and it was found that 0.036° is the lowest angular resolution that can fulfil the requirements of spatial resolution and coverage.

Table 4. Scanning parameters for the classroom.

Scanning Parameter	Value
Scanning device	FARO Focus S70
Scanning location	Classroom centre
Angular resolution	0.036°
Other parameters	Camera enabled

Figure 9. The scanner location and the resulting scan data from the scan at classroom centre.

　　As shown in Table 5, with an angular resolution of 0.036°, the floor has a spatial resolution within 2–16 mm and the coverage is around 30% due to occlusions from desks. Although the ceiling has a spatial resolution between 0.6 mm and 65 mm due to the large incident angles near ceiling edges, around 60% of the ceiling area has a spatial resolution of less than 20 mm. Walls have a spatial resolution between 2 mm and 13 mm and the coverage is around 70% due to occlusions from desks. Doors have a spatial resolution between 12 mm and 13 mm, and all the glass panels are fully covered. Desks have a spatial resolution between 1 mm and 15 mm, and at least two edges of each desk are covered in spite of partial occlusions from chairs. Lastly, since the RGB colours of scan data are required, the camera should be enabled in scanning.

Table 5. Scan data quality with one scan at classroom centre and angular resolution of 0.036°.

Element	Accuracy	Spatial Resolution	Coverage	Other Properties
Floor		2–16 mm	~30%	
Ceiling		0.6–65 mm	~60%	
Walls	1 mm within 25 m	2–13 mm	~70%	RGB colours
Doors		12–13 mm	The glass panel of each door is covered.	
Desks		1–15 mm	At least two edges of each desk are covered.	

4.1.4. As-Is BIM Reconstruction

After scan data acquisition, the building elements are recognized from the scan data using the hard-coded knowledge-based approach. Planes are firstly detected from laser scan data using the RANdom SAmple Consensus (RANSAC) algorithm as shown in Figure 10. The RANSAC algorithm iteratively takes sample points from a dataset and generates an estimated plane based on the sample points. Eventually, the estimated plane that contains the most number of points from the dataset is selected as the final estimated plane. Then, hard-coded knowledge is utilized to classify the planes into floor, ceiling, walls, and desks. As shown in Table 6, the knowledge involves the size (e.g., size around 13 m × 6 m for the floor), position (e.g., height around 2.5 m for the ceiling), and orientation (e.g., vertical for walls) of the different building elements.
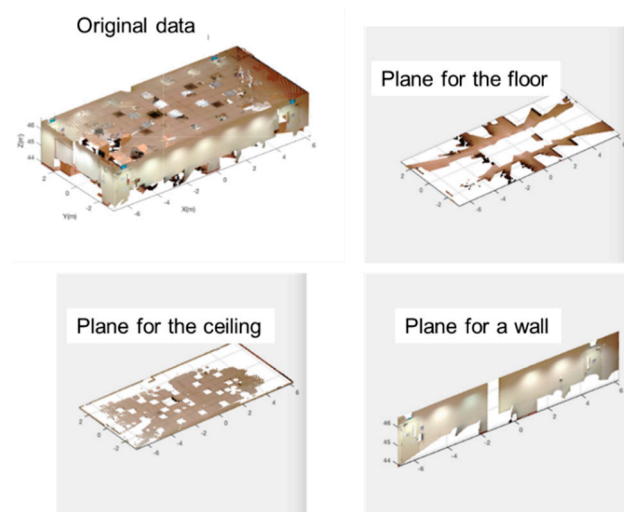

Figure 10. Detection of planes from laser scan data using the RANSAC algorithm.

Table 6. Hard-coded knowledge for classifying planes into floor, ceiling, walls, and desks.

Element	Knowledge
Floor	Horizontal, height around 0, size around 13 m × 6 m
Ceiling	Horizontal, height around 2.5 m, size around 13 m × 6 m
Walls	Vertical, size around or larger than 6 m × 2.5 m
Desks	Horizontal, height around 0.75 m

Because the top surfaces of all the desks are on the same plane, all the desks were detected as one single plane in the previous RANSAC and hard-coded knowledge-based approach. To recognize individual desks, scan data for all desks were further segmented using the region growing algorithm with the criteria in both distance and RGB colours. RGB colours are used because desks have substantially different colours from other objects such as walls and chairs. The region growing algorithm initiates a segment with a seed point and then grows the segment by iteratively adding neighbouring points that fulfil certain criteria into the segment. Here, the criteria include (1) the distance between a neighbouring point and the seed point is less than a certain threshold value, and (2) the colour difference between a neighbouring point and the seed point is less than a certain threshold value. As a result, scan points from desks are easily distinguished from scan points of other objects when taking RGB colours into consideration. As shown in Figure 11a, scan data for all desks are segmented into many segments shown in different colours. While some scan data segments do represent desks, some segments are actually from other objects. To further screen the scan data segments from region growing, prior knowledge of the size of a desk is utilized. Each scan data segment is registered to and compared to the actual shape of a desk. Then, two metrics are computed, namely, degree of matching (DOM) and degree of completeness (DOC), to evaluate how likely a

segment represents a desk. The two metrics were initially proposed in [52], and they measure the degree to which a segment matches the actual shape of a desk and the degree to which a segment is complete relative to the actual shape of a desk, respectively. All scan data segments with DOM > 0.8 and DOC > 0.5 are regarded as representing a desk and are shown in Figure 11b. Note that the completeness of the scan data for each desk ranged from 66% to 100%. Although some desks had a low completeness, the desks were accurately modelled because at least two edges of each desk were covered by the scan data.
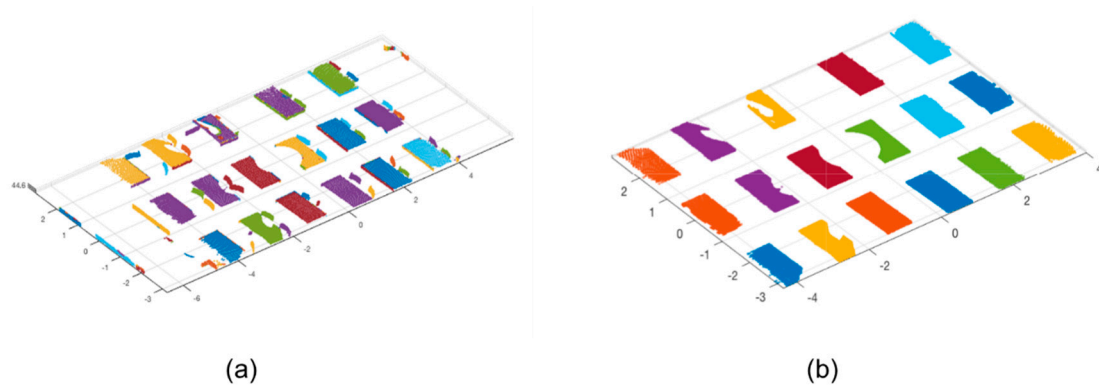

(a)   (b)

Figure 11. Recognition of individual desks: (**a**) Segmentation of scan data using region growing, and (**b**) Selection of scan data segments using DOM and DOC metrics.

Doors are recognized based on the glass panel on each door (Figure 12a). As laser beams travel through glass, there are no scan data on the glass panel. Hence, areas on walls but without any scan data are firstly identified, and the boundaries of such areas are shown in Figure 12b. Then, the location of each door is obtained from the location of the respective glass panel, as marked in red rectangles in Figure 12c.
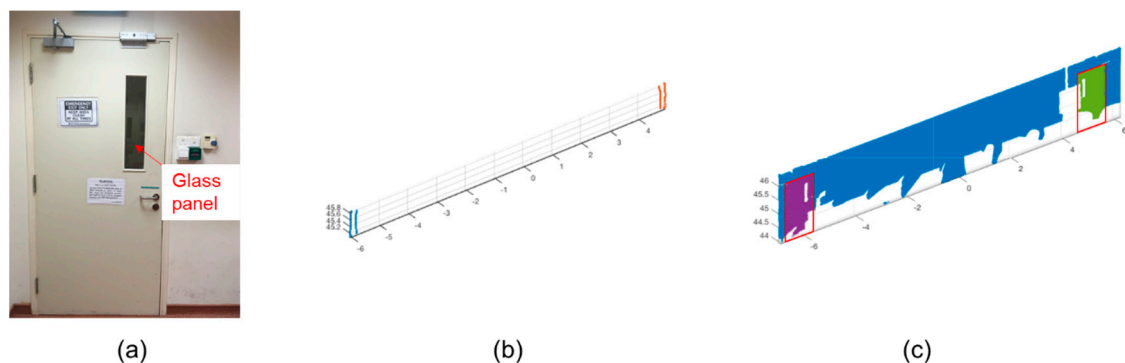

(a)   (b)   (c)

Figure 12. Recognition of doors: (**a**) Glass panel on a door, (**b**) Boundaries of areas with no scan data represent glass panels on doors, and (**c**) Extraction of door locations (red rectangles) based on glass panel locations.

4.1.5. Evaluation

After reconstructing the as-is BIM, the accuracy of the reconstructed BIM was examined by comparing the dimensions of the reconstructed BIM with the ground-truth values from manual measurements. Five groups of dimensions were selected for comparison, as summarized in Table 7. The height of the classroom was measured as the vertical distance from floor to ceiling, and was measured at 10 different locations. The length of the classroom was measured as the horizontal distance between two short walls, and was measured at 10 different locations. The width of the classroom was measured as the horizontal distance between two long walls, and was measured at 10 different locations. The position of the doors was measured as the distance from each door

boundary to its nearest wall boundary, and was measured for four doors. The position of desks was measured as the distances from each desk boundary to its nearest long wall and its nearest short wall, and was measured for 17 desks. As shown in Table 7, the average discrepancies for the above-mentioned five groups of dimensions were 1.3 mm, 1.8 mm, 1.5 mm, 8.2 mm, and 6.6 mm, respectively. For each group, the maximum discrepancies were 2.6 mm, 2.8 mm, 2.1 mm, 12.9 mm, and 16.4 mm, respectively. The results indicate that the reconstructed BIM has achieved the required model LOD (i.e., 20 mm accuracy regarding their size, shape, and location). Among the five groups of dimensions, the discrepancies for door positions and desk positions were relatively large due to relatively low spatial resolution for the scan data of door and desks.

Table 7. Discrepancies between the reconstructed as-is BIM and the ground-truth dimensions of the classroom.

Dimension	Measurement Method	Average Discrepancy (mm)	Maximum Discrepancy (mm)
Height of classroom	Vertical distance from floor to ceiling, measured at 10 different locations	1.3	2.6
Length of classroom	Horizontal distance between two short walls, measured at 10 different locations	1.8	2.8
Width of classroom	Horizontal distance between two long walls, measured at 10 different locations	1.5	2.1
Positions of doors	Distance from each door to its nearest wall boundary, measured for four doors	8.2	12.9
Positions of desks	Distances from each desk to its nearest long wall and its nearest short wall, measured for 17 desks	6.6	16.4

The illustrative example was fully automatic except that the required scan data quality and the hard-coded knowledge for object recognition needed to be determined manually. The total computational time for data processing was 85 seconds using an Intel Core i5-7200U CPU.

4.2. Precast Panel Case Study

4.2.1. Identification of Information Requirements

This case study uses the scan-to-BIM framework to create an as-is BIM model of a precast panel for geometry quality assessment. As shown in Figure 13, the panel is cuboid and has a dimension of around 12.6 m (length) $\times$ 2.5 m (width) $\times$ 0.24 m (height). On the top surface of the panel, there are 25 rectangular shear pockets with an identical size of 0.44 m $\times$ 0.14 m, which are used to connect the panel with the girder. Furthermore, on each long side surface of the panel, there are 23 shear keys with an identical outer size of 0.14 m $\times$ 0.14 m and 14 flat ducts with an identical outer size of 0.13 m $\times$ 0.06 m. Each shear key has five small surfaces (four inclined and one vertical), and each flat duct has four small inclined surfaces. The shear keys and flat ducts are used to connect the panel with adjacent panels.

For geometry quality assessment, the information requirements include the dimensions of the top surface and side surfaces, and the dimensions and positions of shear pockets, shear keys, and flat ducts. Because geometry quality assessment requires an accuracy of 5 mm according to relevant construction codes, the top surface and side surfaces require 5 mm accuracy in size, and shear pockets, shear keys, and flat ducts require 5 mm accuracy in size and position, as illustrated in Figure 14. Note that only the outer sizes of shear keys and flat ducts are required. It means that only the four outer edges of each shear key and each flat duct need to be estimated from the laser scan data. For all the elements, non-geometric attributes are not needed.
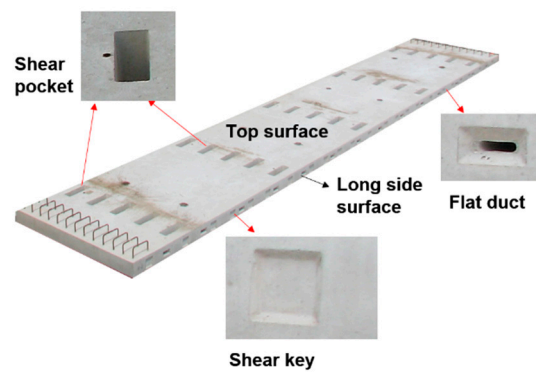
Figure 13. The precast panel used for the illustrative example.

Element	Required LOD (Accuracy)	Non-geometric Attributes
Top surface	Precise geometry (5 mm for size)	
Side surfaces	Precise geometry (5 mm for size)	
Shear pockets	Precise geometry (5 mm for size and position)	
Shear keys	Precise geometry (5 mm for outer size and position)	
Flat ducts	Precise geometry (5 mm for outer size and position)	

Figure 14. The information requirements for the precast panel.

4.2.2. Determination of Required Scan Data Quality

Based on the information requirements, the required scan data quality is determined as shown in Table 8. Since the modelled elements should have 5 mm accuracy in size and position, the accuracy and spatial resolution of scan data for these elements must be equal to or less than 5 mm. Because the geometric quality assessment requires high accuracy and high completeness of scan data, the required coverage is set to 100% for all the elements. Note that all the five small surfaces for each shear key and all the four small surfaces for each flat duct must be covered to provide an accurate estimation of the size and position of each shear key or flat duct. No other properties are required.

Table 8. Required scan data quality for the precast panel.

| Scan Data Parameter	Requirement
Accuracy	≤5 mm
Spatial resolution	≤5 mm
Coverage	100% for all elements
Other properties	N/A

4.2.3. Scan Data Acquisition

Based on the required scan data quality, scanning parameters including scanning devices, locations, resolutions, and other parameters are determined, as shown in Table 9. A commercial terrestrial laser scanner (FARO Focus S70) with 1 mm ranging accuracy within 25 m is adopted to fulfil the scan data accuracy requirement. To acquire high resolution scan data, a higher angular resolution 0.018° was adopted. Numerical simulations were conducted to determine the scanning locations to minimize the number of scans while the required scan data quality is fulfilled. We first determined a set of measurement target points on the precast panel surfaces to be scanned. The spacing between adjacent target points is 5 mm, the same as the required minimum spatial resolution. In addition,

the whole 3D space is divided by grids along X, Y, and Z directions with 0.5 m interval, and only grid points are treated as candidature scanning locations. Then, we used a greedy algorithm to find the optimal set of scanning locations from the candidature scanning locations. The greedy algorithm firstly finds the scanning location that can scan the maximum number of target points with the required spatial resolution. Next, the greedy algorithm continues to find the next scanning location that can scan the maximum number of new target points (i.e., target points that were not scanned in previous scanning locations) with the required spatial resolution. The greedy algorithm terminates when all the target points are scanned. As shown in Figure 15, a total of six scanning locations (L1-L6) were determined from the greedy algorithm.

Table 9. Scanning parameters for the precast panel.

Scanning Parameter	Value
Scanning device	FARO Focus S70
Scanning location	Six different locations
Angular resolution	0.018°
Other parameters	N/A
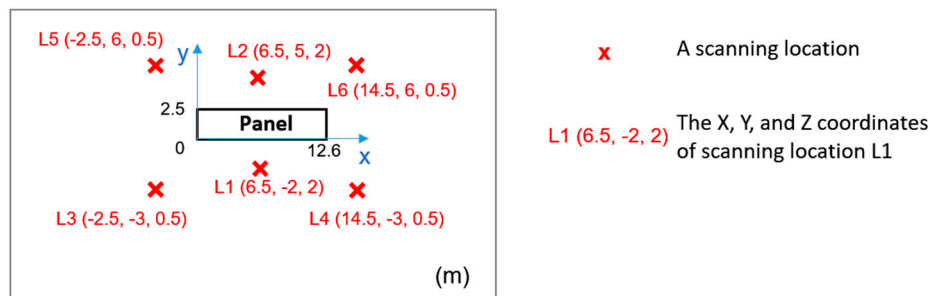
Note: X axis: along the long side of panel; Y axis: along the short side of panel; Z axis: along the height of panel

Figure 15. The scanning locations for the precast panel.

Based on the simulations, the achieved scan data quality is shown in Table 10. Because all the scans are within 25 m from the target, the ranging accuracy is 1 mm. Based on the six scanning locations, the spatial resolutions of scan data are all within 5 mm for all the elements. On average, the top and side surfaces have relatively high spatial resolutions. Shear keys and flat ducts have relatively low spatial resolutions because they have inclined surfaces that result in low spatial resolutions. It is worth noting that because shear pockets are actually holes, there are no scan data on shear pockets. Therefore, the spatial resolution of shear pockets refers to the spatial resolution of scan data near the edges of shear pockets because these scan data will be used to estimate the edges of shear pockets. A coverage ratio of 100% is achieved for all the elements.

Table 10. Achieved scan data quality for the precast panel.

Element	Accuracy	Spatial Resolution	Coverage	Other Properties
Top surface		1.2–3.8 mm		
Side surfaces		0.6–3.2 mm		
Shear pockets	1 mm	1.4–3.6 mm (for scan data near the edges of shear pockets)	100%	N/A
Shear keys		0.9–5 mm		
Flat ducts		0.9–4.8 mm		

4.2.4. As-Is BIM Reconstruction

The scan-vs.-BIM approach is adopted to recognize elements and reconstruct the as-is BIM model. First, the scan data are aligned with the as-designed BIM based on a two-step registration process. The coarse registration is conducted by matching the scan data with the as-designed BIM based on the three principal directions from the principal component analysis (PCA), and the fine registration further improves the registration performance using the iterative closest point (ICP) algorithm. Second, each scan data point is matched to a surface of the as-designed BIM model. The local normal directions of points and distances from points to as-designed surfaces are used to determine which surface a point belongs to. Figure 16a shows an example, where points belonging to the five small surfaces (No. 1–5) of a shear key and points belong to the long side surface (No. 6) of the panel are identified and shown in different colours. Third, edge lines of all the elements are identified. In general, there are two types of edge lines. The first type of edge lines are the intersecting boundaries of two planes where scan data are available for both planes. For the first type of edge lines, the edge lines are obtained as the intersection lines of two intersecting as-is surfaces, and each as-is surface is obtained as the least-squares fitting plane of all points belonging to it. For example, the right edge line of the shear key is obtained as the intersection line of surfaces 2 and 6, as shown in Figure 16b. On the other hand, the second type of edge lines have scan data available from only a single surface without any scan data from the other surface. One example is the edge lines of shear pockets, because the inner vertical surfaces of shear pockets are not scanned. For the second type of edge lines, the edge lines are estimated based on scan points near the edges using the method developed by Wang et al. [4]. Lastly, after obtaining the edge lines of all elements, the as-is BIM is reconstructed as shown in Figure 16c.
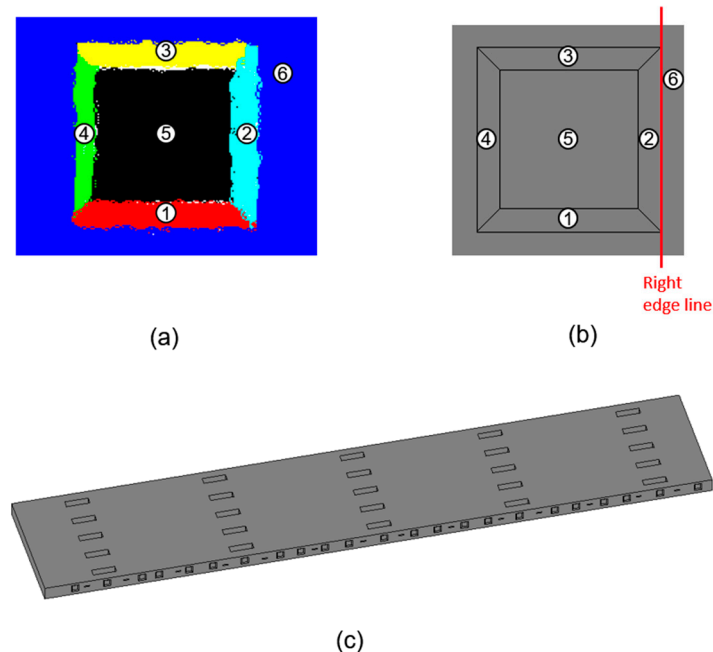

(a)　　　　　　　　　　　　　　　　　　(b)

(c)

Figure 16. As-is BIM reconstruction of the precast panel: (**a**) matching scan points to different as-designed surfaces, (**b**) extraction of edge lines as the intersection lines of as-is planes, and (**c**) the reconstructed as-is BIM model of the precast panel.

4.2.5. Evaluation

To examine the accuracy of the reconstructed as-is BIM, the dimensions of the as-is BIM were compared to the ground-truth values from manual measurements. Six groups of dimensions were selected for comparisons, as summarized in Table 11. The dimensions of shear pockets are measured as the lengths of four boundaries of each shear pocket for a total of 25 shear pockets. The dimensions of

shear keys are measured as the lengths of four outer boundaries of each shear key for a total of 23 shear keys. The dimensions of flat ducts are measured as the lengths of four outer boundaries of each flat duct for a total of 14 flat ducts. The length of the precast panel is measured as the distances between two short side surfaces of the panel, and is measured at 10 different locations. The width of the precast panel is measured as the distances between two long side surfaces of the panel, and is measured at 10 different locations. The height of the precast panel is measured as the distances between the top and bottom surfaces of the panel, and is measured at 10 different locations. As shown in Table 11, the average discrepancies for the above-mentioned six groups of dimensions were 2.2 mm, 1.2 mm, 1.4 mm, 2.6 mm, 2.1 mm, and 2.4 mm, respectively. For each group, the maximum discrepancies were 3.8 mm, 2.4 mm, 2.6 mm, 4.9 mm, 4.4 mm, and 4.0 mm, respectively. The results show that the reconstructed BIM achieved the required model LOD (i.e., 5 mm accuracy).

Table 11. Discrepancies between the reconstructed as-is BIM and the ground-truth dimensions of the precast panel.

Dimension	Measurement Method	Average Discrepancy (mm)	Maximum Discrepancy (mm)
Dimensions of shear pockets	Lengths of four boundaries of each shear pocket, measured for 25 shear pockets	2.2	3.8
Dimensions of shear keys	Lengths of four outer boundaries of each shear key, measured for 23 shear keys	1.2	2.4
Dimensions of flat ducts	Lengths of four outer boundaries of each flat duct, measured for 14 flat ducts	1.4	2.6
Length of precast panel	Distances between two short side surfaces of panel, measured at 10 different locations	2.6	4.9
Width of precast panel	Distances between two long side surfaces of panel, measured at 10 different locations	2.1	4.4
Height of precast panel	Distances between top and bottom surfaces, measured at 10 different locations	2.4	4.0

The case study on precast panel was fully automatic except that the required scan data quality and the measurement target points for scan planning needed to be determined manually. The total computational time for data processing was 380 seconds using an Intel Core i5-7200U CPU.

5. Future Research within the Scan-to-BIM Framework

Although this study proposes an application oriented scan-to-BIM framework, there are still many technical issues within the framework to be resolved in future research. Three future research directions are identified below.

5.1. Information Requirements for BIM Applications

The information requirements of as-is BIM for different applications need to be further studied. In the illustrative example, it is assumed that the created BIM is used for emergency evacuation simulation. For other applications, the information requirements could be substantially different, which remains to be studied in the future. It is often easy to decide the buildings elements to be modelled in an as-is BIM according to the intended BIM application. However, it is more difficult to determine the required LOD and modelling accuracy for the building elements. Higher modelling accuracy usually improves the reliability of the as-is BIM for the intended application; however; it also involves a higher cost for modelling. Hence, a trade-off must be made to balance the model reliability

and modelling costs. In this trade-off, it is desirable to consider the quantitative relationship between the modelling accuracy and the reliability of the as-is BIM.

For example, when using an as-is BIM for QA/QC, the modelling accuracy should be decided based on the dimensional tolerances of elements according to relevant codes and regulations. Assuming that the dimensional tolerance of a floor slab is 10 mm and the modelling accuracy is ±5 mm, an actual dimensional deviation of 11 mm between the as-is and as-designed conditions is likely to be modelled between 6 mm (=11-5) and 16 mm (=11+5). In this case, the deviation of 11 mm, which actually exceeds the tolerance of 10 mm, has a 40% chance of being modelled as less than the tolerance 10 mm by mistake. One recent study also quantitatively investigated the relationship between point cloud quality (which is closely related to modelling accuracy) and the success of identifying building elements for construction progress tracking [35]. Three criteria were defined using a statistical method to measure point cloud quality and experiments showed that the defined quality criteria can be applied to decide the appropriate scanning methods for successful construction progress tracking.

However, research is still lacking to quantify the relationship between modelling accuracy or point cloud quality and the reliability of as-is BIM for other BIM applications in the construction and O&M phases. Particularly, for existing buildings, as-is BIM has great potential to support various FM functionalities in the O&M phase. The information requirements for various FM functionalities and the quantitative relationships between information requirements and BIM reliability remain to be studied.

5.2. Scan Planning

Previous studies have proposed methods to optimize the scanning parameters such that scan data with a certain quality are obtained with the minimum number of scans or minimum scanning time. However, one precondition of these methods is that a BIM model exists for the building to be scanned so that the scan data can be simulated for different scanning parameters through computer simulations. However, in most of the real-world applications for existing buildings, no BIM model exists for the building to be scanned. Hence, the optimization of scanning parameters remains a problem for cases where the BIM model is missing.

Furthermore, the existing studies are all focused on terrestrial laser scanning. Nowadays, with the advances in unmanned aerial vehicle (UAV or drone) technology, UAV-mounted laser scanning has become popular in the construction industry due to its high mobility. Efforts have been made in both industry and academia to deploy UAV-mounted laser scanners for a variety of applications such as building façade inspection, evacuation volume calculation, ground settlement monitoring, etc. However, the scan planning problem remains to be solved for UAV-mounted laser scanning, especially when scanning new environments. One possible solution is to first scan the environment with a low resolution to reconstruct a low LOD BIM model. Then, the low LOD BIM model is utilized for scan planning to eventually reconstruct a high LOD BIM model that fulfils the needs of the intended BIM application. Another possible solution is to monitor scan data quality in real-time while scanning and adjust the scanning parameters once the data quality is too low or too high.

5.3. As-Is BIM Reconstruction Techniques

Although numerous studies are reported for semi-automatic or automatic as-is BIM reconstruction from laser scan data, the existing reconstruction techniques still have a room for improvement regarding accuracy, applicability, and automation. In general, it is difficult to achieve high accuracy, high applicability, and high levels of automation at the same time. For example, the hard-coded knowledge-based object recognition technique can achieve a high accuracy and high automation once the knowledge is hard-coded. However, this technique is applicable only to elements within strict assumptions. Any irregular shape elements such as an inclined wall will be a challenge. As another example, some software packages provide interactive modelling tools, which can automatically recognize the geometry (e.g., plane, cylinder, cuboid) when the user clicks in a certain region. Such tools have a high accuracy and applicability, but a low level of automation.

Currently, most of the existing studies have focused on the automation of the as-is BIM reconstruction process. Hence, assumptions on the elements to be modelled are always needed to narrow down the scope and make sure the technique has a high accuracy. One future research direction is to study how to lose the assumptions such that the techniques have a wider applicability. Regarding accuracy, future research is needed to fuse data from multiple sensors to improve the recognition accuracy including using 3D scan data with colours for more accurate material recognition. For automation, one potential research direction is to eliminate the tedious process of labelling training datasets for supervised learning-based object recognition. BIM object families available online can be utilized to generate labelled training datasets after transforming BIM objects into simulated scan data.

6. Conclusions

This study proposes an application oriented framework for scan-to-BIM, which describes the major steps within a scan-to-BIM process and their relationships. The framework is oriented towards the specific BIM application to be implemented using the created as-is BIM. Firstly, the information requirements of the specific BIM application are identified, including the required building elements to be modelled, the required LOD, and the required non-geometric attributes. Secondly, the required scan data quality is determined based on the information requirements. The scan data quality is defined using four parameters including accuracy, spatial resolution, coverage, and other properties. The required scan data accuracy and spatial resolution are determined by the required LOD of each building element; the required scan data coverage is determined by the required building elements and their LOD requirements; and the other properties are determined by the required non-geometric attributes of each building element. Following the determination of scan data quality, scan data of the building are acquired in the third step. The scanning parameters for data acquisition include scanning devices, scanning locations, angular resolutions, and other parameters. The scanning devices are determined by the required scan data accuracy and coverage; scanning locations are determined by the required scan data accuracy, spatial resolution, and coverage; angular resolutions are determined by the required scan data spatial resolution and coverage; and other scanning parameters are determined by the other properties of the scan data. In the last step, the as-is BIM is reconstructed from laser scan data including the modelling of required building elements with required LOD and the modelling of non-geometric attributes. For the modelling of required building elements, three common scenarios are summarized and their respective modelling strategies are recommended. The proposed framework is summarized and illustrated in Figure 17.

Two illustrative examples are provided to demonstrate the proposed scan-to-BIM framework. The first example is to create an as-is BIM for a classroom for evacuation simulation, and the second example is to create an as-is BIM for a precast panel for geometric quality assessment. The proposed scan-to-BIM framework was successfully implemented on the two cases. By comparing the reconstructed as-is BIM with the ground-truth dimensions, it is validated that the proposed method could provide as-is BIM models that satisfy the needs of the intended application. This study also identifies three main future research directions within the scan-to-BIM framework. First, the information requirements for different BIM applications should be identified and the quantitative relationships between the modelling accuracy or point cloud quality and the reliability of as-is BIM for its intended use should be investigated. Second, scan planning techniques should be further studied for cases in which an as-designed BIM does not exists and for UAV-mounted laser scanning. Third, as-is BIM reconstruction techniques should be improved with regard to accuracy, applicability, and level of automation.
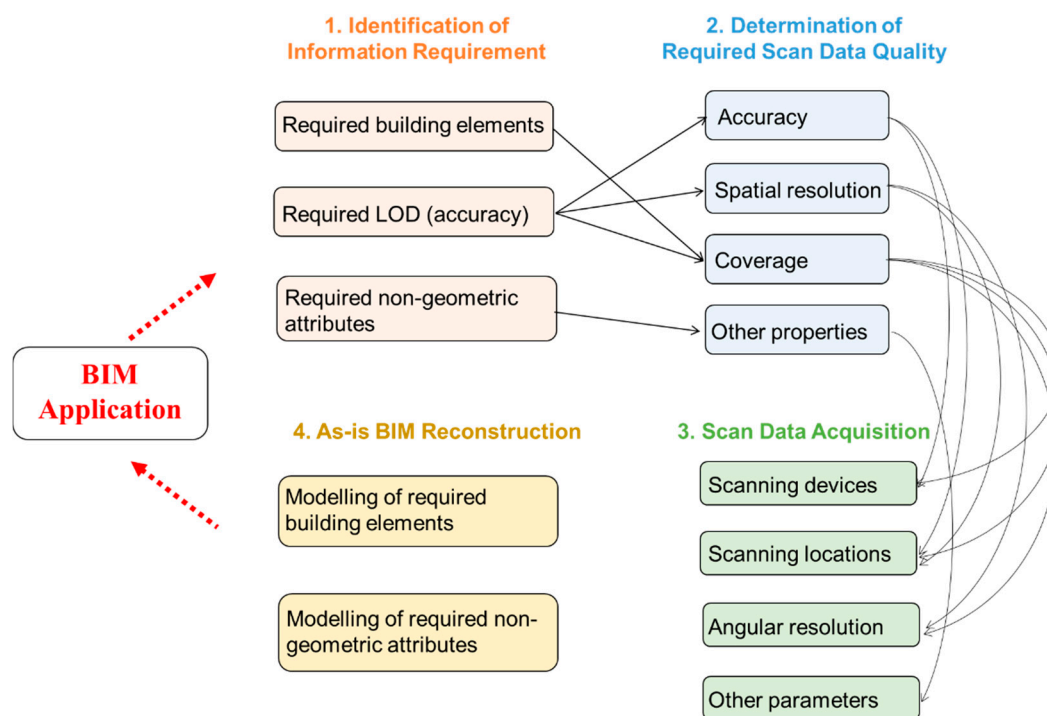
Figure 17. Summary of the proposed scan-to-BIM framework.

Author Contributions: Conceptualization, Q.W., J.G. and M.-K.K.; Methodology, Q.W., J.G. and M.-K.K.; Validation, Q.W. and J.G.; Writing—original draft, Q.W. and J.G.; Writing—review & editing, M.-K.K.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. 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]

2. Bosché, F.; Ahmed, M.; Turkan, Y.; Haas, C.T.; Haas, R. The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components. *Autom. Constr.* **2015**, *49*, 201–213. [CrossRef]

3. Xiong, X.; Adan, A.; Akinci, B.; Huber, D. Automatic creation of semantically rich 3D building models from laser scanner data. *Autom. Constr.* **2013**, *31*, 325–337. [CrossRef]

4. Wang, Q.; Sohn, H.; Cheng, J.C. Automatic As-Built BIM Creation of Precast Concrete Bridge Deck Panels Using Laser Scan Data. *J. Comput. Civ. Eng.* **2018**, *32*, 04018011. [CrossRef]

5. Pu, S.; Vosselman, G. Knowledge based reconstruction of building models from terrestrial laser scanning data. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 575–584. [CrossRef]

6. Brilakis, I.; Lourakis, M.; Sacks, R.; Savarese, S.; Christodoulou, S.; Teizer, J.; Makhmalbaf, A. Toward automated generation of parametric BIMs based on hybrid video and laser scanning data. *Adv. Eng. Inform.* **2010**, *24*, 456–465. [CrossRef]

7. Volk, R.; Stengel, J.; Schultmann, F. Building Information Modeling (BIM) for existing buildings—Literature review and future needs. *Autom. Constr.* **2014**, *38*, 109–127. [CrossRef]

8. 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]

9.　Aydin, C.C. Designing building façades for the urban rebuilt environment with integration of digital close-range photogrammetry and geographical information systems. *Autom. Constr.* **2014**, *43*, 38–48. [CrossRef]

10.　Wang, Q.; Kim, M.-K.; Cheng, J.C.; Sohn, H. Automated quality assessment of precast concrete elements with geometry irregularities using terrestrial laser scanning. *Autom. Constr.* **2016**, *68*, 170–182. [CrossRef]

11.　Bosché, F.; Guenet, E. Automating surface flatness control using terrestrial laser scanning and building information models. *Autom. Constr.* **2014**, *44*, 212–226. [CrossRef]

12.　Tang, P.; Huber, D.; Akinci, B. A comparative analysis of depth-discontinuity and mixed-pixel detection algorithms. In Proceedings of the Sixth IEEE International Conference on 3-D Digital Imaging and Modeling, 3DIM'07, Montreal, QC, Canada, 21–23 August 2007.

13.　Bosché, F. Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction. *Adv. Eng. Inform.* **2010**, *24*, 107–118. [CrossRef]

14.　Kim, C.; Son, H.; Kim, C. Automated construction progress measurement using a 4D building information model and 3D data. *Autom. Constr.* **2013**, *31*, 75–82. [CrossRef]

15.　Son, H.; Kim, C. 3D structural component recognition and modeling method using color and 3D data for construction progress monitoring. *Autom. Constr.* **2010**, *19*, 844–854. [CrossRef]

16.　Turkan, Y.; Bosche, F.; Haas, C.T.; Haas, R. Automated progress tracking using 4D schedule and 3D sensing technologies. *Autom. Constr.* **2012**, *22*, 414–421. [CrossRef]

17.　Nahangi, M.; Yeung, J.; Haas, C.T.; Walbridge, S.; West, J. Automated assembly discrepancy feedback using 3D imaging and forward kinematics. *Autom. Constr.* **2015**, *56*, 36–46. [CrossRef]

18.　Rausch, C.; Nahangi, M.; Haas, C.; West, J. Kinematics chain based dimensional variation analysis of construction assemblies using building information models and 3D point clouds. *Autom. Constr.* **2017**, *75*, 33–44. [CrossRef]

19.　Zhang, S.; Teizer, J.; Lee, J.-K.; Eastman, C.M.; Venugopal, M. Building information modeling (BIM) and safety: Automatic safety checking of construction models and schedules. *Autom. Constr.* **2013**, *29*, 183–195. [CrossRef]

20.　Wang, J.; Zhang, S.; Teizer, J. Geotechnical and safety protective equipment planning using range point cloud data and rule checking in building information modeling. *Autom. Constr.* **2015**, *49*, 250–261. [CrossRef]

21.　Xu, J.; Ding, L.; Love, P.E. Digital reproduction of historical building ornamental components: From 3D scanning to 3D printing. *Autom. Constr.* **2017**, *76*, 85–96. [CrossRef]

22.　Yastikli, N. Documentation of cultural heritage using digital photogrammetry and laser scanning. *J. Cult. Herit.* **2007**, *8*, 423–427. [CrossRef]

23.　Rüther, H.; Chazan, M.; Schroeder, R.; Neeser, R.; Held, C.; Walker, S.J.; Matmon, A.; Horwitz, L.K. Laser scanning for conservation and research of African cultural heritage sites: The case study of Wonderwerk Cave, South Africa. *J. Archaeol. Sci.* **2009**, *36*, 1847–1856. [CrossRef]

24.　Ham, Y.; Golparvar-Fard, M. An automated vision-based method for rapid 3D energy performance modeling of existing buildings using thermal and digital imagery. *Adv. Eng. Inform.* **2013**, *27*, 395–409. [CrossRef]

25.　Balado, J.; Díaz-Vilariño, L.; Arias, P.; Soilán, M. Automatic building accessibility diagnosis from point clouds. *Autom. Constr.* **2017**, *82*, 103–111. [CrossRef]

26.　Riveiro, B.; Caamaño, J.; Arias, P.; Sanz, E. Photogrammetric 3D modelling and mechanical analysis of masonry arches: An approach based on a discontinuous model of voussoirs. *Autom. Constr.* **2011**, *20*, 380–388. [CrossRef]

27.　Becerik-Gerber, B.; Jazizadeh, F.; Li, N.; Calis, G. Application areas and data requirements for BIM-enabled facilities management. *J. Constr. Eng. Manag.* **2011**, *138*, 431–442. [CrossRef]

28.　OmniClass. The OmniClass Construction Classification System. 2018. Available online: http://www.omniclass.org/ (accessed on 14 April 2018).

29.　Messner, J.; Anumba, C.; Dubler, C.; Goddman, S.; Kasprzak, C.; Kreider, R.; Leicht, R.; Saluja, C.; Zikic, N. *BIM Project Execution Planning Guide*; The Computer Integrated Construction Research Program (CIC), Penn State University: University Park, PA, USA, 2010.

30.　Pishdad-Bozorgi, P.; Gao, X.; Eastman, C.; Self, A.P. Planning and developing facility management-enabled building information model (FM-enabled BIM). *Autom. Constr.* **2018**, *87*, 22–38. [CrossRef]

31.　Construction Specifications Institute (CSI) and Construction Specifications Canada (CSC). MasterFormat. 2016. Available online: https://www.csiresources.org/practice/standards/masterformat (accessed on 24 April 2018).

32. NBS. Uniclass 2015. 2015. Available online: https://www.thenbs.com/services/our-tools/introducing-uniclass-2015 (accessed on 24 April 2018).

33. Reddy, K.P. *BIM for Building Owners and Developers: Making a Business Case for Using BIM on Projects*; John Wiley & Sons: Hoboken, NJ, USA, 2012.

34. BIMForum. Level of Development Specification. 2017. Available online: http://bimforum.org/lod/ (accessed on 24 April 2018).

35. 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]

36. GSA. GSA BIM Guide for 3D Imaging. 2009. Available online: https://www.gsa.gov/cdnstatic/GSA_BIM_Guide_Series_03.pdf (accessed on 25 April 2018).

37. Dai, F.; Rashidi, A.; Brilakis, I.; Vela, P. Comparison of image-based and time-of-flight-based technologies for three-dimensional reconstruction of infrastructure. *J. Constr. Eng. Manag.* **2012**, *139*, 69–79. [CrossRef]

38. Tang, P.; Akinci, B.; Huber, D. Quantification of edge loss of laser scanned data at spatial discontinuities. *Autom. Constr.* **2009**, *18*, 1070–1083. [CrossRef]

39. Kaasalainen, S.; Jaakkola, A.; Kaasalainen, M.; Krooks, A.; Kukko, A. Analysis of incidence angle and distance effects on terrestrial laser scanner intensity: Search for correction methods. *Remote Sens.* **2011**, *3*, 2207–2221. [CrossRef]

40. Biswas, H.K.; Bosché, F.; Sun, M. Planning for scanning using building information models: A novel approach with occlusion handling. In Proceedings of the 32nd International Symposium on Automation and Robotics in Construction and Mining: Connected to the Future, Oulu, Finland, 15–18 June 2015.

41. Argüelles-Fraga, R.; Ordóñez, C.; García-Cortés, S.; Roca-Pardiñas, J. Measurement planning for circular cross-section tunnels using terrestrial laser scanning. *Autom. Constr.* **2013**, *31*, 1–9. [CrossRef]

42. Quintana, B.; Prieto, S.; Adán, A.; Vázquez, A.S. Semantic scan planning for indoor structural elements of buildings. *Adv. Eng. Inform.* **2016**, *30*, 643–659. [CrossRef]

43. Zhang, C.; Kalasapudi, V.S.; Tang, P. Rapid data quality oriented laser scan planning for dynamic construction environments. *Adv. Eng. Inform.* **2016**, *30*, 218–232. [CrossRef]

44. Lu, Q.; Lee, S. Image-Based Technologies for Constructing As-Is Building Information Models for Existing Buildings. *J. Comput. Civ. Eng.* **2017**, *31*, 04017005. [CrossRef]

45. Ma, Z.; Liu, S. A review of 3D reconstruction techniques in civil engineering and their applications. *Adv. Eng. Inform.* **2018**, *37*, 163–174. [CrossRef]

46. Zhang, C.; Arditi, D. Automated progress control using laser scanning technology. *Autom. Constr.* **2013**, *36*, 108–116. [CrossRef]

47. Bosche, F.; Haas, C.T. Automated retrieval of 3D CAD model objects in construction range images. *Autom. Constr.* **2008**, *17*, 499–512. [CrossRef]

48. Bosche, F.; Haas, C.T.; Akinci, B. Automated recognition of 3D CAD objects in site laser scans for project 3D status visualization and performance control. *J. Comput. Civ. Eng.* **2009**, *23*, 311–318. [CrossRef]

49. Sanchez, V.; Zakhor, A. Planar 3D modeling of building interiors from point cloud data. In Proceedings of the 2012 19th IEEE International Conference on Image Processing (ICIP), Orlando, FL, USA, 30 September–3 October 2012.

50. Rusu, R.B.; Marton, Z.C.; Blodow, N.; Dolha, M.; Beetz, M. Towards 3D point cloud based object maps for household environments. *Robot. Auton. Syst.* **2008**, *56*, 927–941. [CrossRef]

51. Quintana, B.; Prieto, S.; Adán, A.; Bosché, F. Door detection in 3D coloured point clouds of indoor environments. *Autom. Constr.* **2018**, *85*, 146–166. [CrossRef]

52. Wang, Q.; Cheng, J.C.; Sohn, H. Automated Estimation of Reinforced Precast Concrete Rebar Positions Using Colored Laser Scan Data. *Comput.-Aided Civ. Infrastruct. Eng.* **2017**, *32*, 787–802. [CrossRef]

53. Xu, Y.; Tuttas, S.; Hoegner, L.; Stilla, U. Reconstruction of scaffolds from a photogrammetric point cloud of construction sites using a novel 3D local feature descriptor. *Autom. Constr.* **2018**, *85*, 76–95. [CrossRef]

54. Chen, J.; Fang, Y.; Cho, Y.K.; Kim, C. Principal axes descriptor for automated construction-equipment classification from point clouds. *J. Comput. Civ. Eng.* **2016**, *31*, 04016058. [CrossRef]

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

56. Son, H.; Kim, C.; Kim, C. 3D reconstruction of as-built industrial instrumentation models from laser-scan data and a 3D CAD database based on prior knowledge. *Autom. Constr.* **2015**, *49*, 193–200. [CrossRef]

57. Mallet, C.; Bretar, F.; Roux, M.; Soergel, U.; Heipke, C. Relevance assessment of full-waveform lidar data for urban area classification. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, S71–S84. [CrossRef]

58. Vu, T.T.; Yamazaki, F.; Matsuoka, M. Multi-scale solution for building extraction from LiDAR and image data. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 281–289. [CrossRef]

59. Hassan, M.U.; Akcamete-Gungor, A.; Meral, C. Investigation of Terrestrial Laser Scanning Reflectance Intensity and RGB Distributions to Assist Construction Material Identification. In Proceedings of the Joint Conference on Computing in Construction (JC3), Heraklion, Greece, 4–7 July 2017.