

From crowd to cloud: simplified automatic reconstruction of digital building assets for facility management

From crowd to cloud

401

Srinimalan Balakrishnan Selvakumaran
*Institute of Technology in Architecture, ETH Zurich, Zurich, Switzerland and
Basler and Hofmann AG, Zurich, Switzerland, and*

Received 1 February 2021
Revised 11 June 2021
Accepted 28 July 2021

Daniel Mark Hall
*Institute of Construction and Infrastructure Management, ETH Zurich,
Zurich, Switzerland*

Abstract

Purpose – The purpose of this paper is to investigate the feasibility of an end-to-end simplified and automated reconstruction pipeline for digital building assets using the design science research approach. Current methods to create digital assets by capturing the state of existing buildings can provide high accuracy but are time-consuming, expensive and difficult.

Design/methodology/approach – Using design science research, this research identifies the need for a crowdsourced and cloud-based approach to reconstruct digital building assets. The research then develops and tests a fully functional smartphone application prototype. The proposed end-to-end smartphone workflow begins with data capture and ends with user applications.

Findings – The resulting implementation can achieve a realistic three-dimensional (3D) model characterized by different typologies, minimal trade-off in accuracy and low processing costs. By crowdsourcing the images, the proposed approach can reduce costs for asset reconstruction by an estimated 93% compared to manual modeling and 80% compared to locally processed reconstruction algorithms.

Practical implications – The resulting implementation achieves “good enough” reconstruction of as-is 3D models with minimal tradeoffs in accuracy compared to automated approaches and 15× cost savings compared to a manual approach. Potential facility management use cases include the issue and information tracking, 3D mark-up and multi-model configurators.

Originality/value – Through user engagement, development, testing and validation, this work demonstrates the feasibility and impact of a novel crowdsourced and cloud-based approach for the reconstruction of digital building assets.

Keywords Facility management, Smartphone, Digital assets, Design science research, Cloud BIM, Crowd-sourcing

Paper type Research paper



© Srinimalan Balakrishnan Selvakumaran and Daniel Mark Hall. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

Journal of Facilities Management
Vol. 20 No. 3, 2022
pp. 401-436
Emerald Publishing Limited
1472-5967
DOI 10.1108/JFM-02-2021-0017

1. Introduction

Inefficiencies and lack of interoperability in facility management (FM) are estimated to cost around US\$15.8bn annually in the USA alone (Yalcinkaya and Singh, 2014). Much of this cost is associated with re-creation and re-entry of building data. At the completion of a project, handover of paper-based documentation and “as-built” drawing information can take several weeks. An inordinate amount of time is required by the FM team to verify facility and project information. Furthermore, older buildings lack any sort of digital representation that can be used during operations and maintenance. FM must source information from paper-based documentation and blueprints that can be inaccurate or difficult to find.

In search of digital representation for use during FM, researchers and technology providers have developed processes to automatically capture and extract FM-specific geometry and attributes from existing buildings. This is because the manual reconstruction of three-dimensional (3D) models can be unreasonably difficult, time-consuming and error prone. As an alternative, semi-automated workflows generate point cloud data using laser scans and convert them manually to as-built building information models (BIMs). While found to be accurate, the process is expensive and demands extremely skilled labor (Xiong *et al.*, 2013). In comparison to the typical digital skillset of a facility manager and end-user, these automated approaches are far too technical and computationally intensive.

The objective of this paper is to develop a simplified and intuitive approach to generate as-is BIM models. To accomplish this, we follow a design science research (DSR) methodology. DSR is driven by deep engagement with the real-life problems or opportunities of potential users (van Aken *et al.*, 2016). By providing a structured approach to designing a solution to fulfill a purpose, testing and validating that the solution does serve the intended purpose and then improving the solution as needed (March and Smith, 1995), DSR represents one method to conduct “bottom-up” engineering informatics research in conversation with engineering practitioners (Hartmann and Trappey, 2020).

In this paper, the DSR methodology led to consideration and eventual combination of two emerging phenomena. The first is the use of cloud-based BIM. Cloud-BIM is the integration of growing cloud-computing technologies with the BIM practices enabling more collaboration and dynamic real-time communication among all stakeholders (Bello *et al.*, 2021). The second is the concept of crowdsourcing. Crowdsourcing is a method of involving a group of people toward a common goal (ideas, financing or work) through a chosen medium, such as the internet, social media or mobile applications (Estellés-Arolas and González-Ladrón-de-Guevara, 2012). The combination results in the development of a crowdsourced workflow to generate “as-is” BIM models using smartphones as the capture device, as will be described later in the paper.

The paper proceeds as follows. First, we review the existing literature on as-built and as-is BIM, including the practical limitations to create these models using laser scanning and image reconstruction approaches. Next, we follow the DSR approach to define the problem further, define a set of solution objectives, design and development the “crowd-to-cloud” workflow, demonstrate and test the workflow for participant use through practical implementation, and evaluate the tool for future implementation. The paper concludes with a discussion of the implications of the proposed workflow and limitations to the research.

2. Departure

BIM is a digital representation for physical and functional characteristics of buildings that has increasingly been accepted in the Architecture, Engineering and Construction (AEC) sector. BIM offers significant benefits to FM, including improved processes for restoration, documentation, operation, maintenance, quality control, energy and space management

(Macher *et al.*, 2017; Yi-Kai and Nai-Pin, 2017). However, despite emerging BIM processes for FM, most existing buildings are not maintained, refurbished or deconstructed with BIM yet (Volk *et al.*, 2014).

2.1 From as-built to as-is building information model

One reason for the lack of BIM in FM is the lack of “as-is” models that are available. The creation of a design BIM during the design stage of a new project has become common. By contrast, the handover of an as-built BIM – a BIM that accurately reflects the facility’s geometry and attributes after the completion of construction on-site (i.e. “as-built”) – is much less common. Even when an as-built BIM is provided, subsequent changes and renovations mean the as-built BIM might not reflect the current state of the building “as-is” (Pătrăucean *et al.*, 2015). Past scholarship has continuously insisted on the necessity to move from as-built to as-is BIM due to significant variations between the constructed and current condition of a building, subsequent changes and renovations arising to buildings over time, absence of design BIM and/or the lack of blueprints (Jung *et al.*, 2015; Tang *et al.*, 2010; Xiong *et al.*, 2013). Therefore, facility managers who wish to operate and manage facilities using an as-is BIM (Dore, 2014) must face the challenging but necessary process to construct and update data models continuously.

To help FM with this updating, there has been a consistent effort in automatically generating as-is BIMs for existing facilities. This workflow includes four stages: data acquisition, 3D reconstruction, 3D modeling and application (Figure 1). Early approaches to as-is BIM used laser scans to generate point clouds and then used 3D reconstruction algorithms and/or manual methods to create 3D models. Recent scholarship has focused on the photogrammetric translation of indoor/outdoor scenes to semantically enriched 3D models. We briefly summarize the two approaches – laser scanning and image reconstruction – below, including a critical perspective on limitations with respect to the FM end-users.

2.2 Laser scanning

Laser scanning systems are the most adopted and well-experimented technologies for data acquisition of as-BIMs in the AEC industry. Laser scanners capture point cloud data with speeds of 1,000–100,000 point measurements per second (Tang *et al.*, 2010) and with great accuracy [1] (~12 mm at 100 m) and point density (Bosché, 2009). Laser scanned point clouds are the simplest form of 3D data representing coordinates and texture of the real-world scene. From here, it is possible to manually model the building using the laser-scanned point clouds merely as visual support (Macher *et al.*, 2017). However, this does not take advantage of the effort or cost to capture the data and has been classified as time-consuming, error prone and demanding skilled labor (Wang and Cho, 2014).

A more effective 3D reconstruction of the point cloud data uses segmentation algorithms. The acquired raw point clouds must first undergo noise removal and structuring. De-noising and outlier removal methods include (Ning *et al.*, 2018) the discontinuous operator-based method (Wang *et al.*, 2013; Zhang *et al.*, 2009; Huang *et al.*, 2010) and the surface fitting-based method (Weyrich *et al.*, 2004). Next, segmentation methods group points into subsets



Figure 1.
3D reconstruction of existing buildings summarized as an ideal workflow

defined by common characteristics and classification assign these points to specific labels according to different criteria. Post segmentation, automatic modeling has been previously achieved using heuristics (Pu and Vosselman, 2006; Rusu *et al.*, 2009), ontologies (Ben *et al.*, 2012), context and prior knowledge (Guo *et al.*, 2014).

Although research in this field has demonstrated success and potential, industry adoption has been limited by the cost. This technology suffers from high initial equipment costs – previous estimates range from US\$50,000 and US\$100,000 for standard versions (Bosché, 2009) – making it infeasible for small- and medium-sized projects. The usage of the equipment takes time to learn and requires skilled knowledge workers. Usually, scanning from a single location cannot cover the entire facility and scans must be obtained from multiple locations, consequently increasing the capturing time spent on-site. Furthermore, the registration of point clouds and de-noising the data needs additional manual effort and checking. 3D reconstruction algorithms require fine-tuning of different parameters according to the captured scene.

2.3 Image reconstruction

Image-based technologies in practice can be divided into: Photogrammetry and Videogrammetry. Photogrammetric techniques are adopted for 3D restitution of scenes using images taken from different points of view. These techniques depend on the geometry or surface characteristics of building elements. Image-based technology holds the advantages of being automatic and having low equipment cost, fast on-site data acquisition and high portability. However, it has the downsides of reduced accuracy and its unfamiliarity to existing surveyors (Dai *et al.*, 2013).

3D reconstruction for images requires computational algorithms to transform images into a 3D point cloud/mesh. The resulting point cloud is unordered and unstructured. Detection of geometric primitives (points, patches, edges and lines) is the fundamental step in identifying objects in images. Image processing researchers have proposed several algorithms, for example, SIFT, SURF, MSER and DAISY for image registration, image differencing and image recognition (Lu and Lee, 2017). Using these algorithms as the base, computer vision and civil engineering scholars developed methods to identify and recognize buildings.

At the core of all methods, the task is to correlate some sort of visual information (regions, edges, lines or points) between the image-pairs in the collected data set. Among all methods, the most widely used and representative for building recognition is image-based point cloud generation. Structure from Motion (SfM) is a method to recover 3D scenes (point cloud) from unordered two-dimensional (2D) images. Several open-source implementations of SfM, such as Bundler [2], OpenMVG [3], VisualSfM [4] and Colmap [5] are available. SfM has been successful in achieving automatic image-based construction monitoring (Golparvar-Fard *et al.*, 2011), image-based indoor navigation (Dong *et al.*, 2015; Gao *et al.*, 2014; Noreikis *et al.*, 2018) and indoor reconstruction (Lehtola *et al.*, 2014). The resulting 3D point cloud from a SfM can be reconstructed into a mesh using one of several available meshing algorithms, such as Poisson Surface Reconstruction, Smooth Signed Surface Distance reconstruction and Floating Scale Surface Reconstruction.

Modeling for image-based point clouds can be categorized as model-driven methods and data-driven methods. Model-driven methods rely on detecting differences between manually created models and structural relationships for object recognition. Data-driven methods (Guo *et al.*, 2014) for image-based point clouds use captured data and its features for segmentation. Image-based 3D reconstruction has been disregarded in the past due to accuracy trade-offs compared to laser scanners. With improvement in algorithms to

reconstruct image-based point clouds, this field has been gaining attention recently. Photogrammetry is considered a suitable method when the scenes to be captured contain a lot of texture and gradient. Laser scanners are adopted when accuracy is the priority. Another major drawback of attaining image-based point clouds is the intense computational power consumed by the reconstruction algorithms.

2.4 Design science research

DSR concerns itself with the designing and making of artefacts to fulfill a purpose, testing and validating that those artefacts serve the intended purpose and do not have unintended consequences, and then improving the artefacts as needed (March and Smith, 1995). Design science research (DSR) is driven by field problems or opportunities because it believes instrumental knowledge is developed by deep engagement with these real-life problems or opportunities (van Aken *et al.*, 2016). DS has been applied in multiple domains related to this paper, including construction management (Tommelein, 2020; Voordijk, 2009), construction informatics (Chu *et al.*, 2018; Fang *et al.*, 2018), information systems (March and Smith, 1995; Peffers *et al.*, 2007) and operations management (van Aken *et al.*, 2016).

Hartmann and Trappey (2020) have called for researchers in advanced engineering informatics to approach problems in a “bottom up manner closely involving engineering practitioners” while making use of social science methods to help formalize and validate knowledge. Bottom up “co-creation” of digital services for FM requires dialogue between the project parties and a shared understanding about both the FM’s needs and the way that a digital service can satisfy those needs (Lavikka *et al.*, 2017). In particular, there is a need to explore advanced computational methods in combination with novel information acquisition systems including understanding practitioner needs for graphical user interfaces, workflows for knowledge capture using cloud-based technologies and methods to support engineers and FM teams in knowledge intensive tasks. DSR offers a structured approach to uncover these needs from FM users, formalize solution targets, design a solution prototype and then demonstrate and evaluate the proposed workflow.

2.5 Research gap

One limitation of the modeling algorithms discussed in the previous sections is the investigation of only partial segments of the complete workflow, with limited involvement or consideration of the work routines and needs of facility managers. Research interest has focused on the best geometric representation and increasing the speed or efficiency of the algorithm. In addition, modeling approaches are developed for working locally, which requires powerful machines/clusters. In a practical setting, for a firm working on a smaller project, this would mean a high investment cost which overshadows the benefit from such a reconstruction.

The required efforts to create an acceptable as-is BIM from these workflows can overshadow the benefits of the process (Kim, 2012). The time consumed to scan large buildings is very high and structuring the data can still require manual processes resulting in high costs. Furthermore, there is a need to understand and analyze the entire chain that leads to an enriched 3D model (Hichri *et al.*, 2013) beneficial to the end FM user. Much of previous research has emphasis on methods, algorithms and techniques to extract a geometrically and semantically rich BIM model. These have resulted in successful technical breakthroughs, but we note the absence of attention to a complete workflow from data acquisition to value-added application in the field.

In summary, there is a gap between “top-down” algorithmically-strong methods with “bottom-up” industry-ready workflows. Ideal 3D digital asset reconstruction should be

economical, efficient, accurate, and practical (Lu and Lee, 2017). Furthermore, scholarship finds additional benefits can occur when new technologies and data workflows are co-created alongside existing FM teams (Lavikka et al., 2017). Critics of BIM find it to act as a standalone system framework that restricts project stakeholders' access to a common set of data (Chuang et al., 2017). As an alternative, the concept of cloud-based BIM as an emerging technology is needed to enable higher levels of cooperation, collaboration and provide an effective real-time communication platform for involved stakeholders (Li et al., 2014).

3. Research approach

To address this gap, we follow the DSR methodology to develop a crowdsourced, cloud-based approach to create as-is BIM models. First, we conduct semi-structured interviews with two groups of stakeholders to understand the specific limitations of BIM for FM in today's practice. Second, we use a series of exploratory tests to develop solution targets regarding automation, efficiency, accuracy, practicality and scalability. Third, we design the initial prototype including the development of the mobile client, cloud architecture and Web client. Fourth, we conduct an experiment to demonstrate implementation of the prototype. Finally, we evaluate user and expert feedback and propose potential further use cases based on this feedback. The specific DSR process is summarized in Figure 2.

3.1 Problem identification and motivation

To better identify the specific problem, we began the research by conducting 12 semi-structured interviews lasting between 30 and 45 min each. This dialogue includes experts from two distinct fields (Table 1): AEC and computer science (CS). All interviewees were identified through industry partners and research labs, contacted digitally by the first author.

AEC interviewees were chosen for their expertise in FM, energy retrofit and spatial development. These interviewees worked during the operations and maintenance phase of

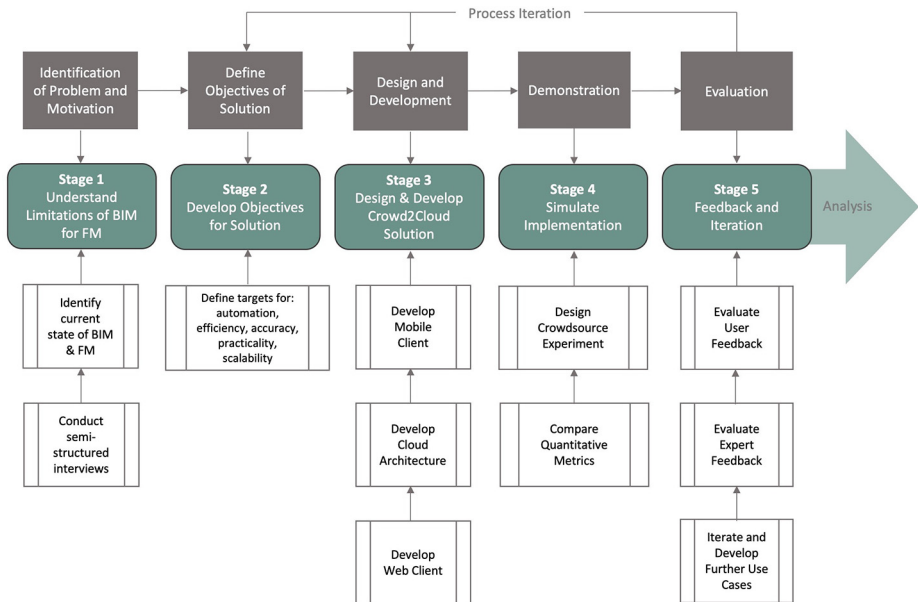


Figure 2.
DSR flowchart
material

the building lifecycle; therefore, they would most benefit by solving the observed problem (Franz *et al.*, 2018). During the interviews, AEC experts were asked if the lack of a digital asset is a primary issue, their needs and problems on regular projects, their current workflow to tackle the issue, the potential use cases, knowledge on reality capture and surveying methods and their readiness to adopt an innovation. In addition, experts from geoinformatics and surveying were also interviewed to understand the current surveying technologies in the AEC industry.

From CS, the interviewees were chosen for their expertise in computer vision, reality capture and 3D reconstruction. During the interviews, CS experts were questioned about reality capture methods to reconstruct building-related scenes, the extent of current work, implementation in the field, their knowledge on the needs of the AEC industry and if they were aware of any application focused pipeline.

Data collection occurred through note-taking by the first author. These notes were compared and contrasted to develop a set of common responses regarding the demands and problems of FM (AEC) and the state-of-the-art solutions (CS). The quotes selected for presentation in the paper were later shown to the interviewees for confirmation that they represent the speaker’s views. A synthesis of the main problem identification is shown in Table 2.

3.1.1 *Synthesis of architecture, engineering and construction expert knowledge.* AEC experts first stated an *urgent need for digitalization* of the existing building stock. One

Designation	Specialization	Industry	Experience (years)
Project Leader	FM	AEC	13
Department Head	Energy	AEC	10+
Project Leader	Energy	AEC	5+
Project Leader	Energy	AEC	4+
Project Leader	Energy	AEC	4+
Senior Engineer	Buildings	AEC, CS	4
Department Head	Geoinformatics	AEC, CS	10+
Project Leader	Surveying	AEC, CS	8+
Senior Researcher	Computer vision research	CS	3+
Researcher	Computer vision research	CS	2+
Researcher	Computer vision research	CS	2+
BIM Leader	Digital innovation	AEC, CS	10+

Table 1.
Profile of interview participants

Topic	AEC experts	CS experts
Digitalization	Urgent need for digitalization. Many projects are existing buildings. Creation of as-is BIM would be very helpful	Many state-of-the-art methods exist today to digitalize existing building stock
Accuracy	Approximation runs their daily job. Only need basic geometrical insights; there is no need for “perfect BIM model”	Research today is highly focused on better algorithms to improve accuracy
Learning	Too much learning required to develop their own workflows. There is need for “one-click solution”	Present algorithms could be modified according to the needs of practitioners, but this must come from AEC
Data	Often collected but not used or shared effectively	

Table 2.
Synthesis of findings from semi-structured interviews

participant stated, “more than 90% of our projects are existing buildings” and “a good deal of them are at least 50 years or older.” If existing assets could be digitized, it would “make their lives easier.” In particular, there is a benefit to having the as-is digital model during the concept phase to understand the current and potential functionality of the spaces. The current documentation is frustrating and there needs to be a “connection between paper and digital model of the asset.”

Next, AEC experts described how *approximation runs their daily job*. As opposed to the state-of-the-art research, which focuses on high accuracy for as-is BIM, practitioners do not need a “perfect BIM model.” Instead, the need is “very basic geometrical insights such as volume of the room, area of windows and room relationships,” which “could add a lot of value.” Most typical work is approximated via hand calculations and excel sheets, with low levels of automation, so reconstructed digital assets with high levels of accuracy are simply not needed at this time.

AEC experts described how the new techniques *demand too much learning* and wished there was a one-click solution. One engineer who was experienced with scanning-based modeling algorithms stated that “the learning time and effort was way too high for a planner to go through it.” From a cost perspective, one MEP engineer with point cloud experience voiced displeasure with the “high number of hours to manually model the structure and elements.” Overall, there was no clear and simple “one-click solution” without disrupting the daily routine of planners and facility managers.

Finally, AEC experts explained that *data is collected but not used effectively*. Many described a recent shift of techniques from manned surveying to laser scanning. Many had conducted experimental scanning with terrestrial laser scanner and drones but seldom had they successfully passed on this data as part of a meaningful project. One interviewee described that the existing in-house laser scanner has been “hardly used.” An expert in energy modeling claimed that “the collected data becomes unusable if it is only acquired for one specific purpose and restricted in modifications.”

3.1.2 Synthesis of computer science expert knowledge. CS experts first described how *many state-of-the-art methods* for reality capture and reconstruction of real-world scenes have been developed or are in development. CS experts described how the reconstruction of scenes related to buildings are among the most complex systems to reconstruct and therefore are of high interest for computer vision scientists. In contrast to AEC experts, these algorithms are so well-established in the CS domain that the researchers could mention several techniques and research groups without hesitation.

Next, CS experts viewed existing CS work as *highly-focused on algorithm development*. A researcher in semantic segmentation mentioned that “most current research is focused on optimizing existing algorithms in terms of accuracy, speed, efficiency and computing power. Research work toward an entirely new framework is very rare”. Another participant expressed his view that “to work on a reconstruction algorithm for building-sized structures is really time-consuming and painstaking” and that the present algorithms could be modified according to the needs of practitioners.

Finally, CS experts expressed curiosity at the *absence of service-oriented tools* in AEC. There were curious questions about the current state of practice in AEC, such as one CS expert asking, “Why do construction practitioners use a laser scanner instead of a total station?” Another interviewee brought up the “absence of complete workflows to install on the engineer’s desk.” Although existing algorithms are extremely advanced, they do lack an industry-specific workflow. However, this adaptation should come from AEC as CS experts would not know the specific needs of industry practitioners.

3.2 Define objectives for a solution

From the above synthesis of the interviews, a preliminary set of objectives was defined to explore the potential workflows to automate reconstruction for FM practitioners. To do so, potential workflows were evaluated based on the following objectives:

- *level of automation*: least manual intervention and potential for “one-click” solution;
- *efficiency*: time, cost and computing load;
- *accuracy*: degree of difference between model and target scene;
- *practicality*: user-friendly, adaptability and customization; and
- *scalability*: generic, large-scale implementation possibilities.

A series of initial tests were performed on two different potential workflows. The initial testing focused on identifying a suitable transformation (input to output) of the development pipeline. Both workflows were built on the same underlying algorithm, SfM and Multi-view stereo. The first reconstruction pipeline, Colmap, is an open-source research tool which takes 2D images as input and reconstructs a dense 3D point cloud. The second pipeline, Autodesk Forge ReCap API is a cloud-based reconstruction engine API from a commercial software provider in the AEC industry. This API allows users to custom build tools for reconstructing 2D images into a 3D mesh.

Two different indoor typologies, a corridor and an office room, both with maximum clutter were tested to understand the capabilities and limitations of the workflows. The acquisition, computing, cleaning times and the result mesh were analyzed. An Image Acquisition Manual was defined specifying guidelines for a user on best practices for image capture to generate significant 3D meshes/point clouds. Incorporating the learnings and addressing limitations of the two methods, a further field test was carried out in an indoor auditorium to provide generality and draw strong conclusions for further design. A common drawback of both methods was inefficiency in reconstructing windows, reflective and texture-less surfaces. This was addressed by testing the auditorium as both non-textured (as-is) and textured (manually adding colored notes) scenes. Predefined control points were introduced to the indoor scenes, measured and used to analyze the geometric accuracy of the reconstructed results. A detailed description of these initial tests and results are available in the supplemental material.

From these initial tests, the scope and idea for the solution were further iterated to consider that the model should be crowdsourced and cloud-based. From here, the specific requirements needed for a solution could be more precisely defined.

3.2.1 Mobile client for data collection. Multiple users should be allowed to collect image data when and as they wish, including simultaneous capture. Therefore, the workflow has to include a user-friendly mobile client (application). This smartphone application must provide access to different users and instruct them on how to gather images for the best possible reconstruction.

3.2.2 Data visualization. To facilitate a structured collection of consecutive images, all users must be able to visualize the current state of reconstruction of the zone/room. On this ground, the developed smartphone application needs to provide means to visualize collected images and reconstruction results.

3.2.3 Cloud processing and computing. A crowdsourced reconstruction process based on the input of multiple scanning devices requires continuous aggregation and processing of scanned data (Franz *et al.*, 2018). Since mobile devices have restricted capabilities and setting up a local server would mean additional hardware expenditure and maintenance, a

cloud processing server is considered best. Staging all the computing to a remotely located cloud also provides flexibility for parallel processing.

3.2.4 Network channel/communication infrastructure. Users with mobile devices must have continuous access (upload and download) of all acquired and reconstructed data. The stability of the network infrastructure is of prime importance to enable multi-directional data exchange between the client and the cloud. The system should be designed to work on both an independent network long term evolution and established Wi-Fi.

3.2.5 Hardware equipment agnosticism. The target is beyond users with specific skill sets or equipment. The workflow should be designed agnostic to the type of smartphone (model and make), operating systems (Android, iOS) and should be scalable for images taken on any other device.

3.2.6 Web client (or mobile client) for application. The goal of such a reconstruction workflow goes beyond just the data collection and processing. Users should be able to use the reconstructed result, use it for their needs and solve routine problems they face. This insists on the requirement of a user-oriented functionality to add/modify and delete data onto the model. The developed concept must include a Web-based application (and/or) an interface in the mobile client to add such functionalities.

3.3 Design and development

As the development is hardware independent and can be deployed on any smartphone supporting either Android or iOS operating system, this section focuses more on the software architecture (Figure 3) of the framework. A conceptual framework was designed based on the objectives and requirements (Section 3.2). The proposed framework consists of three major modules: Mobile Client, Cloud Architecture and Web Client. The mobile application developed would incorporate features for crowdsourced image data collection and dynamic visualization. The cloud computing platform should handle uploading collected images to the reconstruction engine and translate reconstructed meshes for visualization. The external Web application should dynamically retrieve reconstructed meshes and provide customized functional features. The reconstruction engine chosen as

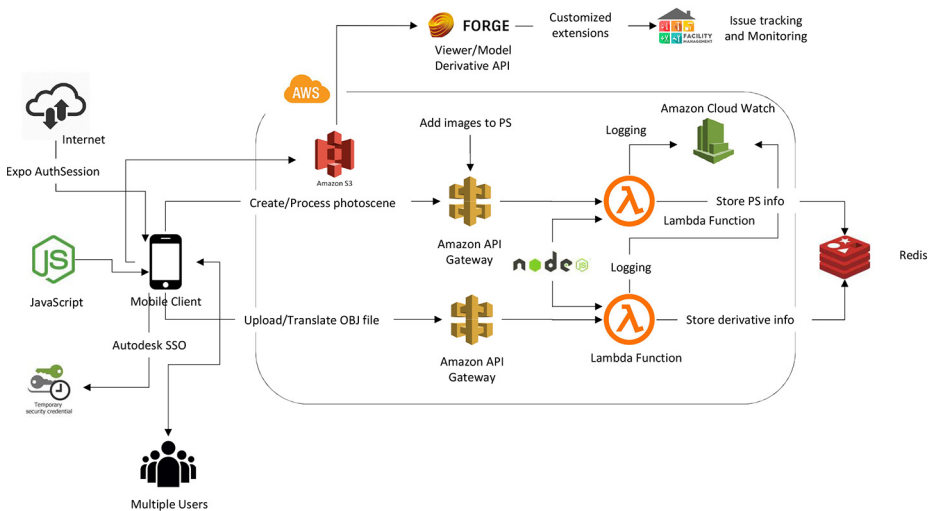


Figure 3.
Application
architecture

ReCap API, it was reasonable to build the entire development around the comprehensive Forge API. The defined framework was proposed for an Autodesk Forge Accelerator [6] and Autodesk Forge developers supported the first author with the application development during the five-day intensive AEC hackathon.

From this framework, a specific implementation was developed and prototyped. Below outlines algorithms, technicalities and source code that were developed. The inspiration for the source code of the mobile application and the cloud architecture was from public development repositories published by Autodesk Forge Specialists. This specific prototype includes additional features, such as multi-user data collection through the smartphone application, dynamic retrieval and visualization of partially reconstructed results by users at any point of time, adding new set of images to an existing photo scene, interactive instruction screen for guiding the user experience, Web-client interfacing with the storage to retrieve reconstructed data and three distinct Web applications with identified use cases.

3.3.1 Mobile client. To bypass double work in mobile implementation (Android and iOS), Expo.io and OS agnostic integrated development environment (IDE) was preferred. The mobile client was scripted on JavaScript supported by out of the box packages from Expo. Expo is an IDE that enables developers to create React Native [7] applications that can run in any native environment containing the Expo Software Development Kit.

The construct of the mobile client was based on some general considerations:

- user-friendliness;
- security;
- visualization/user guidance;
- user instructions;
- API management;
- network; and
- backend.

Conceptually, the first step was to authenticate the user to access the secure mobile application which was achieved by using Autodesk Single Sign-On.

The demand for a user-friendly experience designed the succeeding steps:

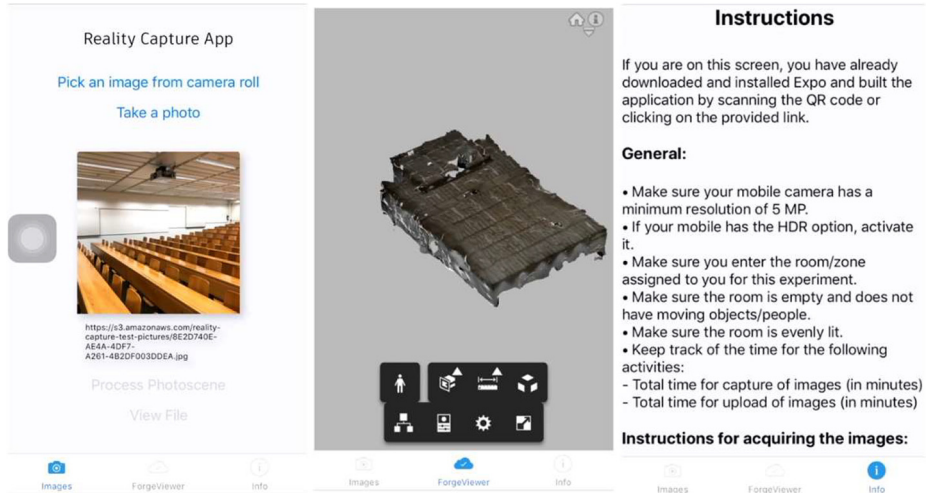
- The users are expected to select captured images from their mobile camera or album and once satisfied.
- The users are expected to send the selection to Autodesk Cloud (hosted on Amazon Web Services [AWS]) for processing.
- On completion of the image set processing by the Autodesk ReCap API, the user should be notified that the reconstructed mesh data (3D model) is ready to be retrieved through the Autodesk Model Derivative API and visualized in the viewer on the mobile client.

Steps 1 and 4 are implemented locally on the device. Steps 2 and 3 overlap with the backend functionalities demanding network related activities (discussed in Cloud Architecture).

As the workflow depended on crowdsourcing, the users of the application had to be instructed on how exactly the images had to be captured. The application reflected this requirement by including an Instruction Screen (Figure 4). These instructions directly adopted the learnings in the “Image Acquisition Manual.”

The mobile application was devised to perform simple network activities like image uploads, processing requests and visualizing retrieved data. Intensive processing was

Figure 4.
Screenshots of the
reality capture
smartphone
application prototype



moved to the cloud infrastructure to ensure reliability (Cloud architecture). Such a workflow has precedence (Franz *et al.*, 2018) and proved beneficial to evade network issues and lower computing capabilities of a mobile device.

3.3.2 Cloud architecture. As detailed in the previous section, Steps 2 and 3 of the application concept demanded strong network-related activities. It was decided to adopt this as [Node.js](#) backend service apps:

- The first app handled creating, processing and deleting the Reality Capture photo scene (bucket of uploaded images) on request by the user through the mobile client.
- The second app included code for handling the reconstructed (by the Forge reconstruction engine) output in the Autodesk Object Storage Service and translating it into a Forge viewable format, .svf.

Prior works have successfully hosted intensive processes on a single server backend built and maintained on powerful local machine/s. The disadvantages, though were setup and management of the servers, hardware requirements and specificity of the backend service.

The particularity of currently implemented architecture is the choice of Serverless Computing. Serverless Computing is a cloud-computing execution model in which the cloud provider runs the server and dynamically manages the allocation of machine resources. The strength lies in the fact that any type of backend service can be deployed with no administration troubles and pay-per-use pricing. It was of prime importance to adopt this new cloud technology that could strengthen the ease of adoption and positively influence a potential business model.

AWS was picked as the cloud provider due to its comprehensive support on Serverless Computing and successful integration with the Autodesk Forge infrastructure. Four distinct serverless computing modules were incorporated (Figure 3):

- *AWS Lambda functions:* AWS Lambda is an event-driven, serverless computing platform provided by Amazon.
- *AWS API Gateway:* is an AWS service for creating, publishing, maintaining, monitoring and securing REST and WebSocket APIs.

- *Amazon S3*: Amazon S3 is a “simple storage service” offered by AWS that provides object storage through a Web service interface.
- *Claudia JS*: Claudia is a JavaScript library that handles deployment, modification and destruction of AWS Lambda functions.

3.3.3 Web client. The Web client is not an integral part of the smartphone application but rather a use-case-based Web application which retrieves the reconstructed mesh data from the same cloud infrastructure. Three distinct modular but integrable use cases were identified for facility managers (primary target, but could also be used with slight modifications by other disciplines) to use the reconstructed 3D mesh model:

- *Issue and information tracking*: dynamically add issues and information to a simple 3D mesh model on the Web. Issue monitoring panel enlisting current and closed issues and meta-properties panel (for static text, image, link, document formats) were the two major components to ensure smoother workflow collaboration.
- *3D mark-up*: a tool for on-site personnel to record as-is condition by attaching images to created pushpins/markups on the model. Each mark-up associated itself with a customizable icon (Issue, request for information, Status and Warning) and adopted to the camera position of the user, ensuring smooth rendering on mobile devices.
- *Multi-model configurator*: space management tool for dynamically modified spaces (meeting rooms, cafeteria and conference) to efficiently test varying configurations on the Web (move and play) and aid in decision-making.

All the applications were built on a Forge Viewer. The Viewer is a WebGL-based JavaScript library for 3D and 2D model rendering. Concerning the Web client for the prototype, a Web viewer was developed as a [Node.js](#) application scripted on JavaScript. The Web application consists of both client-side HTML/CSS code responsible for the Web interface and server-side code handling the backend and API requests. All the use cases were developed as modular viewer extensions. Viewer extensions are JavaScript classes written specifically by the developer to achieve customized functionality.

3.4 Demonstration

As demonstration of the artifact, an experiment was devised to understand if the proposed tool could solve the observed problem when used by non-experts. The experimental design is described in [Table 3](#). Ten participants were chosen in total through an expression of interest form. The first five were put into the unguided crowdsourcing group and the rest into guided crowdsourcing. A regular seminar auditorium was chosen as the case study for this experiment. Selected metrics, both quantitative and qualitative were obtained and analyzed by conducting the experiment.

Irrespective of the group, each participant was responsible for collecting exactly 40 pictures. This number was identified from initial tests as an intuitive minimum to achieve

Experiment design criteria	Unguided crowdsourcing (G1)	Guided crowdsourcing (G2)
#participants	5	5
#images per participant	40	40
Guidance	No	Yes
Information screen	Yes	Yes

Table 3.
Crowdsourcing
experiment design

partial reconstruction (geometrically meaningful). Participants in G1 were asked to follow the instruction screen (Section 3.3.1) and take pictures without any further guidance. G2 participants could additionally access the Viewer screen on the mobile application and visualize the partially reconstructed model from the previous participant in their group. This difference was introduced to analyze the advantage of visual guidance in crowdsourcing, as the reconstruction on the cloud was dynamic.

The perimeter of the room was measured manually using a meter scale. Other dimensions, such as wall length, height and door dimensions were also measured. Two fixed distances, 1.15 m each, were marked on opposite walls of the auditorium using “X” marks on paper (Control points), remaining unchanged for all participants. It was made sure that the room was empty during the capture and any movable objects located in its original position. The lighting conditions were maintained the same.

The participants were required to attend an online survey after completing their experiment. The questions were related to the metrics of analysis and usage of the smartphone application.

For both the guided and unguided activities, the 3D reconstruction was triggered in steps of 40 (40, 80, 120, 160 and 200). The reconstructed mesh model was saved after every consecutive step and analyzed as follows:

- *Quality of mesh with respect to the number of images:* Mesh size, Mesh faces and Mesh points.
- *Boundary fulfillment:* Percentage of actual perimeter reconstructed in every step.
- *Total time:* Sum of time of capture and time of upload.
- *Geometric accuracy:* Distance between the control points on the model versus the manually measured distance.

The 40 images from each participant were also reconstructed individually without cumulating to evaluate the aggregation effect and modularity of the partial models.

The mobile devices (4 Android, 6 iOS) used by the participants were at least 12 MP in camera resolution or higher. The entire participant set was covered over a period of four weeks according to the availability of the auditorium.

For both groups, the captured 40 images were uploaded by the participant in the relevant storage bucket. The core analysis was the consecutive [8] reconstruction with additional 40 images in every step.

G1 participants required an average time of 4.4 min to capture 40 images each compared to an average capture time of 5.5 min by G2 participants. The higher time of capture for the guided group is justified by the time taken to refer to the partially reconstructed model, comprehend the mesh and identify locations to take further pictures.

The time to upload depended on the participant’s affinity to interact with their phone. There was no identified pattern differentiating the groups. However, Android users required slightly more time than iPhone users due to the user interface design (two pictures per row) of their image gallery. It took an average of 15.8min and 16.4min by G1 and G2 participants, respectively, to upload 40 images each.

The reconstructed meshes from individual participants (40 images) and cumulative image sets (40, 80, 120, 160 and 200) for G1 and G2 are presented in Figures 5 and 6, respectively. It can be observed that the individual mesh reconstructions in unguided participatory crowdsourcing are very random, reconstructing almost nothing (P1 and P5) to achieve a complete boundary (P3). It was observed there was no order in the collection of images and hence the resulting meshes. The cumulative results showed that a highly

Individual reconstructions

40 40 40 40 40



Cumulative reconstructions

40 80 120 160 200



From crowd to
cloud

415

Figure 5.
Results of unguided
participatory
crowdsourcing

Individual reconstructions

40 40 40 40 40



Cumulative reconstructions

40 80 120 160 200



Figure 6.
Results of guided
participatory
crowdsourcing

satisfactory result (geometric accuracy and boundary fulfillment) was achieved at 120 images. Images added further produced redundant meshes containing bumpy surfaces and poor feature matching. This can be attributed to excessive similar images containing too much overlap which are not required by the algorithm. This redundancy did not intrinsically reduce the effective reconstructed area. Of worth mention is a manual error by P5, where the participant stood in the center of the room and took panoramic shots of the room. The instructions required the participants to stop along the perimeter and capture the opposite side of the room. This resulted in individual P5 reconstruction (40 images) to fail. However, the images captured from the center of the room did contribute to the cumulative reconstruction providing some overlap to existing holes.

The guided participatory crowdsourcing in which the participants had the choice to refer to the previously reconstructed mesh (partial) produced modular meshes (Figure 6) from individual reconstructions (40 images) which fit as perfect pieces to complete the auditorium. The cumulative reconstruction built up slowly on the back (P1 + P2), on three

sides (P1 + P2 + P3) and finally the complete auditorium (P1 + P2 + P3 + P4). Images from P5 reconstructed an individual mesh but did not contribute to the cumulative model because the participant captured too much of the ceiling without overlap to the existing model. This participant was indeed expected to capture the ceiling, as that was the only missing part in the previous mesh. However, the visual guidance also needs to align with the prerequisites of the algorithm itself to improve the results.

The reconstruction time (Figure 7) on the cloud for the cumulative image sets varied significantly between the two groups. In the unguided scenario, the number of minutes required increased with an increasing number of input images, peaked at the third participant (120 images) and then decreased for 160 and 200 images. Again, this trend is rationalized by the redundant images after 120 inputs. The algorithm dropped out/did not match all the images and consumed less time. The mesh size (Figure 8) of the unguided group also followed an identical trend. It was interesting to note that this reduction in reconstruction time and mesh vertices did not affect the geometric accuracy (Supplementary material) or the completeness of the reconstructed model. However, it brought forward the sensitivity of the reconstruction engine to the images being added. It can be safely concluded that a higher number of unordered images does not always mean a more accurate reconstruction.

In the guided scenario, the cumulative reconstructions required increasing time (Figure 7) with increased inputs. The mesh sizes (Figure 8) increased linearly for every consecutive model. The data for the 5th participant is excluded as the images added failed to contribute to the cumulative reconstruction. The observations from guided participatory crowdsourcing conclude the effectiveness of visual guidance in reconstructing a satisfactory mesh step by step as modular partial models. The total reconstruction time for 160 images in the guided scenario was 42.4 min compared to 95.7 min in the unguided scenario. This was additional validation to the easier and faster processing of guided and ordered images.

Boundary fulfillment (Figure 9) of reconstructed models were calculated as the percentage of the total perimeter covered by each cumulative reconstruction. The model



Figure 7.
Total reconstruction
time for both the
groups

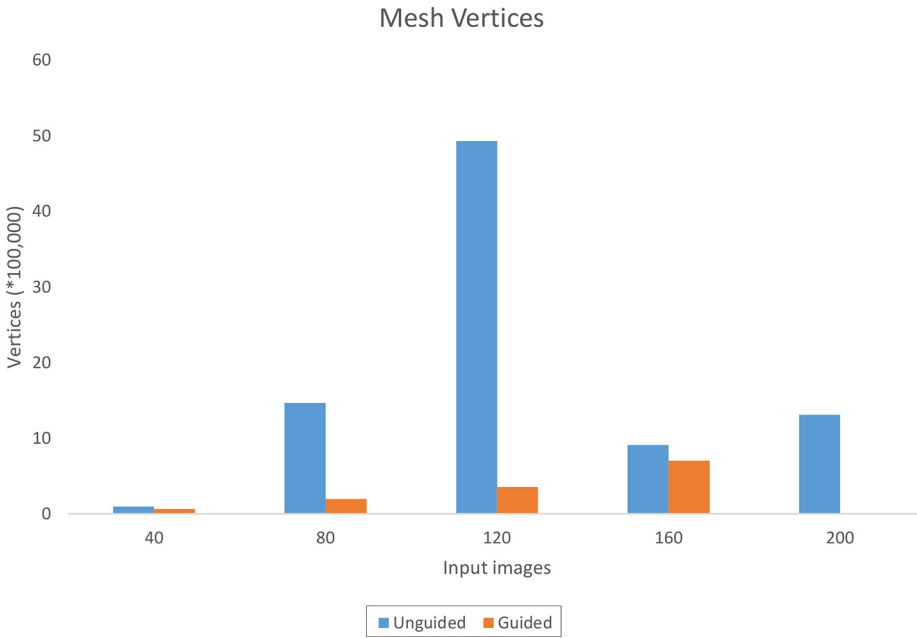


Figure 8.
Size of the reconstructed 3D mesh

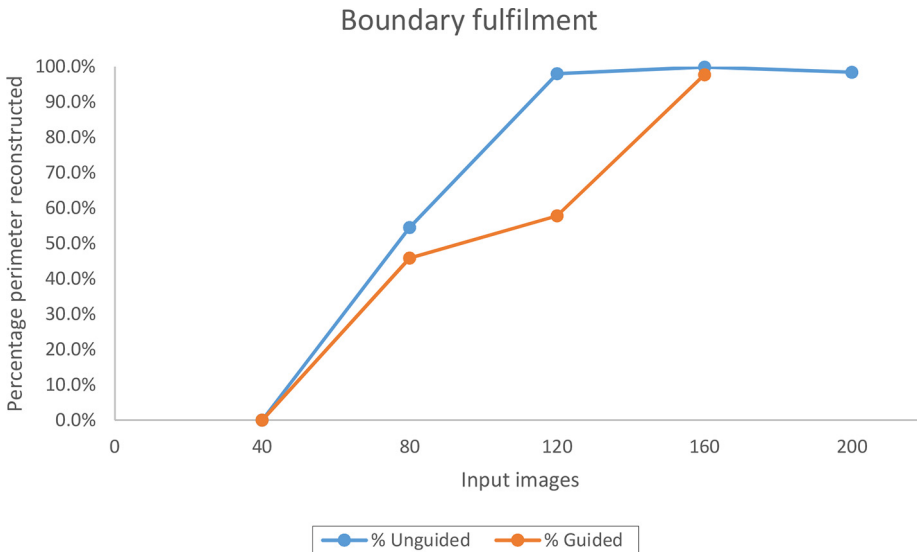


Figure 9.
Percentage perimeter covered by the reconstructed 3D models

perimeter was measured as dimensions of walls on the mesh and documented (supplementary material). The unguided models swiftly achieved a high perimeter cover compared to the gradual stepwise growth of the guided models. At 120 images, the unguided group already had several overlapping images and reconstructed a complete

model (98%). However, images added after proved to be of less use as they only added redundancy and decreased surface smoothness. Interpreting the guided models, the boundary fulfillment grew through 0%, 45.8%, 57.8% and 97.7% for 40, 80, 120 and 160 images, respectively. This highlighted the optimal usage of input images and nearly absent redundancy. In the field, redundancy can be directly related to increased time, cost and inefficiencies to arrive at the same output.

In summary, the validation experiment yielded noteworthy results (Table 4). One significant limitation is the absence of a sensitivity analysis of the reconstruction to the number of input images. An optimum was observed in the unguided group, but it is still unclear what is expected out of the field user in terms of image acquisition. The take of the author on this area is the clear advantage of using visually guided crowdsourcing for indoor scene reconstruction.

3.5 Evaluation

The evaluation seeks to understand how well the artifact supports the needs of the industry practitioners and supports further improvement. Evaluation was solicited in two stages. First, the smartphone application requested subjective responses when users completed the task. Second, industry experts were presented with a demonstration of the prototype or allowed to use the app themselves. Qualitative feedback was solicited and recorded.

3.5.1 Smartphone application responses. The participant survey used in the smartphone application recorded some subjective responses based on the experiment design. A short analysis of these responses is presented below. More than half the participants were interested in having a 3D model of their home or office. In general, participants expressed a willingness to perform this task without a specific incentive other than the final creation of a 3D model of their working space. In contrast, those participants who felt they would require an incentive of some sort for the opportunity cost mentioned incentives, such as money, gamification and food or beverage. The participants also provided feedback on potential use cases. These included interior design for a rental advertisement, Airbnb including mesh models of their homestays, 3D printing and so on.

All 10 participants answered that they were aware of what to capture after reading the instructions. This validated the usefulness of the Image Acquisition Manual. Two minor comments from the participants whose model did not contribute to the reconstructions (either individual or cumulative, P5 in both groups): there was confusion on how to align the mobile camera both vertically and horizontally to get the widest field of view. When asked if the partially reconstructed model assisted in capturing the set of images, the majority answered that the visualization of the model on the application guided them to acquire their pictures. However, one participant specifically stated that the model resolution was not good enough to visualize where to take the pictures.

Table 4. Quantitative results, validation experiment (*4th cumulative reconstruction, guided group)

Acquisition metrics	Unguided	Guided
#images	200	200
Total capture time (min)	22	27.5
Total upload time (min)	79	82
Reconstruction time (min)	75.9	42.4*
Mesh size (100,000 vertices)	9.1	7.2*
Boundary fulfillment (%)	98.4	97.7

3.5.2 Industry feedback. Forge experts expressed that ReCap API has been scarcely used for indoor scene capture, consequently providing promising feedback for the workflow presentation.

One mechanical, electrical and plumbing expert in the field tested the application to reconstruct a meeting room. Impressed with the results, provided confirmation that it is an easier, faster and effective workflow compared to manually modeling models from unorganized point clouds. The FM lead, happy with the results, perceived it as a digital platform where they could link all their static data. Other experts in digital twins foresaw huge potential in such a workflow and showed interest in integrating their solutions with their own reconstructed meshes to create a holistic service offering to the market. Another digital twin portfolio manager asserted the need for a complete pipeline. They proposed a potential use case to integrate the meshes with dynamic IoT data from their systems and achieve a digital communication platform.

3.6 Iteration and further use case development

Following the results and the analysis from the validation study, some of the use cases developed using Forge Viewer and other APIs are presented.

The classroom mesh was uploaded on the Issue and Information tracking application and tested for entering meta-data and tracking issues (Figure 10). Image, link and file data were tagged to the mesh. Information and issues were added as text and monitored at a later point using the developed panel on the user interface (Figure 11).

The 3D markup tool was also tested with the classroom mesh uploading some test images collected by the participants. The images were randomly assigned an identification and a customized icon (Issue, Information and Warning) (Figure 12).

In the multi-model configurator, a corridor mesh was chosen due to the frequent arrangement modification for exhibitions, displays and events. Three object models, a chair,

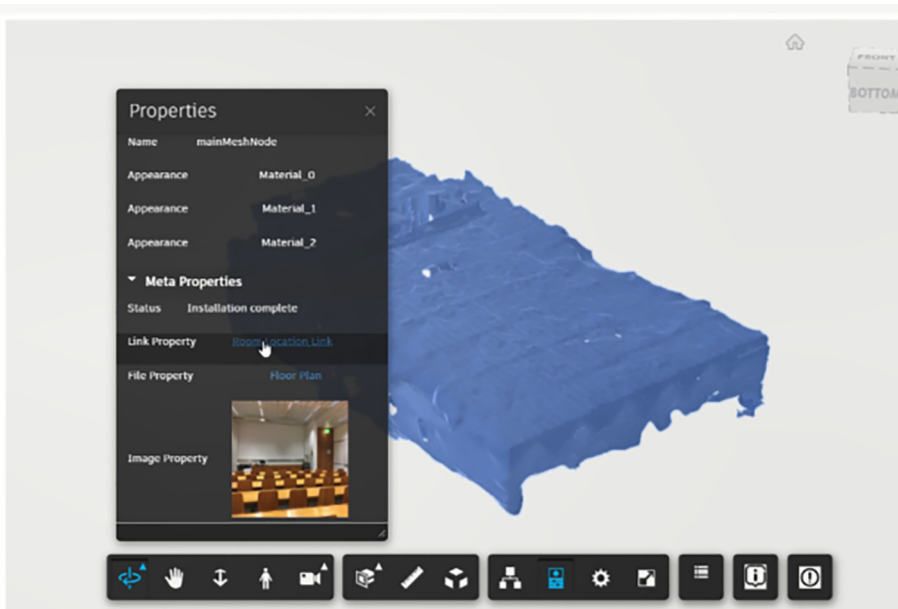


Figure 10.
Issue and information tracking on the Web browser

Figure 11.
3D markup on the
browser

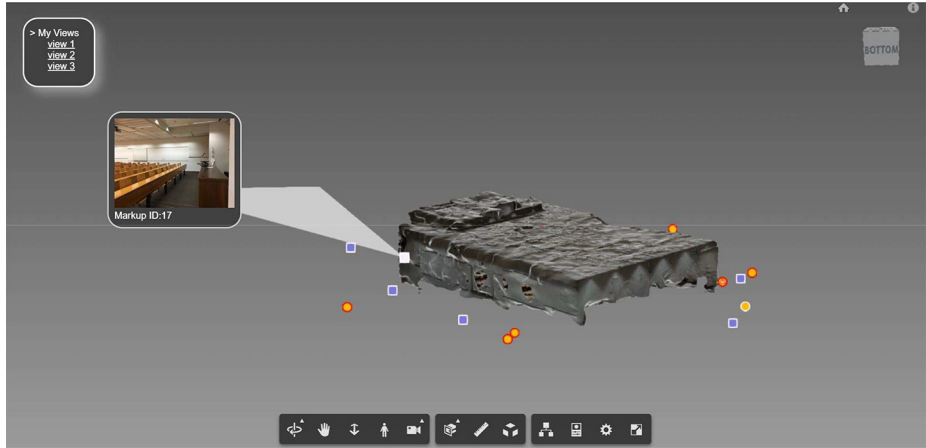


Figure 12.
Multi-model
configurator



a table and a white board were uploaded to be configured within the space. The object models can be scaled, transformed and rotated according to the user requirements. This opened doors for configuring spatial arrangements of event spaces, the interior of houses, storage facilities and any space whose functionality is dynamic.

4. Discussion

The resulting implementation achieves a realistic 3D model characterized by different typologies, minimal trade-off in accuracy, and low processing costs. By crowdsourcing the images, the proposed approach can reduce costs for asset reconstruction by an estimated 93% compared to manual modeling and 80% compared to locally processed reconstruction algorithms. The approach shows how techniques from CS and Machine Vision can be used to fulfill practitioner needs within a fragmented and conventional sector (AEC) that suffers from a lack of coherence and interconnection. Scarce adoption of prevailing tools could be perceived as

a question of acceptance and cultural baggage within the construction sector, simultaneously raising concerns on the knowledge gap between “the provider” and “the user.”

4.1 *Need-based paradigm*

This work successfully interlinked the two different fields by using a bottom-up approach. The goal definition started from questioning the everyday hurdles of the AEC practitioners. The concept focused less on algorithm development rather took advantage of existing tools and optimized it for the demanded use case. It can be conclusively claimed that the users are more satisfied and readier to adopt even with a small enhancement in the process efficiency.

4.2 *Building information model-as-complete vs building information model-as-needed*

The past scholarly discourse has focused on achieving the conventional industry notion of BIM. BIM, at its completion, contains all information and relationships attached to the geometric representation of the asset. It is important to highlight the difference between perception and perfection in such a study. A simple 3D model, easy to navigate, containing only the required information, could pass as *perfect* for a specific user within the divided industry.

Using a simplified approach resulting in *BIM-as-needed* can evidently reduce costs, time and increase efficiency within the industry. In addition, the requirement of the practitioners mostly revolved around the decision-making phase. A geometrically relevant, visually aiding and manipulative model on the browser can serve this need with high confidence.

The above discussion does not imply an anti-BIM notion. The resulting realistic mesh model from the workflow is flexible in terms of usage. It can be modeled manually or automatically to reach an enriched BIM status. Realistically, not all of the existing buildings can be digitalized. The ones that have or want to be can be converted smoother if innovations are designed as a speed breaker and not as a diversion in this journey.

4.3 *Building information model as a basis for digital twins*

Cloud-BIM integration is considered as the second generation of BIM development and is expected to produce another wave of changes across the industry (Li *et al.*, 2014). It does not take into consideration the current understanding and minimal exploitation of BIM within the industry. Furthermore, for existing buildings, BIM is near absent concept and the definition is commonly misunderstood.

This work discusses the potential of BIM as a basis for Digital Twins, a virtual duplicate of the physical environment. The suggested workflow already opens possibilities for such a concept. Buildings are living entities with human involvement. Data is smarter and more dynamic. It is evident that enriching models during data collection is a more feasible workflow rather than collecting data and then attempting to reverse engineer it. A mesh model on the Web, stored and retrieved in the cloud is one step closer to implement technologies like internet of things and Big Data. In addition, FM requires data capturing the current states and process of the building, which is made easier through a Web interface eliminating the need of desktop software.

4.4 *Limitations and future work*

Despite validation of the implemented workflow, some limitations are brought forward in this section. As explained in the literature review, a typical as-is BIM reconstruction pipeline

is divided into two transformations. This work consciously dropped the automated modeling after attempting some existing algorithms. Future scope includes analysis on what possible modeling methods could be used and integrated with this workflow.

The technical capabilities of the smartphone application are bounded by its development environment and API provider (Autodesk Forge). At this point, users can upload only one image at a time (from camera roll/instantaneous shot) using the prototype. A common benchmarking of the benefits of automatic reconstruction is necessary. At what point does automation surpass manual effort? The size and investment of the projects must be the deciding factors in calculated adoption of such a tool.

As all data is hosted on a cloud storage architecture, the questions of governance, ownership and security become key.

This work concludes the adoption of guided crowdsourcing is more suitable for reconstruction of indoor scenes. A sensitivity analysis on the number of pictures is not considered, especially in the unguided crowdsourcing task. This might lead to redundant effort and costs overshadowing the automation.

Reconstructed models are manually oriented, not locally or globally referenced to the building. Existing algorithms allow automatic orientation, often needing another data point such as an inertial sensor or Wi-Fi for indoor localization. Mesh gluing; attaching different individual reconstructed zones to one big model is currently manual. Scholarship in this aspect is scarce and could be a beneficial addition to this implementation.

Some other concepts that deviate from the core of this work but could potentially contribute are:

- Segmentation by crowdsourced annotation of mesh model, i.e. involving humans earlier in the loop for accuracy.
- Reconstructing remote buildings by scraping images from internet websites, such as Facebook, Instagram and Twitter. Firms with a global presence could reduce project costs with such an implementation.

5. Conclusion

Using the DSR, this work provides a firm base for innovative research and potential business models in the operation and maintenance phase of the building lifecycle. This work stands out in developing, testing and validating a mobile-to-mobile novel crowdsourced cloud-based approach for the reconstruction of digital building assets. It started as a question of technical possibility and, in due course, transformed into a “one-click, all done” solution by incorporating the needs of the industry. We use futuristic technologies, such as Autodesk Forge and Serverless Computing, turning the reconstruction process into a parallel and dynamic computing process. The crowdsourcing reduced the huge burden in data collection, secondarily aiding in understanding the importance of human behavior and mentality, which acts as a prime factor in the construction industry. Results of study demonstrated that a realistic 3D model of different typologies can be achieved with minimal trade-off in accuracy and roughly one-fifteenth of the manual costs. The workflow is also five times cheaper than using a representative automation pipeline existing today. Use cases were developed as Web applications serving different demanded functionalities. The retrospective approach is a strong ground to conclude industry-readiness. As for any work, the assumptions and limitations pave the way for future research.

Notes

1. Trimble, GX 3D.
2. Bundler, www.cs.cornell.edu/~snively/bundler/
3. OpenMVG, <https://github.com/openMVG/openMVG>
4. VisualSfM, <http://ccwu.me/vsfm/>
5. COLMAP, <https://colmap.github.io/>
6. <https://forge.autodesk.com/accelerator-program>
7. React Native, <https://facebook.github.io/react-native/>
8. Consecutive: Step 1: P1, Step 2: P1 + P2, . . . , Step 5: P1 + P2 + P3 + P4 + P5.
9. The hourly rate was taken as standard pay for a planner in Zurich. Autodesk Forge cost included the cloud credits for the gigapixels processed.
10. Dell XPS 15, 16 GB RAM, 6 core processor and 4 GB Nvidia graphics memory.
11. ETH Services and Resources, <https://ethz.ch/services/en/service/rooms-and-buildings/building-orientation/gebaeude.html?args0=HIL>

References

- Bello, S.A., Oyedele, L.O., Akinade, O.O., Bilal, M., Delgado, J.M.D., Akanbi, L.A., Ajayi, A.O. and Owolabi, H.A. (2021), “Cloud computing in construction industry: use cases, benefits and challenges”, *Automation in Construction*, Vol. 122, pp. 103441, doi: [10.1016/j.autcon.2020.103441](https://doi.org/10.1016/j.autcon.2020.103441).
- Ben, H., Cruz, C., Boochs, F. and Nicolle, C. (2012), “From unstructured 3D point clouds to structured knowledge – a semantics approach”, in *Semantics – Advances in Theories and Mathematical Models*, doi: [10.5772/37633](https://doi.org/10.5772/37633).
- Bosché, F. (2009), “Automated recognition of 3D CAD model objects and calculation of as-built dimensions for dimensional compliance control in construction”.
- Chu, M., Matthews, J. and Love, P.E.D. (2018), “Integrating mobile building information modelling and augmented reality systems: an experimental study”, *Automation in Construction*, Vol. 85, pp. 305-316, doi: [10.1016/j.autcon.2017.10.032](https://doi.org/10.1016/j.autcon.2017.10.032).
- Chuang, T.-H., Lee, B.-C. and Wu, I.-C. (2017), “Applying cloud computing technology to BIM visualization and manipulation”, in *28th International Symposium on Automation and Robotics in Construction (ISARC 2011)*, doi: [10.22260/isarc2011/0023](https://doi.org/10.22260/isarc2011/0023).
- Dai, F., A.M., Rashidi, A., S.M., Brilakis, I., A.M. and Vela, P. (2013), “Comparison of image-based and time-of-flight-based technologies for three-dimensional reconstruction of infrastructure”, *Journal of Construction Engineering and Management*, Vol. 139 No. 1, pp. 69-79, doi: [10.1061/\(ASCE\)CO.1943-7862.0000565](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000565).
- Dong, J., Xiao, Y., Noreikis, M., Ou, Z. and Ylä-Jääski, A. (2015), “iMoon”, doi: [10.1145/2809695.2809722](https://doi.org/10.1145/2809695.2809722).
- Dore, C. (2014), “Semi-Automatic generation of as-built BIM façade geometry from laser and image data”, *Journal of Information Technology in Construction*, Vol. 19 No. 2, pp. 20-46.
- Estellés-Arolas, E. and González-Ladrón-de-Guevara, F. (2012), “Towards an integrated crowdsourcing definition”, *Journal of Information Science*, Vol. 38 No. 2, pp. 189-200, doi: [10.1177/0165551512437638](https://doi.org/10.1177/0165551512437638).
- Fang, W., Ding, L., Zhong, B., Love, P.E.D. and Luo, H. (2018), “Automated detection of workers and heavy equipment on construction sites: a convolutional neural network approach”, *Advanced Engineering Informatics*, Vol. 37, pp. 139-149, doi: [10.1016/j.aei.2018.05.003](https://doi.org/10.1016/j.aei.2018.05.003).

- Franz, S., Irmeler, R. and Ruppel, U. (2018), "Advanced engineering informatics real-time collaborative reconstruction of digital building models with mobile devices", *Advanced Engineering Informatics*, Vol. 38, pp. 569-580, doi: [10.1016/j.aei.2018.08.012](https://doi.org/10.1016/j.aei.2018.08.012).
- Gao, R., Zhao, M., Ye, T., Ye, F., Wang, Y., Bian, K. and Li, X. (2014), "Jigsaw: indoor floor plan reconstruction via mobile crowdsensing*", doi: [10.1145/2639108.2639134](https://doi.org/10.1145/2639108.2639134).
- Golparvar-Fard, M., Peña-Mora, F. and Savarese, S. (2011), "Integrated sequential as-built and as-planned representation with D4AR tools in support of decision-making tasks in the AEC/FM industry", *Journal of Construction Engineering and Management*, Vol. 137 No. 12, doi: [10.1061/\(asce\)co.1943-7862.0000371](https://doi.org/10.1061/(asce)co.1943-7862.0000371).
- Guo, Y., Bennamoun, M., Sohel, F., Lu, M. and Wan, J. (2014), "3D object recognition in cluttered scenes with local surface features: a survey", doi: [10.1109/TPAMI.2014.2316828](https://doi.org/10.1109/TPAMI.2014.2316828).
- Hartmann, T. and Trappey, A. (2020), "Advanced engineering informatics – philosophical and methodological foundations with examples from civil and construction engineering", *Developments in the Built Environment*, Vol. 4, p. 100020, doi: [10.1016/j.dibe.2020.100020](https://doi.org/10.1016/j.dibe.2020.100020).
- Hichri, N., Stefani, C., De Luca, L. and Veron, P. (2013), "Review of the «as-built bim» approaches".
- Huang, H., Li, D., Zhang, H., Ascher, U. and Cohen-Or, D. (2010), "Consolidation of unorganized point clouds for surface reconstruction".
- Ikehata, S., Yang, H. and Furukawa, Y. (2015), "Structured indoor modeling", *In Proceedings of the IEEE International Conference on Computer Vision*, pp. 1323-1331, available at: www.cv-foundation.org/openaccess/content_iccv_2015/html/Ikehata_Structured_Indoor_Modeling_ICCV_2015_paper.html
- Jung, J., Hong, S., Yoon, S., Kim, J. and Heo, J. (2015), "Automated 3D wireframe modeling of indoor structures from point clouds using constrained least-squares adjustment for as-Built BIM", *Journal of Computing in Civil Engineering*, Vol. 30 No. 4, p. 04015074, doi: [10.1061/\(asce\)cp.1943-5487.0000556](https://doi.org/10.1061/(asce)cp.1943-5487.0000556).
- Kim, B.-K. (2012), "A study on the developing method of the BIM (building information modeling) software based on cloud computing environment".
- Lavikka, R.H., Lehtinen, T. and Hall, D. (2017), "Co-creating digital services with and for facilities management", *Facilities*, Vol. 35 Nos 9/10, pp. 543-556, doi: [10.1108/F-12-2016-0101](https://doi.org/10.1108/F-12-2016-0101).
- Lehtola, V., Kurkela, M. and Hyyppä, H. (2014), "Automated image-based reconstruction of building interiors – a case study", doi: [10.17690/014241.1](https://doi.org/10.17690/014241.1).
- Li, H., Wong, J. and Li, H. (2014), "A review of cloud-based bim technology in the construction sector", Vol. 19, pp. 281-291.
- Li, M., Wonka, P. and Nan, L. (2016), "Manhattan-world urban reconstruction from point clouds", *In European Conference on Computer Vision*, Springer, Cham, pp. 54-69.
- Lu, Q. and Lee, S. (2017), "Image-based technologies for constructing as-is building information models for existing buildings", Vol. 31 No. 4, doi: [10.1061/\(ASCE\)CP.1943-5487.0000652](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000652).
- Macher, H., Landes, T. and Grussenmeyer, P. (2017), "From point clouds to building information models: 3D semi-automatic reconstruction of indoors of existing buildings", *Applied Sciences*, Vol. 7 No. 10, pp. 1030, doi: [10.3390/app7101030](https://doi.org/10.3390/app7101030).
- March, S.T. and Smith, G.F. (1995), "Design and natural science research on information technology", *Decision Support Systems*, Vol. 15 No. 4, pp. 251-266, doi: [10.1016/0167-9236\(94\)00041-2](https://doi.org/10.1016/0167-9236(94)00041-2).
- Ning, X., Li, F., Tian, G. and Wang, Y. (2018), "An efficient outlier removal method for scattered point cloud data", *PLoS One*, Vol. 13 No. 8, doi: [10.1371/journal.pone.0201280](https://doi.org/10.1371/journal.pone.0201280).
- Noreikis, M., Xiao, Y., Hu, J. and Chen, Y. (2018), "SnapTask: towards efficient visual crowdsourcing for indoor mapping", in *Proceedings – International Conference on Distributed Computing Systems*, doi: [10.1109/ICDCS.2018.00063](https://doi.org/10.1109/ICDCS.2018.00063).

- Pătrăucean, V., Armeni, I., Nahangi, M., Yeung, J., Brilakis, I. and Haas, C. (2015), "State of research in automatic as-built modelling", *Advanced Engineering Informatics*, Vol. 29 No. 2, pp. 162-171, doi: [10.1016/j.aei.2015.01.001](https://doi.org/10.1016/j.aei.2015.01.001).
- Peffer, K., Tuunanen, T., Rothenberger, M.A. and Chatterjee, S. (2007), "A design science research methodology for information systems research", *Journal of Management Information Systems*, Vol. 24 No. 3, pp. 45-77, doi: [10.2753/MIS0742-1222240302](https://doi.org/10.2753/MIS0742-1222240302).
- Pu, S. and Vosselman, G. (2006), "automatic extraction of building features from terrestrial laser scanning".
- Rusu, R.B., Marton, Z.C., Blodow, N., Holzbach, A. and Beetz, M. (2009), "Model-based and learned semantic object labeling in 3D point cloud maps of kitchen environments", *In 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009.*, doi: [10.1109/IROS.2009.5354759](https://doi.org/10.1109/IROS.2009.5354759).
- Tang, P., Huber, D., Akinci, B., Lipman, R. and Lytle, A. (2010), "Automation in construction automatic reconstruction of as-built building information models from laser-scanned point clouds: a review of related techniques", *Automation in Construction*, Vol. 19 No. 7, pp. 829-843, doi: [10.1016/j.autcon.2010.06.007](https://doi.org/10.1016/j.autcon.2010.06.007).
- Tommelein, I.D. (2020), "Design science research in construction management: multi-disciplinary collaboration on the SightPlan system", *Construction Management and Economics*, Vol. 38 No. 4, pp. 340-354, doi: [10.1080/01446193.2020.1718723](https://doi.org/10.1080/01446193.2020.1718723).
- van Aken, J., Chandrasekaran, A. and Halman, J. (2016), "Conducting and publishing design science research", *Journal of Operations Management*, Vols 47/48 No. 1, pp. 1-8, doi: [10.1016/j.jom.2016.06.004](https://doi.org/10.1016/j.jom.2016.06.004).
- Volk, R., Stengel, J. and Schultmann, F. (2014), "Building information modeling (BIM) for existing buildings – literature review and future needs", *Automation in Construction*, Vol. 38, pp. 109-127, doi: [10.1016/j.autcon.2013.10.023](https://doi.org/10.1016/j.autcon.2013.10.023).
- Voordijk, H. (2009), "Construction management and economics: the epistemology of a multidisciplinary design science", *Construction Management and Economics*, Vol. 27 No. 8, pp. 713-720, doi: [10.1080/01446190903117777](https://doi.org/10.1080/01446190903117777).
- Wang, C. and Cho, Y.K. (2014), "Automatic as-is 3D building models creation from unorganized point clouds crane safety view project construction automation view project", doi: [10.1061/9780784413517.094](https://doi.org/10.1061/9780784413517.094).
- Wang, J., Xu, K., Liu, L., Cao, J., Liu, S., Yu, Z. and Gu, D. (2013), "Consolidation of low-quality point clouds from outdoor scenes", *Eurographics Symposium on Geometry Processing*, Vol. 32 No. 5, pp. 207-216
- Weyrich, T., Pauly, M., Keiser, R., Heinzle, S., Scandella, S. and Gross, M. (2004), "Post-processing of scanned 3D surface data", *Eurographics Symposium on Point-Based Graphics*.
- Xiong, X., Adan, A., Akinci, B. and Huber, D. (2013), "Automation in construction automatic creation of semantically rich 3D building models from laser scanner data", *Automation in Construction*, Vol. 31, pp. 325-337, doi: [10.1016/j.autcon.2012.10.006](https://doi.org/10.1016/j.autcon.2012.10.006).
- Yalcinkaya, M. and Singh, V. (2014), "Building information modeling (BIM) for facilities management – literature review and future needs, (January 2016)", pp. 1-10, doi: [10.1007/978-3-662-45937-9_1](https://doi.org/10.1007/978-3-662-45937-9_1).
- Yi-Kai, J. and Nai-Pin, H. (2017), "BIM-Based approach to simulate building adaptive performance and life cycle costs for an open building design", *Applied Sciences*, doi: [10.3390/app7080837](https://doi.org/10.3390/app7080837).
- Zhang, K., Hutter, M. and Jin, H. (2009), "A new local distance-based outlier detection approach for scattered real-world data", in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, doi: [10.1007/978-3-642-01307-2_84](https://doi.org/10.1007/978-3-642-01307-2_84).

Appendix

1. Testing and comparative analysis

1.1 Test setup

1.1.1 Initial testing. Two different channels were tested after a justified filtering from the literature review and interviews (Figure A1). The filtering and matching were performed bearing the scope and feasibility in mind. It co-incidentally aligned that the first reconstruction pipeline (Colmap) was an academic research algorithm and the second (Autodesk Forge ReCap API), a cloud-based reconstruction engine API from a commercial software provider in the AEC industry.

The learning tests were carried out in two zones of different typologies. The large corridor (Section 1) is located on the E floor of the HIL building in Eidgenössische Technische Hochschule Zürich Hönggerberg. The second zone is a regular office room. This conscious choice of zones was to test the differences in the reconstruction of a bigger and a smaller zone, both with maximum clutter.

The analysis performed on the resulting model was based on preset factors. It was planned not to analyze geometric accuracy in the initial tests as these were designed to learn and test the effectiveness of the reconstruction tools. The resulting output (point cloud/mesh) was cleaned to remove noise and fed into the automated modeling algorithms to explore the possibilities.

Additional to the analysis, preparation of an Image Acquisition Manual (Section 2) was considered a major goal of the initial testing phase. It was expected that this manual outlined in detail the step-by-step practices to capture images for the highest quality of the information in the reconstructed geometric model.

1.1.2 Actual testing. A field test was planned to incorporate the learnings, address the limitations of the two methods, check the generality and provide strong conclusions for the comparative analysis. The common observation from both the reconstruction methods was the inefficiency in reconstructing windows, reflective and texture-less surfaces. Hence, a closed classroom (with no windows was selected as the location for the field test). In addition, another level of testing (Figure A2) was added to manually add texture to the room walls and compare the result with the as-is reconstruction. The average amount of clutter and the geometric complexity were similar to the previous test zones.

The image acquisition path (Figure A3) was clearly defined from the learning tests. The number of images per user location to achieve the required overlap was also laid out in the manual and followed for this test.

Not textured: The pictures were taken with no changes to the as-is condition of the room.

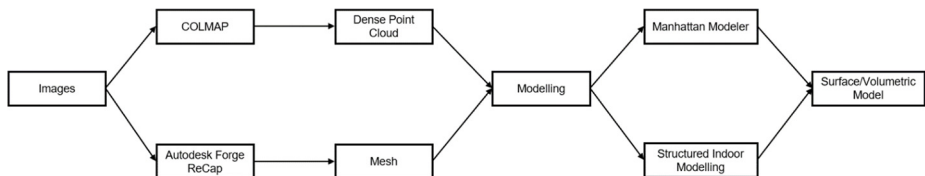
Textured: The completely texture-less white walls of the room were manually textured (increased gradient variation in the pixels of the images captured) using sticky notes of different colors.

Both textured and non-textured tests were performed identically with an equal number of images and exactly the same shoot location.

Some control points were measured within the room to analyze the geometric accuracy of the reconstructed point cloud/mesh. These were points that were most likely to be completely reconstructed (containing texture, gradient and standard dimensions like the door width).

The rest of the metrics were recorded exactly as in the learning tests; the final output was cleaned and input to the geometric modeling algorithms Table A2.

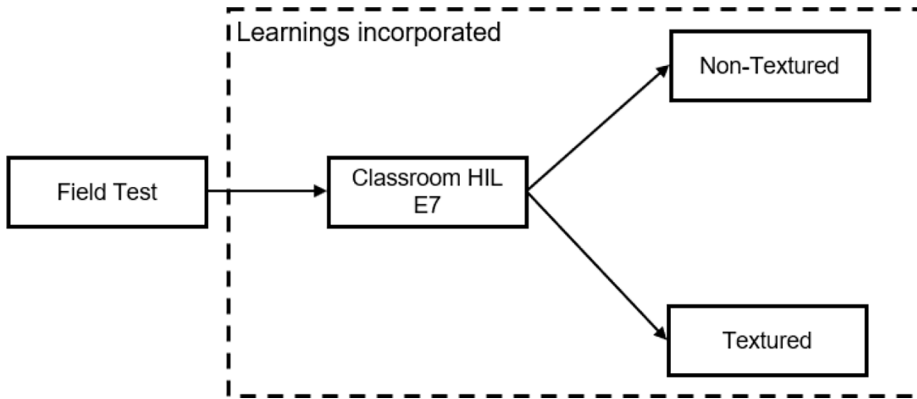
Figure A1.
Testing pipelines
(Manhattan Modeler,
Li et al., 2016),
structured indoor
modeling (*Ikehata
et al., 2015*)



1.2 Findings

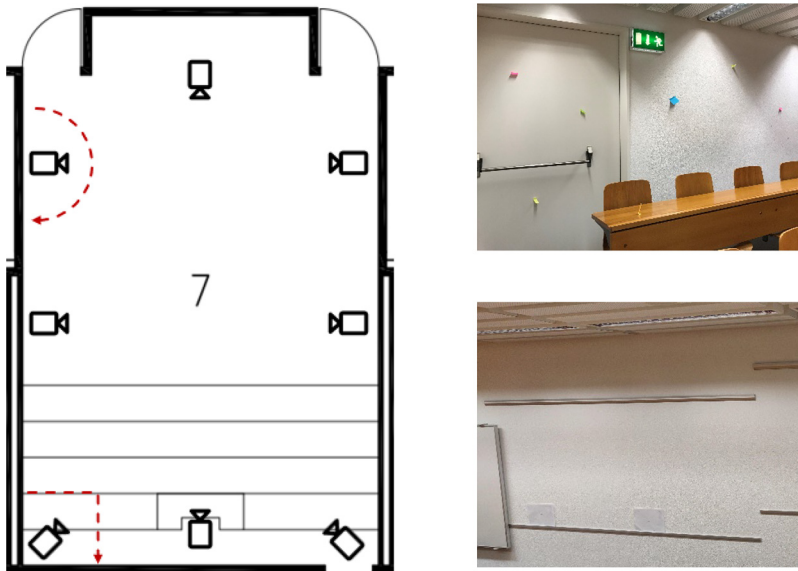
1.2.1 Initial testing. For the corridor, a total of 108 pictures were captured. These images were used for reconstruction through Colmap and Forge ReCap API. Colmap reconstructs through an intermediate sparse reconstruction step to arrive at the denser mesh [Figure A4(a)]. The Forge ReCap reconstruction [Figure A4(b)] was managed through CURL commands.

From crowd to cloud



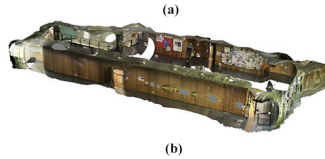
Notes: Nontextured: scene as-is; textured: additional texture with sticky notes

Figure A2. Field test classification



Notes: (Right top) Manual texturing with sticky notes; (right bottom) Control points measured for analyzing geometric accuracy

Figure A3. (left) Acquisition path in HIL E7



428

Figure A4.
(a) Colmap reconstructed corridor;
(b) forge reconstructed corridor

Table A1. Test setup	Camera	12 MP
	Coverage	At least 60% overlap
	Position	Around the perimeter
	Lighting	Completely lit
	Others	No occupants/moving objects

Table A2. Table metrics of comparison used for learning tests	Acquisition time	Total time to capture images
	Computing time	Time required to reconstruct 3D model using the chosen pipeline
	Cleaning time	Time for denoising the output model
	Cost	As a sum of acquisition cost, processing and post-processing cost

The raw output of Colmap was a dense point cloud, while that from the Forge reconstruction engine is a triangulated mesh. The results (Table A3) discussed pertain to cleaned (removing unwanted visual noise) point cloud/mesh. Colmap and Forge reconstruct the output in an arbitrary scale with relative distances between the surfaces. This is transformed to metric (real world) scale by one controlled measurement. Total cost [9] was calculated as the product of active manual time and the average planner cost per hour.

Active time: minutes/hours of manual effort to arrive from real world scene to reconstructed model.

Passive time: minutes/hours of machine effort.

Acquisition metrics	Colmap	Autodesk Forge
#images	108	108
#images matched	89	108
Acquisition time (min)	5	5
Reconstruction time (min)	110	83
Cleaning time (min)	45	3
Total time (min)	160	91
Cost (CHF)	80	16

Table A3.
Learning test results for corridor

The Colmap graphical user interface (GUI) required the user to be on screen for the four distinct steps: Feature Matching, Feature Extraction, Sparse Reconstruction and Dense Reconstruction. In-built meshing algorithms were not suitable for indoor scenes due to the amount of clutter involved. The algorithm was computationally intensive occupying full process memory and graphics memory on a high-end laptop [10]. Consequently, the machine time overlapped with human effort while using Colmap. The situation was entirely different for the Forge engine as the only local activity was to upload the images using the CURL commands. A stable Wi-Fi connection proved more than sufficient. The wait time during the reconstruction could be used as man-hours on some other activity. This conclusively eliminated a lot of active time and cost.

The presence of windows on one side of the corridor influenced Colmap to reconstruct a lot of external noise. Autodesk Forge engine differentiated the internal and external zones and reconstructed the window as an overlaying frame.

Colmap results are output as .Ply and interconvertible to several other point cloud/mesh formats (such as .pts, .xyz, .obj). The Forge reconstruction results can be requested from the cloud cluster in .rcm (high resolution streaming ReCap textured mesh) or .obj (open-source textured mesh), which can be post processed to any other format. The Forge API also provides management for notifications through e-mail once the reconstruction is complete, which was implemented in the testing.

Similarly, for the office, a total of 42 images were captured and processed through Colmap [Figure A5(a)] and Forge [Figure A5(b)].

The detailed results are tabulated (Table A4). The distortion and noise due to the presence of full wall sized window in the office room turned out higher than the corridor case in both algorithms. This can be attributed to the increased variation in the external environment which SfM cannot recognize as features of the captured zone. Colmap matched 34 out of 42 images requiring a total time of 71.5 min. The cloud reconstruction took 39.5 min, with just about 16% active manual time.

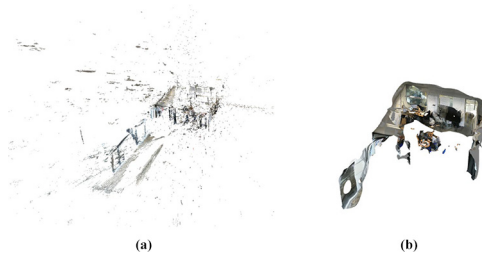


Figure A5
(a) Colmap
reconstructed office;
(b) forge
reconstructed office

Acquisition metrics	Colmap	Autodesk Forge
#images	42	42
#images matched	34	42
Acquisition time (min)	3.5	3.5
Reconstruction time (min)	47	33
Cleaning time (min)	21	3
Total time (min)	71.5	39.5
Cost (CHF)	35.75	7.25

Table A4.
Learning test results
for office

These tests were used as a learning method to understand the intricacies of image acquisition, limitations of the reconstruction algorithms and guidance for the field test. The outcome of this test was twofold. The first being the “Image Acquisition Manual,” which is a step-by-step best practice guide for users to capture data for an optimal reconstruction. The other outcome was the understanding of time (influencing cost) consumption by the different tools, influencing the field test design.

The conversion of point cloud/mesh to a surface/volumetric model was attempted with two chosen academic algorithms. Cleaned and ordered corridor and office point cloud/meshes were input into Manhattan Modeler and Structured Indoor Modeling. Manhattan Modeler produced appreciable result (Figure A6) for the corridor case whereas failed with the office for a high amount of clutter. This is justified by the specific algorithmic design as it was developed for reconstructing exterior urban scenes from point clouds and did not include high tolerance to indoor objects. The source code of Structured Indoor Modeling was modified by the author to accommodate this study. However, the output from the implementation was unsatisfactory as the script worked with free space and point space evidences (Ikehata *et al.*, 2015) of registered and ordered point clouds. The resulting geometry was a reduced cube neglecting the actual boundaries of the generated mesh.

1.2.2 Field testing. Both the non-textured and textured scenarios of the classroom had 136 images captured, all factors remaining the same. The outcomes of non-textured (Figure A7) and textured (Figure A8) differed in how the texture-less surfaces were reconstructed. Metrics similar to the initial tests are recorded (Table A5) with addition of geometric accuracies (Table A6).

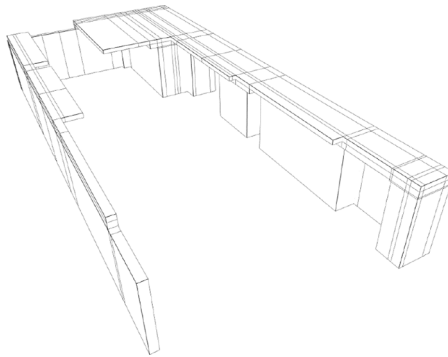


Figure A6.
Automated geometric modeling of corridor using Manhattan Modeler



Figure A7.
Non-textured reconstructions of auditorium (a) Colmap; (b) Forge



Figure A8.
Textured reconstructions of auditorium (a) Colmap; (b) Forge

The non-textured room was only partially reconstructed due to the absence of any variation in color or surface gradient on the other three walls. The sticky notes behaved as differentiating texture rendering a more suitable reconstruction. The differences between Colmap and Forge were similar to the initial tests for both non-textured and textured studies. It was observed that the number of images dropped in the textured case was considerably lower (3.6% as compared to 19% in the non-textured), highlighting the fundamental need of gradient in the images for feature matching. The numbers showed that Forge reconstruction is faster compared to Colmap. However, the difference in speed of reconstruction between the tools decreased as the number of images increased.

The geometric accuracy of the reconstructed models was assessed by manually measuring 11 distances within the test room. After scaling, the meshes were sectioned and the corresponding points were identified and measured. Both the pipelines resulted in highly tolerant reconstructions. The maximum difference in distances was 38 cm for Colmap versus 17 cm for Forge.

A comparative analysis was conducted using the results from the field test. To assess the generality, the same analysis was also performed for the initial tests. The corridor and textured auditorium are considered as full reconstructions for this analysis (Table A7).

Interpretation of the results concludes that transforming mobile captured into images into point cloud/mesh models are faster with the Autodesk Forge API compared to Colmap. Using the Forge infrastructure eliminated the need for computing power, which previous scholarship has classified as the largest hindrance for graphical reconstruction. The processing cost is lower in the cloud pipeline. This can be attributed to the requirement of manual intervention and inability to use the local

Acquisition metrics	Colmap	Autodesk Forge	Colmap	Autodesk Forge
	Non-textured		Textured	
#images	136	136	136	136
#images matched	44	58	131	136
Acquisition time (min)	6	6	24	24
Reconstruction time (min)	66	37	194	181
Cleaning time (min)	12	2	35	3
Total time (min)	84	45	253	208
Cost (CHF)	42	9.5	126.5	25.5

Table A5.
Field test results

Element	Manual measurement (m)	Colmap (m)	Autodesk Forge (m)	Deviation Colmap (%)	Deviation Forge (%)
Wall 1, length	12.39	12.76	12.31	2.99	-0.65
Wall 2, length	8.87	9.14	8.74	3.04	-1.47
Wall 3, length	12.39	12.77	12.32	3.07	-0.56
Door 1, height	2.1	2.03	2.23	-3.33	6.19
Door 1, width	1.25	1.33	1.42	6.40	13.60
Door 2, height	2.1	1.99	2.1	-5.24	0.0
Door 2, width	1.25	1.21	1.24	-3.20	-0.80
Front, height	2.9	2.60	2.85	-10.34	-1.72
Back, height	2.3	2.33	2.28	1.30	-0.87
White board, height	1.18	1.21	1.15	2.54	-2.54
White board, length	1.50	1.51	1.51	0.67	0.67
<i>Mean absolute error</i>	-	-	-	3.83	2.64

Table A6.
Geometric accuracy of reconstructed models

machine for other activities during Colmap runtime. Two major positives of Forge APIs are modularity and scalability. Both the pipelines were geometrically accurate, Forge API performing considerably better than Colmap. However, Colmap performed extremely well in terms of robustness and structural representation of the output. This conclusion might be biased because of different output formats. Colmap results are visualized and point clouds as compared to meshed Forge models. Meshing algorithms use surface smoothing and this might be the cause of uneven surfaces in the Forge model. Of worth mentioning is the accessibility of the source code for Colmap. In theory, it was possible to customize the tool and call the functionality from any other self-developed software. This demanded user learning and the final tool could only be implemented locally. The Forge pipeline overtook in this aspect with several code samples, extensive support in development and Web/mobile implementation possibilities. The level of automation is perceived as the distribution of total processing time between active (manual) and passive (machine) time (Figure A9). It was observed in all cases that a satisfactory expected reconstruction was achieved with lower manual intervention through the Forge pipeline.

The automated modeling algorithms did not provide successful results for the field tests. This was due to the indoor clutter in the auditorium. The use cases attended did not demand a highly accurate volumetric model rather a digital asset that served their purpose. In this work, this approach is defined as service-oriented or need-based paradigm.

Table A7.
Comparative
analysis (*16 GB
RAM, 4 GB graphics
card)

Criteria	Corridor	Auditorium	Corridor	Auditorium
	Colmap		Autodesk Forge	
Level of automation	High, needs manual intervention		Very high, completely automated	
Processing time (min)	160	253	91	208
Computing load	Intensive, entire process and graphics memory consumed*		Light, no local activities, processing split on cloud clusters	
Accuracy	-		-	
Practicality	Existing GUI, source code optimization, comprehensive documentation		Command line interface, freedom of customization, notification management	
Scalability	Not possible, single instance per runtime		Parallel processing, any number of requests	

Distribution of processing time

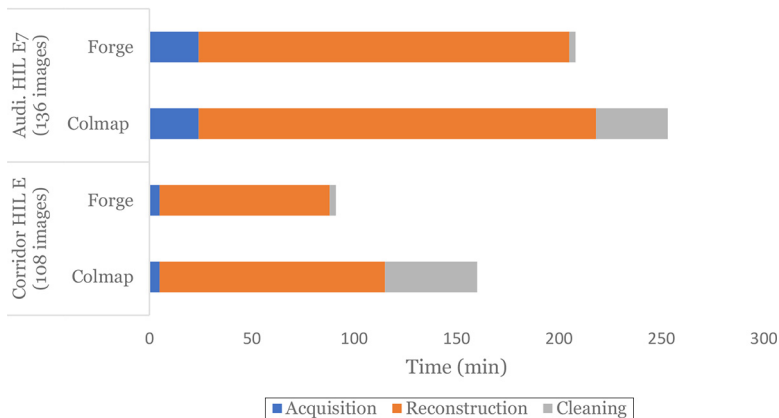


Figure A9.
Active and passive
time for corridor and
auditorium (textured)
with Colmap and
Forge

On completion of the comparative analysis, the cloud-based reconstruction engine using Autodesk Forge ReCap API performed suitably. The second half of transformation, reconstructed mesh → 3D model did not provide expected results with the chosen algorithms. Although previously scholarly discourse mainly focused on achieving a surface/volumetric geometric model from point clouds, a highly structured model grammar was not the core focus of this work. Owing to the service-oriented paradigm of this work, the automated modeling was consciously dropped justified by the gap in the literature review, actual demands of the practitioners in the industry and the possibility to arrive at a solution with the resulting realistic mesh (instead of surface model).

2. Image acquisition manual

This section outlines the learnings of the initial testing and documents the best practices to capture images for the best possible reconstruction. The major points in this section were transferred to the Info_Screen of the mobile application to guide the participants in acquiring images:

- A minimum camera resolution of 5 MP (Reference: iPhone6 has a 8 MP camera).
- Do not change the focal length (Fixed focal length while shooting), do not zoom in or zoom.
- Out on a mobile, follow the autofocus.
- If your phone has high dynamic range, activate it.
- Try to eliminate shaking or blurriness in images.
- Do not take too many images on texture-less/featureless surfaces (surfaces that are plain, same color, no change in gradient).
- Avoid areas that are shiny, transparent or reflective.
- Make sure the room is evenly lit and perfectly still (in terms of objects/people in the room).
- Moving objects/people will result in bad reconstruction due to highlights.
- Capture a photo roughly every 15°–20° (horizontal and vertical).
- Make sure there is at least a 60% overlap.
- Get the widest field of view in every picture.
- Move a few steps after you have some pictures at the first location. Simply turning your camera to get all sides will not work.
- Walk along the perimeter, follow the example markups.

3. Layout of test and experiment rooms (Figure A10)

4. Interfacing with Colmap and Autodesk Forge (Figure A11)

All the required parameters were preset in the Colmap source code and most of the interfacing was through the GUI. Autodesk Forge was only an API provider which meant complete scripting and interfacing development by the user (Figure A12). To begin with, simple command line tools were used and then, later on, a API testing tool, Postman was incorporated with developer background to interface much faster.

5. Measurements – crowdsourcing experiment

5.1 Geometric measurements of the reconstructed meshes (Table 8)

5.2 Mesh characteristics (Table 9)

Figure A10.
Floor plan
HIL E – ETH
Honggerberg [11]

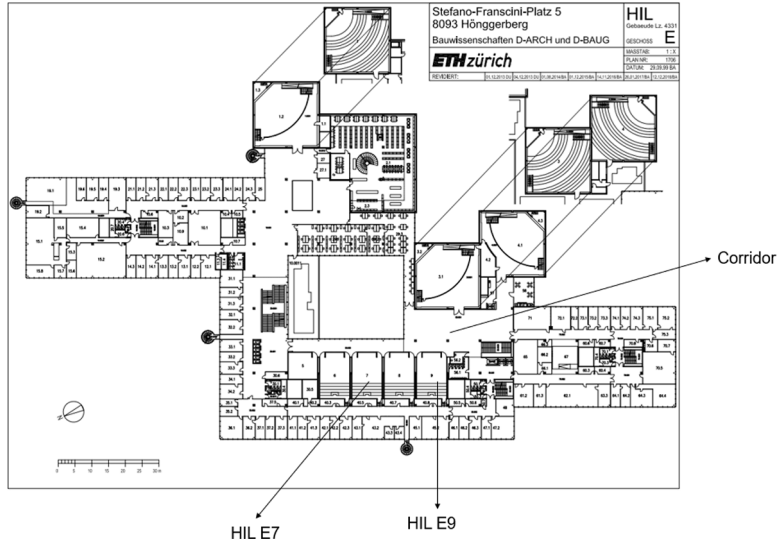
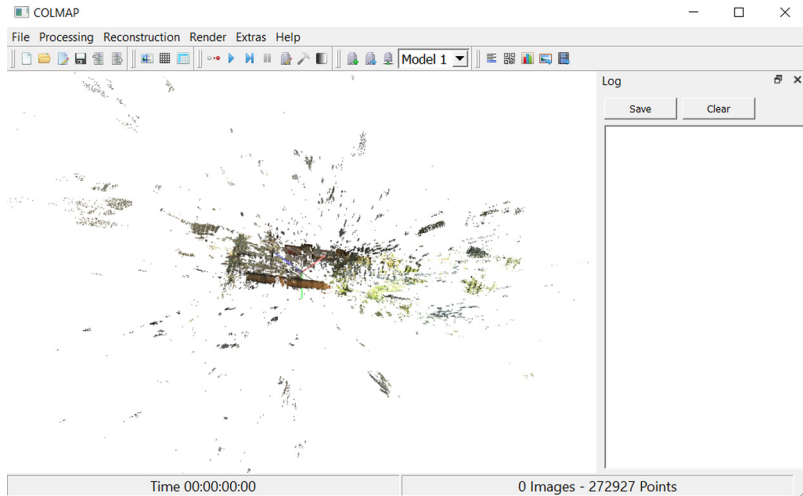


Figure A11.
Graphical user
interface – Colmap



From crowd to cloud

435

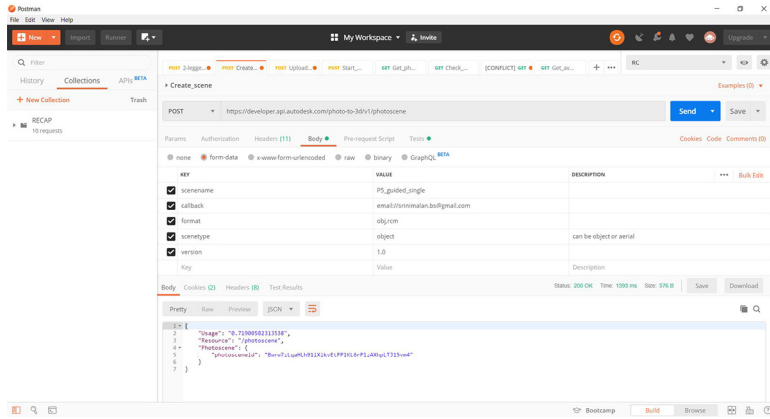


Figure A12.
Postman, an application programming interface testing tool

Element	Manual	Partial Complete Complete Complete					Partial Partial Complete Error				
		P1	P2	P3	P4	P5	P1	P2	P3	P4	P5
Wall 1, length	12.39	—	12.16	12.35	12.43	12.32	—	7.19	9.06	12.35	
Wall 2, length	8.87	—	8.1	8.85	8.79	8.82	—	8.89	5.77	8.74	
Wall 3, length	12.39	—	4.53	12.38	12.73	12.41	—	4.79	1.73	12.42	
Door 1, height	2.1	—	—	1.58	2.02	2.01	—	—	1.47	1.73	
Door 1, width	1.25	—	—	1.31	1.35	1.3	—	—	1.31	1.16	
Door 2, height	2.1	—	—	1.51	1.96	2.02	—	—	1.55	1.66	
Door 2, width	1.25	—	—	1.22	1.3	1.25	—	—	1.13	1.15	
Front, height	2.9	—	—	2.52	2.75	2.79	—	—	2.17	2.34	
Back, height	2.3	—	2.1	1.47	2.24	2.23	—	1.54	0.97	1.4	
White board, height	1.18	—	1.22	1.17	1.24	1.18	—	1.16	1.19	1.18	
White board, length	1.5	—	1.54	1.49	1.54	1.49	—	1.49	1.49	1.48	
Wall 4 (sum of strips on the entrance wall)	11.87	—	—	11,034	11,519	11,226	—	—	9,834	10,972	
Control Point Wall 1	1.15	—	—	1.16	1.17	1.14	—	—	1.13	1.13	
Control Point Wall 3	1.15	—	1.15	1.15	1.16	1.16	—	—	—	1.14	

Table A8.
Geometric accuracies

Mesh properties	Manual	P1	Partial P2	Complete P3	Complete P4	Complete P5	P1	Partial P2	Partial P3	Complete P4	Error P5
Vertices	–	96,495	1,466,906	4,933,987	910,134	1,310,294	63,809	194,904	356,276	702,479	
Faces	–	172,063	2,631,920	8,873,178	1,647,127	2,373,636	117,953	342,644	627,379	1,257,159	
Surface area	–		257	320	319	317	891	95	118	229	
Volume	–		356,084	272	274,247	335,241	8,847,733	351,976	154,788	238,694	
File size (MB) – RCM	–	29.8	203	687	187	245	16.9	32.6	61.6	151	
File size (MB) – OBJ		16.8	283	992		252	11.5	35.8	65.1	131	
Perimeter	45.52	0	24.79	44,614	45,469	44,776	0	20.87	26,394	44,482	

Table A9.

Mesh characteristics

Note: RCM = Reality capture mesh

Corresponding author

Daniel Mark Hall can be contacted at: hall@ibi.baug.ethz.ch