BIM reconstruction from 3D point clouds: A semantic registration approach based on multimodal optimization and architectural design knowledge

Fan Xue, Weisheng Lu, Ke Chen, Chris Webster

This is the peer-reviewed post-print version of the paper:
Xue, F., Lu, W., Chen, K., & Webster, C. J. (2019). BIM reconstruction from 3D point clouds: A semantic registration approach based on multimodal optimization and architectural design knowledge. *Advanced Engineering Informatics*, 42, 100965. Doi: 10.1016/j.aei.2019.100965
The final version of this paper is available at: https://doi.org/10.1016/j.aei.2019.100965.

Abstract

Reconstructing semantically rich building information model (BIM) from 2D images or 3D point clouds represents a research realm that is gaining increasing popularity in architecture, engineering, and construction. Researchers have found that architectural design knowledge, such as symmetry, planarity, parallelism, and orthogonality, can be utilized to improve the effectiveness of such BIM reconstruction. Following this line of enquiry, this paper aims to develop a novel semantic registration approach for complicated scenes with repetitive, irregular-shaped objects. The approach first formulates the architectural repetition as the multimodality in mathematics. Thus, the reconstruction of repetitive objects becomes a multimodal optimization (MMO) problem of registering BIM components which have accurate geometries and rich semantics. Then, the topological information about repetition and symmetry in the reconstructed BIM is recognized and regularized for BIM semantic enrichment. A university lecture hall case, consisting of 1.9 million noisy points of 293 chairs, was selected for an experiment to validate the proposed approach. Experimental results showed that a BIM was satisfactorily created (achieving about 90% precision and recall) automatically in 926.6s; and an even more satisfactory BIM achieved 99.3% precision and 98.0% recall with detected semantic and topological information under the minimal effort of human intervention in 228.4s. The multimodality model of repetitive objects, the repetition detection and regularization for BIM, and satisfactory reconstruction results in the presented approach can contribute to methodologies and practices in multiple disciplines related to BIM and smart city.

Keywords

Building information model, architectural repetition, multimodal optimization, semantic enrichment, 3D point cloud

1 Introduction

The research reported in this paper is positioned in a small but rapidly growing body of literature on reconstruction of building information models (BIM) (Tang et al. 2010; Valero et al. 2016; Belsky et al. 2016; Xue et al. 2018). A BIM is a digital representation of physical and functional characteristics of a facility to enhance data interoperability and information sharing in the building lifecycle (NIBS 2015). The keyword in the term 'BIM' is 'information' (Lu et al. 2019). Schlueter & Thesseling (2009) classified BIM information into three categories including geometric, semantic and topological, whereby geometric information directly relates to the building form in three dimensions; semantic information describes the properties of components (i.e., more advanced rule and function information); and topological information captures the dependencies of components. Semantics will become more important as BIM grows into a mature technology in architecture, engineering, construction, and even smart city development.

Not many existing built facilities have a semantically rich BIM. One approach to make up such fault is to enrich a BIM's semantics manually. The manual method, albeit accurate, is tedious and time-consuming (Chen et al. 2015). The cost of manual semantics enrichment may far exceed the values that the enriched BIM semantics can generate. Therefore, researchers in recent years have endeavored to deploy various semi-automatic or automatic methods to reconstruct BIM from high-quality yet inexpensive measurement data, e.g., satellite images, or 3D point clouds, and enrich semantics in the BIM (Huber et al. 2011; Xiong et al. 2013; Barazzetti 2016; Jung et al. 2016; Thomson & Boehm 2015; Pătrăucean et al. 2015).

Volk et al. (2014) categorizes the BIM reconstruction methods into two subclasses, i.e., "data-driven" and "model-driven," based on their principles; while another well-known taxonomy distinguishes "Scan-vs-BIM" from "Scan-to-BIM" regarding the involvement of as-designed BIM in inputs (Bosché, et al. 2013). However, most approaches in the literature of each subclass relied on a generic computer vision process called 'semantic segmentation', whereby every point in a 3D point cloud (or a pixel in a 2D image) is assigned to a semantic label first. The semantic segmentation has its fair share of shortcomings which can be largely alleviated by an emerging segmentation-free paradigm (Andreopoulos & Tsotsos 2013; Xue et al. 2019b). For example, the 'semantic registration' approach fits semantically rich components into an intermediate BIM by maximizing the similarity (or minimizing the errors) between the reconstructed BIM and the whole measurement data, and subsequently registering them with detected semantic information and topological information (Xue et al. 2019b).

Along with the considerable progress in developing methods for BIM reconstruction and semantics enrichment, some researchers (e.g. Chen et al. 2018; Wang et al. 2018) serendipitously discovered that architectural design knowledge can be utilized to improve the efficiency and effectiveness of these methods. Architectural features such as symmetries,

planarity, parallelism, and orthogonality in relationships between building components, contain rich semantics in their own right. Properly retrieved, they can be rich semantics to be reconstructed into the building information models. Such architectural features are not accidental. Rather, they are the result of functions, economics, mechanics, manufacturing, and aesthetics (Mitra 2008; 2012) and they therefore represent clues in matching geometric pattern with meaning and the related symbolism of language. They can also be applied as constraints to effectively eliminate noise in measurement data and to reduce the search space of formulated problem (Chen et al. 2017).

Such advances in semantic registration and utilizing architectural domain knowledge have not fully overcome difficulties in dealing with complicated scenes with repetitive, complex-shaped objects. For example, multiple identical furniture measured as identical point cloud patches are the multiple optima, i.e., 'modes,' in registering the furniture; but, a unimodal algorithm often wastes computational resources on re-explorations without incorporating efficient search space structures such as the 'neighborhood' topology (Du et al. 2015). Thus, the overall efficiency and effectiveness of the existing unimodal registration are relatively low, which is in line with unimodal algorithm's inferior results on various multimodal benchmark datasets (Chen et al. 2010; Li et al. 2013). The multimodal nature triggered us to apply multimodal optimization (MMO) – a well-discussed problem in applied mathematics – to BIM reconstruction in such circumstances.

In addition, shape and pattern repetition, which conveys important semantics about the design, function, and organization of a facility, has not been widely used for BIM reconstruction. Only a few recent studies have explored the modeling of repetitive structural ribs and piers of a bridge (Hidaka et al. 2018), openings on walls (Dore & Murphy, 2014; Previtali et al. 2018), boundary patterns of rooms (Jung et al. 2018), and indoor furniture (Wang et al. 2018) but they focused on simple-shaped objects or noise-free measurement data. Repetition is actually an ordering principle in architecture leading to sophisticated patterns and structural regularity to support life and well-being in the buildings comprising a city (Ching 2007; Fan et al. 2017). Examples of architectural repetitions exist in furniture setups of conference rooms, patterns or windows on building facades, as well as city blocks. The repetition-based reconstruction for such complex-shaped objects is a research gap that we now address.

We aim to advance the semantic registration approach for BIM reconstruction by experimenting multimodal optimization algorithms and applying architectural knowledge like repetition patterns. The remainder of the paper is organized as follows. The next section is a literature review covering (a) the paradigmatic change from 'semantic segmentation' to 'semantic registration' for BIM reconstruction and semantic enrichment; and (b) the opportunities to enhance such semantic registration approaches. The third section is a detailed description of the MMO-based approach in different languages, e.g., mathematical language

and pseudocode. The fourth section presents an experiment using a university lecture hall case, consisting of 1.9 million noisy points of 293 chairs as a testbed. Section 5 is an in-depth discussion of the implications and limitations of the research and the last section concludes.

2 Literature review

2.1 From 'semantic segmentation' to 'semantic registration'

In the context of reconstructing BIM from 2D images and 3D point clouds, there are two well-known taxonomies of the BIM reconstruction methods. Volk et al. (2014) categorized the shape-based, shape descriptor-based, and material-based matching methods as "data-driven" methods while knowledge and context-based methods are "model-driven." Alternatively, Bosché et al. (2013) regarded the BIM reconstruction with referencing to as-designed BIMs as the "Scan-vs-BIM" subclass and those without referential BIM resources as "Scan-to-BIM." However, most BIM reconstruction applications in the literature, no matter data-driven, model-driven, Scan-to-BIM, or Scan-vs-BIM, relied on semantic segmentation for extracting the object surfaces and creating BIM components (Barazzetti 2016; Babacan et al. 2017).

Semantic segmentation is a computer vision process that assigns each point (or pixel in 2D images) to a semantic label (Shamir 2008). The segmentation methods involved in BIM reconstruction can be broadly grouped into four types: (i) *a priori* rules, (ii) geometric shape descriptors, (iii) supervised machine learning classifiers, and (iv) a combination of these methods for multiple types of BIM components. The *a priori* rules for BIM component recognition utilize the regularities of individual components, such as the region-growth regarding the planarity of walls and ceilings (Huber et al. 2011) and the prism boundary reconstruction of indoor space (Valero et al. 2012). Explicit shape descriptors extract the characteristic geometric features for shape matching, such as local convexity (Son & Kim 2017) and the Laplace-Beltrami filtering (Wang et al. 2018). Supervised machine learning classifiers such as stacking of logistic regression (Xiong et al. 2013), convolutional neural network (Babacan et al. 2017), and random forest (Bassier, et al. 2019) have also been applied. Many studies employed a combination of multiple segmentation methods for multiple types of BIM components, e.g., Nguyen and Choi (2018) removed planar primitives before the RANSAC fitting of the cylindrical piping systems, and Czerniawski et al. (2018)'s point density-based clustering followed by a bagged decision tree for planar objects. However, these semantic segmentation-based methods have three weaknesses in common: (1) unsatisfactory results for complex-shaped objects (e.g., decorations, furniture, and appliances) (Wang et al. 2018; Zou et al. 2018); (2) reliance on *a priori* rules or labeled data set for training the correlational models; and (3) failure to reuse as-designed or online open BIM resources (Bosché et al. 2013; Xue et al. 2018).

Segmentation-free methods have thus been developed recently for overcoming these weaknesses. Xue et al. (2018; 2019b) proposed a semantic registration approach that essentially

reassembles individual BIM components into a complete model in iterations by minimizing the overall error (or maximizing the similarity) between the reconstructed BIM and the whole measurement data. So far, semantic registration was validated, on both 2D images and 3D point clouds of both indoor and outdoor scenes, e.g., about 80% precision and recall in reconstructing 293 theater chairs, using unimodal algorithms such as covariance matrix adaptation evolution strategy (CMA-ES) (Xue et al. 2019b). Hidaka et al. (2018) developed another segmentation-free method, in which similar regions of template CAD models were first adaptively localized then fine-tuned by the iterative closest point (ICP) algorithm. These segmentation-free methods proven to be successful in making use of existing BIM resources (e.g., components collected from open BIM libraries) to enrich the reconstructed BIM with semantics, topology, and fine details.

However, the segmentation-free methods still suffer from a few limitations (Xue et al. 2019b). The first limitation is that existing segmentation-free methods are not effective and efficient enough for complicated scenes with repetitive, irregular-shaped objects. One of the reasons for this is that existing semantic registration approaches rely on 'unimodal' problem solving, in which only one optimal solution can be found in one time. For complicated scenes with repetitive objects, unimodal problem solving suffers the unnecessary re-exploration of the problem search space in the component-by-component processing. The second limitation is the method's proneness to input errors such as noise, clutters, and occlusion, due to the adoption of the objective functions such as the SSIM (structural similarity), RMSE (root-mean-square error), and the descriptor-based similarity in Hidaka et al. (2018). The last, but not the least, limitation is the availability of online open or as-designed BIM resources and annotated topological requirements, so that they may not work on unique and tailor-made components.

2.2 Two opportunities to enhance the 'semantic registration' approaches

Multimodal optimization (MMO) can enhance the segmentation-free methods by addressing the first limitation, i.e., to enhance effectiveness and efficiency for complicated scenes. MMO is a class of non-linear optimization that aims to find *all* the optimal solutions (*i.e.*, 'modes') to a multimodal problem (Das et al. 2011). For instance, the problem "arg $\min_{x\in[0,10\pi]} \cos(x)$" has five solutions, *i.e.*, 1, 3π, 5π, 7π, and 9π. Due to the multimodality, an MMO algorithm can find all the five values, while a unimodal algorithm can only find one (Xue et al. 2019b; Kim et al. 2013). MMO algorithms have been applied to many complicated problems with multiple local optima, such as protein structure prediction (Wong et al. 2010) and engineering design (Forrester & Keane 2008). Recent MMO competitions showed that NMMSO (niching migratory multi-swarm optimizer), RS-CMSA-ES (covariance matrix self-adaption evolution strategy with repelling subpopulations), and NEA2+ (niching the CMA-ES via nearest-better clustering) are among the best algorithms (Li et al. 2013; Fieldsend 2014; Ahrari et al. 2017; Qu et al. 2012). For example, NMMSO dynamically manages a large set of Particle Swarm Optimization (PSO) processes for a balanced search for all solutions. MMO algorithms'

exclusive capability to handle multimodal problems thus provides an opportunity to complement the existing research of unimodal algorithms for segmentation-free methods for BIM.

The use of architectural domain knowledge can help address the problem of noisy data and occlusion, i.e., the second limitation of the segmentation-free methods. Several previous studies have made use of architectural domain knowledge in the building reconstruction process. For example, Fisher (2003) described the application of standard feature relations to enhance the building reconstruction. De Luca et al. (2006) proposed a reconstruction approach that used architectural design knowledge to interpret architectural shapes from 3D point clouds. Likewise, Liu and Wu (2016) presented a rule-based method to reconstruct historical building with different architectural styles. In addition, Chen et al. (2018) applied a fundamental regularization rule to rooftop elements from noisy LiDAR point clouds and reconstructed over one thousand buildings located in Hong Kong Island. However, most of these studies were limited by the use only of parallel or orthogonal relationships between building components, paying less attention to the repetitions that embed meaningful architectural domain knowledge. Our anticipation in starting this study was that the two opportunities, i.e., MMO and repetition as architectural domain knowledge, will enhance the 'semantic registration' approaches to BIM reconstruction.

3 Methodology

To reiterate, this study focuses on developing a novel semantic registration approach to BIM reconstruction by exploring MMO algorithms and making good use of architectural repetition. The approach proposed, as shown in Figure 1, consists of three steps: mathematical formulation, reconstruction of an intermediate BIM based on multimodality of repetition, and regularization of the intermediate BIM using repetition formations. The first two steps are fully automated for an intermediate BIM output based on candidate components with accurate geometry and rich properties, while the last one is semi-automated for the detection and regularization of topological relationships about repetition and symmetry for a final BIM output. There are two differences between the proposed approach and previous semantic registration applications in Xue et al. (2018; 2019b). First, the problem formulation changes from constrained optimization to multimodal optimization, while the solving algorithms which convey automatic reconstruction evolve from unimodal algorithms to MMO ones. Secondly, the architectural design knowledge, such as repetition and symmetry, is applied to recognition and regularization of repetition formations for BIM reconstruction, so that the approach can benefit from both data-driven and model-driven principles.

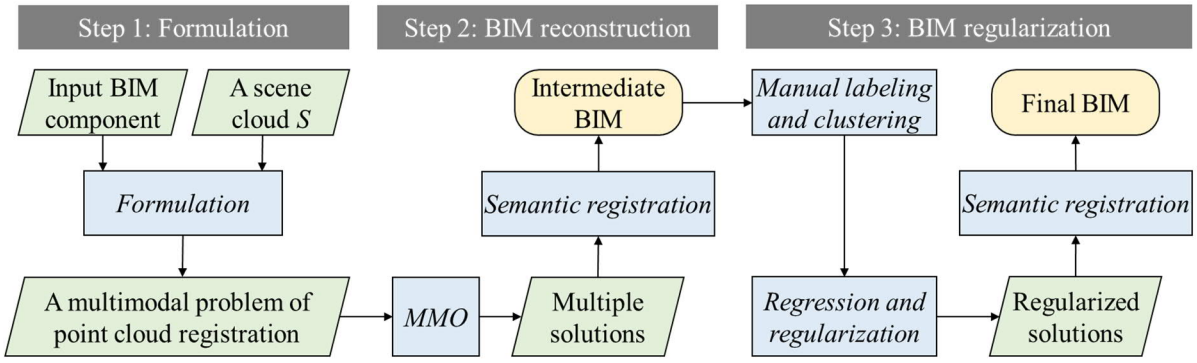


Figure 1. A general framework of the proposed MMO-based semantic registration approach

3.1 Mathematical formulation

The semantic registration approach requires two inputs, i.e., measurement data (e.g., a 3D point cloud or 2D images) and a set of BIM components annotated with topological relationships (Xue, et al. 2019b). To highlight the repetition patterns of components, the BIM reconstruction task in this paper has two data inputs. One input is a 'scene cloud' $S = \{ p_1, p_2, \ldots, p_n \} \subset \mathbb{R}^3$ of n points of repetitive objects. The other is a candidate BIM component of the repetitive objects, such as a parametric Revit family, from which a 'component cloud' $\mathcal{C} = \{ p_1, p_2, \ldots, p_m \} \subset \mathbb{R}^3$ of m visible surface points is evenly sampled. The task of BIM reconstruction is thus equivalent to an optimization problem that finds all the instances of $\mathcal{C}$ in S:

$$
\begin{aligned}
&\textbf{arg min} \quad f(x) = RMSE(\mathcal{C}(x), S) = \left[\frac{1}{m} \sum_{p \in \mathcal{C}(x)} \| p - N(p, S) \|^2\right]^{1/2} \\
&\textbf{s.t.} \qquad C(x) \le 0, \\
&\qquad\quad \mathcal{C}(x) = \{T_x(p) \mid p \in \mathcal{C}\}, \\
&\qquad\quad T_x(p) = \mathbf{R}p + [t_x, t_y, t_z]^{\mathrm{T}}, \\
&\qquad\quad \mathbf{R} = \begin{bmatrix} \cos r_z & -\sin r_z & 0 \\ \sin r_z & \cos r_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos r_y & 0 & \sin r_y \\ 0 & 1 & 0 \\ -\sin r_y & 0 & \cos r_y \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos r_x & -\sin r_x \\ 0 & \sin r_x & \cos r_x \end{bmatrix}, \\
&\qquad\quad x = [t_x, t_y, t_z, r_x, r_y, r_z]^{\mathrm{T}} \in \mathbb{R}^6
\end{aligned}
\tag{1}
$$

where x indicates the six degrees of freedom (DoFs) about 3D translation (i.e., t_x, t_y, t_z) and rotation (i.e., r_x, r_y, r_z), *RMSE* is the *root-mean-square error* function to minimize, $N(p, S)$ returns the nearest point of p in S, C represents the topological constraints such as "a window must reside on a wall" and "a desk sits on horizontal surfaces" (Belsky et al. 2016; Xue et al. 2018). The expression "$C(x) \le 0$" is a general form of constraint equations and inequalities, e.g., a constraint $c_1(x) \ge a$ is equivalent to $c_1'(x) = a - c_1(x) \le 0$, and $c_2(x) = 0$ is equivalent to $c_2(x) \le 0$ and $c_2'(x) = -c_2(x) \le 0$. The point cloud $\mathcal{C}(x)$ is a permutated instance of $\mathcal{C}$ transformed by a Euclidean transformation T_x defined on x. $\mathbf{R}$ is a 3×3 orthogonal matrix of rotation, *i.e.*, $\mathbf{R}$ $\mathbf{R}^{\mathrm{T}} = \mathbf{I}_3$, $[r_x, r_y, r_z]^{\mathrm{T}}$ in x is the (proper) Euler angular vector about the axes, and $t = [t_x, t_y, t_z]^{\mathrm{T}}$ is the translation vector of the origin.

It should be noted that Eq. (1) involves one component for clarity of presentation; it, however, does not degrade the generality of the formulation. The reason lies in the incremental build

phase of semantic registration: Given an i-th component and a set $\{ \mathcal{C}_1(x_1), \mathcal{C}_2(x_2), \ldots, \mathcal{C}_{i-1}(x_{i-1}) \}$ of $i-1$ reconstructed components, we can note $\mathcal{C}(x) = \mathcal{C}_i(x) \cup \mathcal{C}_1(x_1) \cup \mathcal{C}_2(x_2) \cup \ldots \cup \mathcal{C}_{i-1}(x_{i-1})$ in Eq. (1) to represent the whole model. In the current study, we only uses *RMSE* in Eq. (1) for simplicity of the objective function. More as-demanded metrics, such as the *non-correspondence rate* (NCR) (Van Kaick, et al. 2011), values of point colors (grayscale or true color), and laser reflectance, can extend the f in Eq. (1) in practice.

An ideal mode (optimal transformation parameters) x^* yields $f(x^*) = f_{\min}$ (e.g., 0), that is, minimal geometric error (or fully corresponded in case of Eq. (2)) between the permutated instance cloud $\mathcal{C}(x)$ and the measured scene cloud S. If there are multiple instances of a parametric BIM component, there should exist multiple modes to Eq. (1). Therefore, it is clear that the formulated problem is a multimodal problem. However, the point cloud of a real building or area inevitably has instrumental, environmental, and calibration errors. In addition, the points of different instances inevitably have heterogeneous point density, geometric accuracy, occlusion, and clutters. Therefore, the ideal condition of "*arg min*" in Eq. (1) is often relaxed to a satisfactory condition $f(x^*) \leq \varepsilon$, where ε is a small error tolerance. The set X^* of multiple satisfactory (Note: rather than ideal or optimal) solutions thus are:

$$X^* = \{ x^* \mid f(x^*) \leq \varepsilon \}, \tag{3}$$

where x^* indicate *one* satisfactory solution (mode) to Eq. (1).

3.2 BIM reconstruction based on multimodality of repetition

The detailed processes of the proposed MMO-based BIM reconstruction is shown in Figure 2. In general, this step includes the two phases of semantic registration, i.e., the incremental build phase and the fine-tuning phase. The difference between the proposed approach and Xue et al. (2019b) is the employment of NMMSO, one of the best MMO algorithms, instead of the unimodal CMA-ES algorithm due to intrinsic multimodality in the formulated problem. Due to the evolutionary searching strategies of NMMSO, the unnecessary re-exploration of search space, a drawback of unimodal semantic registration, was largely eliminated. The proposed approach was implemented in an in-house developed software plugin COBIMG-Revit (Constrained Optimization-based Building Information Model Generator-Revit; source code available at: https://github.com/ffxue/cobimg).

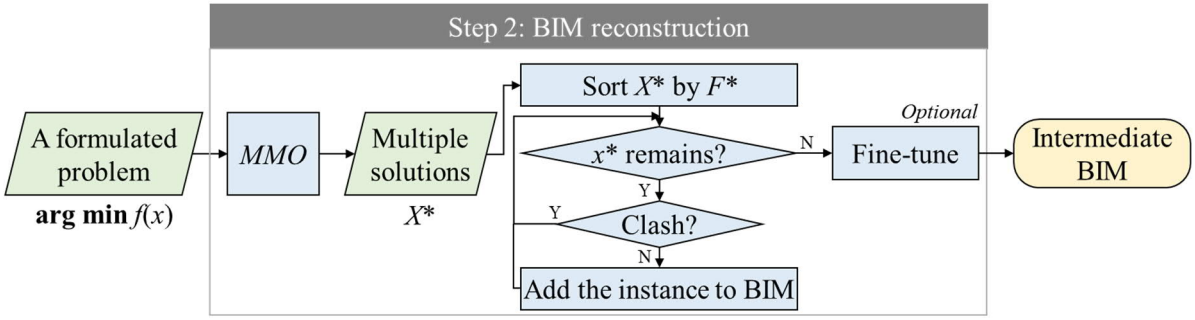


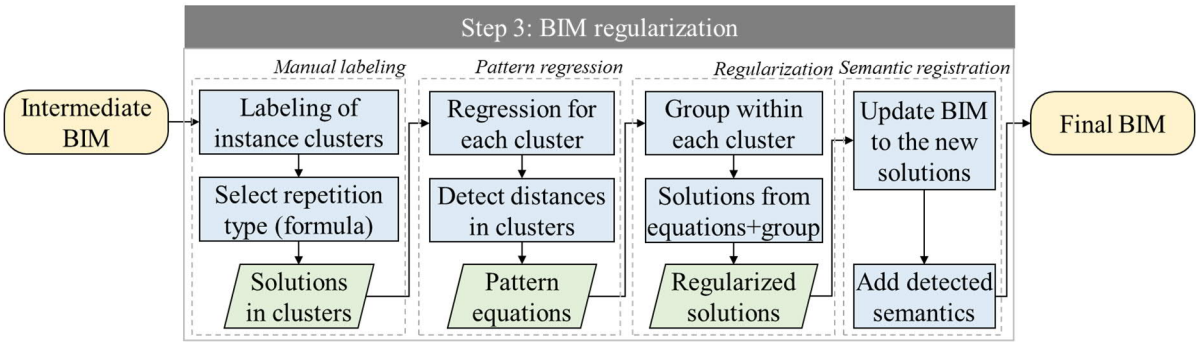
Figure 2. Zoom-in of the proposed MMO-based BIM reconstruction for repetitive objects

As the output of the MMO algorithm, the set X^* in Figure 2 is the multiple satisfactory solutions about the input BIM component, as defined in Eq. (3). However, not every satisfactory solution leads to a feasible instance, this being due to the ε-relaxation which extends the set X^* to include solutions near the optimal solutions. For example, the problem "arg $\min_{x\in[0,10\pi]} \cos(x)$" has five exact solutions while the relaxed problem "$\cos(x) < -1 + \varepsilon,\ \varepsilon > 0, x \in [0, 10\pi]$" has many. MMO algorithms have native strategies, called distance "tolerance", before forking into new modes, which handles this problem in part. However, the distance tolerance is in 6D in our study, instead of the Euclidean 3D. As a result, there still exist a number of clashes between BIM components if the full set X^* is used for creating new instances.

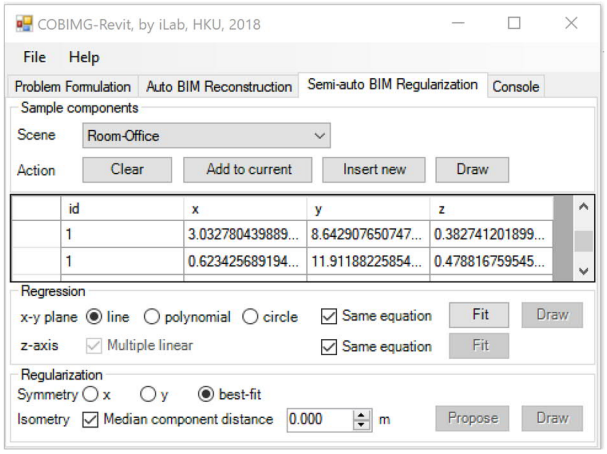
A greedy process, as shown in Figure 2, is employed to cleanse the MMO solutions X^* by accepting non-clashed solutions in the reconstruction. First, the objective values F^* of the solutions in X^* are evaluated, so that the set X^* can be sorted in an ascending order of F^*. Then, the solutions in the sorted X^* are tested one by one in order. For each solution, if there is no clash detected between its instance and the BIM comprising of BIM component instances of previous solutions, the new instance is added to the BIM; otherwise, the solution is skipped. After every solution in X^* is tested, the reconstructed BIM is fine-tuned and output as the intermediate BIM. Besides, the object-level semantics, such as materials, production, and usage, is registered to the BIM. It should be noted that the systematic fine-tuning can be omitted in case the MMO algorithm has performed an equivalent processing during its problem solving. After the Step 2, an intermediate BIM is automatically reconstructed with repetitive BIM components.

3.3 BIM regularization using repetition formations

The intermediate BIM is regularized to create the final BIM. The regularization step aims to correct the errors that come from input data noises or the context-free MMO-based component registration. As shown in Figure 3.a, four modules, i.e., Manual labeling, pattern regression, regularization, and semantic registration, are designed to achieve the aim. In the four modules, only the first one requires human intervention. The regularization was also implemented in our COBIMG-Revit, as shown in Figure 3.b.



(a) Zoom-in of the BIM regularization



(b) Graphical interface of the COBIMG-Revit plugin

Figure 3. The BIM regularization process and implementation in an in-house developed software plugin

In the first module of manual labeling, as shown in Figure 3.a, a human modeler is needed to select a set of components. For example, a row of windows or desks can be quickly labeled as a cluster by dragging a selection box using a mouse and clicking the “Insert new” button in Figure 3.b. In addition, the type of repetition is also chosen by observation of the intermediate BIM. The interim output of the first module is the manually clustered solutions.

The second module detects the patterns as equations. The pattern equation of each cluster, such as lines or a circle, as well as the uniformed equation of the cross-cluster formation, such as parallel lines and concentric circles, are then regressed for the labeled clusters. COBIMG-Revit realizes the multiple linear regression and the least squares methods for other regression models using two Python scientific libraries *scipy* (version 1.1.0) and *sklearn* (version 0.19.1). The median of nearest distance between components is detected within the clusters.

The third module proposes new, regularized solutions. Based on the median distance, each cluster can be segmented into smaller groups. The centroids and range of each group are computed from the locations of its members. If an approximate symmetry is detected or defined, two symmetric groups will have perfectly symmetric centroids and ranges. Based on the

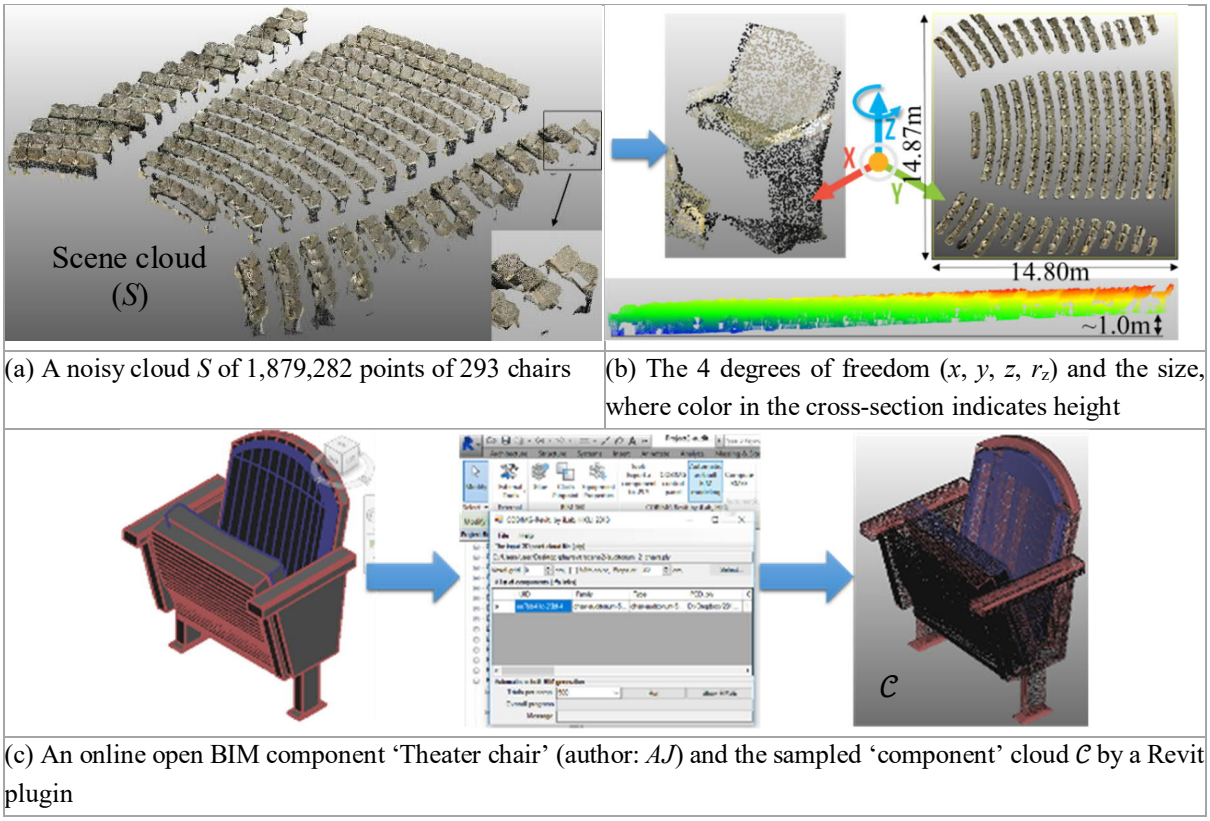
median distance and the centroid of each group, all the solutions in a cluster can be calibrated isometrically to have a uniform Euclidean distance to their neighbors. In the isometric calibration, new regularized solutions will be proposed along the curve of the pattern equation, while the uniform distance is equal to the median distance. The interim output of this module is the regularized solutions.

The final module revise the intermediate BIM. First, the reconstructed components in BIM are updated using the regularized solutions. Then newly detected topological relationships, including the symmetry, cluster, group, sequence in group, and nearest neighbors, enriches the components to form the final semantically rich BIM. Due to the limited involvement of human intervention, the BIM regularization, as shown in Figure 3, can be regarded as a semi-automatic process.

4 Experimental tests

4.1 Experimental settings

A case of a university lecture hall, which is the "Area_2 Auditorium_2" instance in the Stanford 2D-3D-S dataset (Armeni et al. 2017), was selected for validation. The Stanford 2D-3D-S is an open benchmark dataset including a cloud of 695 million annotated indoor points produced from multi-view photos inside a university building (available at http://buildingparser.stanford.edu/dataset.html). One reason for targeting an indoor dataset was that it is more challenging in general. The case is also the largest indoor instance in the dataset. The standard exemplar also allows the results of our experiment to be compared with that of a unimodal algorithm CMA-ES in Xue et al. (2019b). To focus on the repetition itself, the semantic labels in the dataset were used to filter 1,879,282 points of 293 theater chairs as the 'scene' cloud S, as shown in Figure 4.a, by removing other building elements such as walls and doors which were already annotated in the dataset (Note: For unlabeled indoor scenes, the planarity and normals can segment such elements (Thomson & Boehm 2015)). Some parts of S were noisy, incomplete, and cluttered as shown in Figure 4.a. We assumed that a chair has a possible rotation (heading direction) around the z-axis (Figure 4.b). Thus, there were four degrees of freedoms (DoFs), *i.e.*, the 3D centroid (t_x, t_y, t_z) and the heading direction r_z, for each chair. Ground truth values of the positions of the 293 chairs were extracted from the noisy point cloud and manually validated within an error threshold at 10cm; the true values of heading directions were manually measured with an error threshold at 5°. An online open BIM component 'Theater chair,' freely shared by *AJ* at *3DWarehouse.com* (See Figure 4.c), was downloaded for the semantic registration. The input component was selected because it was proven better than some others in Xue et al. (2019b). A 'component' cloud $\mathcal{C}$ of 1,802,939 dense points, as shown in Figure 4.c, was then downsampled from the polygon surface of the volumetric component using an Autodesk Revit (version Educational 64-bit) software plugin developed in Xue et al. (2019b).



(a) A noisy cloud S of 1,879,282 points of 293 chairs

(b) The 4 degrees of freedom (x, y, z, r_z) and the size, where color in the cross-section indicates height

(c) An online open BIM component 'Theater chair' (author: *AJ*) and the sampled 'component' cloud $\mathcal{C}$ by a Revit plugin

Figure 4. A university lecture hall case and a BIM component using *Autodesk Revit 2015*

The computational experiments were conducted on a workstation (dual Intel Xeon E5-2690 v4 2.6GHz, 64 GB memory, Windows 10 Enterprise, 56-threading in all tests), with point cloud library (version 1.8.1) and fast library for approximate nearest neighbor (FLANN, version 1.8.4) for efficient point cloud processing. The mathematical formation step followed Eq. (3) with additional settings from Xue et al. (2019b). In the BIM reconstruction step, a C++ version of the NMMSO algorithm was applied. To make full use of the multi-threading CPUs, the problem-solving was realized by 110 parallel NMMSO threads, with a maximum of 10,000 iterations. Thus, the maximum number of iterations for BIM reconstruction was equivalently 1.1 million, which is comparable to the 1.4 million (5,000 × 289) iterations of CMA-ES in Xue et al. (2019b). The swarm size was 300 and the floating-point error tolerance was set to default (1^{-6}). The results of the NMMSO algorithm automatically registered the template BIM chair to various positions through the COBIMG-Revit for *Autodesk Revit 2015*. In the BIM regularization step, the estimated formations of repetition were concentric circles, and the reflection symmetry and isometry between BIM chairs were also assumed.

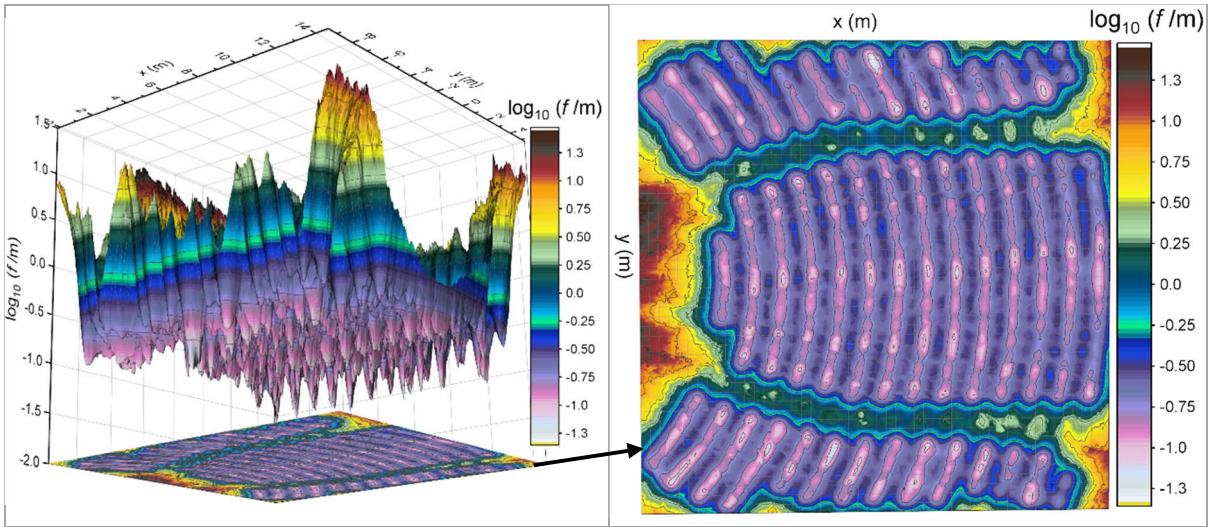
4.2 Experimental results

4.2.1 Problem formulation

Based on Eq. (3), the problem of BIM reconstruction for the university lecture hall was thus:

$$
\begin{aligned}
\textbf{arg min} \quad & f(x) = RMSE(\mathcal{C}(x), S) \\
\textbf{s.t.} \quad & x = [t_x, t_y, t_z, r_z]^{\mathrm{T}} \\
& parent(\mathcal{C}(x)) = Ground \\
& [t_x, t_y, t_z]^{\mathrm{T}} \in boundingbox(S) \\
& r_z \in [0, 360) \\
& f(x) \leq \varepsilon = 0.25 diag_{\mathcal{C}} \approx 0.01 diag_S
\end{aligned}
\tag{4}
$$

Where *parent* is a function that returns the "parent" component that $\mathcal{C}(x)$ attaches to, *boundingbox* indicates the 3D bounding box of the scene cloud S (see Figure 4.b), $diag_{\mathcal{C}}$ stands for the diagonal length of $\mathcal{C}$, and $diag_S$ is the diagonal length of the scene cloud S. The tolerance ε is a constant of minimum requirement for a new BIM component, about 1/4 $diag_{\mathcal{C}}$ and 0.01 $diag_S$, and it can be changed (e.g., to 0.1m or $0.05 diag_{\mathcal{C}}$) for other scenes. Figure 5 visualizes the jagged fitness landscape of Eq. (4) over the x-axis and y-axis. The t_z and r_z in the parameters x were set – after an exhaustive search independent to the experiment – to indicate the best possible f in Figure 5, due to considerably less variance in the z-axis and heading direction than those in the x-y plane. The axes x and y of the spikes in the surface chart, as shown in Figure 5.a, are the $[t_x, t_y]^{\mathrm{T}}$ in the demanded modes (optimal transformation parameters); while the best possible f (unit in meter) is shown in logarithm to the base 10 in color to emphasize the modes. It is obvious that there exist many modes for Eq. (4). Figure 5.b shows the contour map, i.e., the vertical projection, of Figure 5a. It can be observed from Figure 5b that the modes were highly correlated to the position of the chairs. A very regular formation of the chairs can also be seen, which could be the concentric circles resulting from the architectural acoustic input to the theatre's design (Mehta et al. 1999).



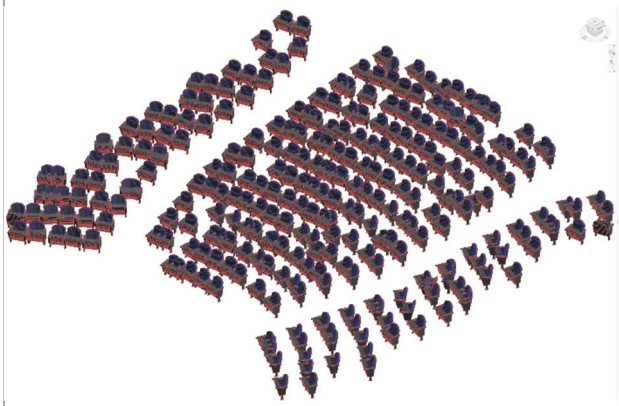
(a) Surface chart of the f (in $\log_{10}$) in Eq. (4), where the demanded modes are the spikes at the bottom

(b) 2D Contour chart, i.e., the bottom projection in (a)

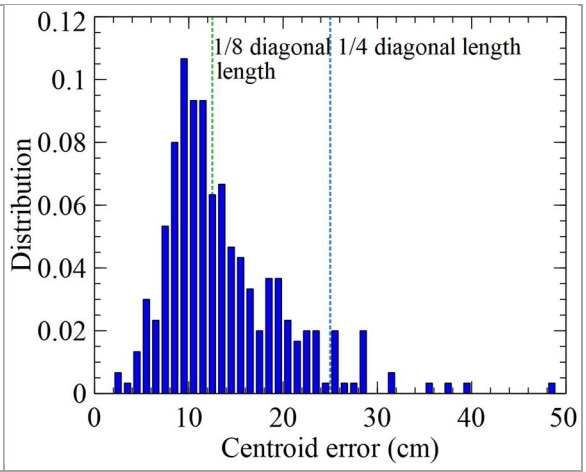
Figure 5. Visualization of the multimodal fitness landscape of the test case over the x-y plane, where z and r_z were assigned as the best values

4.2.2 BIM reconstruction

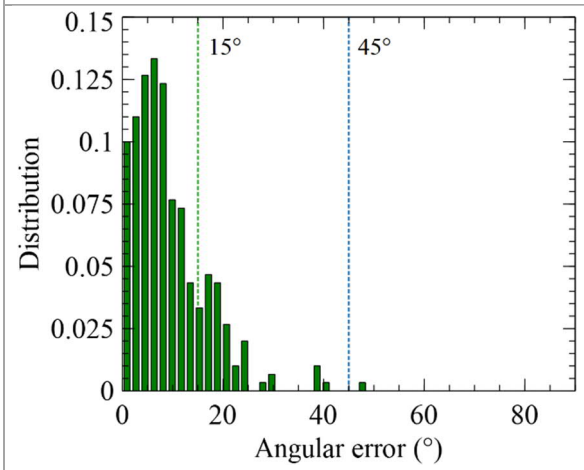
The NMMSO algorithm spent 414.5s to find 300 solutions to Eq. (4) in the incremental build phase using the greedy processing shown in Figure 2. In the second phase of semantic registration, COBIMG-Revit spent 512.1s on fine-tuning the 300 chairs in BIM using CMA-ES. The reconstructed intermediate BIM is shown in Figure 6.a. The overall geometry of the reconstructed BIM looks similar to the input scene, but has some notable gaps with missing objects. Figure 6.b shows the distribution of centroid errors of the 300 chairs, where the mean error = 13.7cm, standard deviation = 6.5cm, and about a half of chairs were placed within the range of 1/8 diagonal length (i.e., the green dashed line of 12.5cm). Figure 6.c visualizes the distribution of errors of their heading directions, where mean error = 10.4°, standard deviation = 13.6°, and a considerable portion of chairs were no less than 15° from the referential directions.



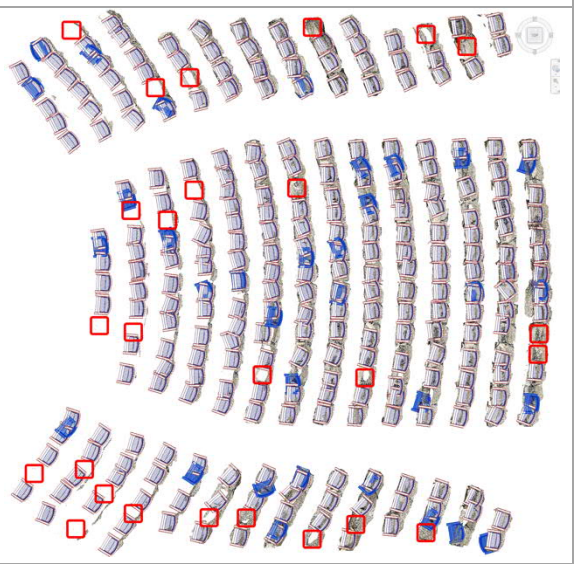
(a) Automatically created intermediate BIM, consisting of 300 instances (modes) of 'Theater chairs'



(b) Distribution of position errors of the instances, mean = 13.7cm, standard deviation = 6.5cm



(c) Distribution of angular errors of the instances, mean = 10.4°, standard deviation = 13.6°



(d) Top view of 267 true positive chairs (wired frames), 33 false positive (highlighted in blue), and 26 false negative (red boxes), where the gray points indicate the cloud *S*

Figure 6. Intermediate BIM automatically reconstructed in 926.6s by NMMSO

In order to compare with previous results such as Xue et al. (2019b), the two blue lines were adopted as the acceptance thresholds, i.e., position error ≤ 25cm and angular error ≤ 45°. In the 300 chairs in the reconstructed intermediate BIM, 267 chairs were true positive, 33 were false positive (i.e., wrongly reconstructed). Figure 6.d shows a top view of the chairs in Autodesk Revit, where the false positive are highlighted in blue and 26 false negative (i.e., missing) chairs are in red boxes. Thus,

$$\begin{aligned} precision &= \frac{\text{true positive}}{\text{true positive} + \text{false positive}} = 267/300 = 89.0\%, \\ recall &= \frac{\text{true positive}}{\text{true positive} + \text{false negative}} = 267/293 = 91.1\%, \\ F_1 &= 2 \times precision \times recall \,/\, (precision + recall) = 90.1\%. \end{aligned} \tag{5}$$

It should be noted that the three metrics may decrease if the acceptance thresholds are changed.

4.2.3 BIM regularization

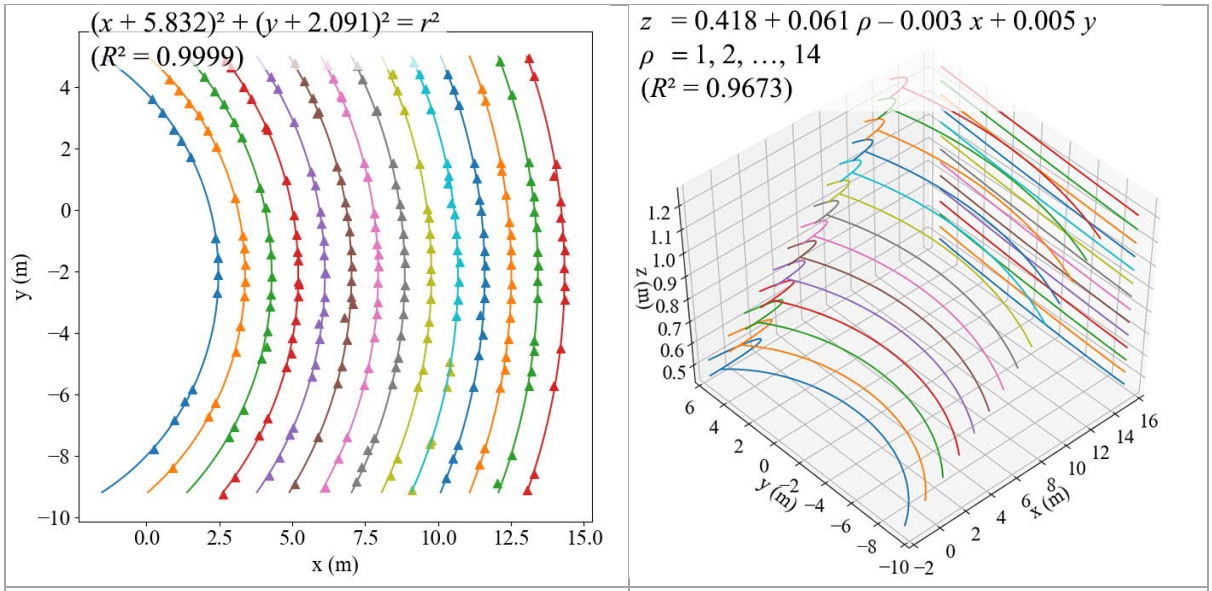
Manual labeling grouped the chairs in correct repetition formations as 14 clusters (rows) in 223.7 seconds. The centroids of the chairs, as shown in Figure 7.a, were used for the regression. The median distance of two neighboring chairs was 55.26cm. The regression of the equations of the concentric circles (acoustic design patterns) on the *x*-*y* plane was conducted using the least square method in 0.06s. The equations obeyed by all chairs were:

$$\begin{aligned} &(x + 5.832)^2 + (y + 2.091)^2 = r^2 && \textit{Concentric circles on the x-y plane} \\ &r = 0.913\,\rho + 7.387 && \textit{with a linear increment on radius} \\ &\rho \in \{1, 2, \ldots, 14\} && \textit{The row number from } 1 \textit{ to } 14 \\ &(R^2 = 0.9999) && \textit{A highly satisfactory regression} \end{aligned} \tag{6}$$

where the center of the 14 circles, i.e., the stage center, was at (–5.832, –2.091), and the radius r increased linearly against the row number ρ. The equation of z values of the chairs' centroids were further obtained by multiple linear regression against x, y, and ρ using the least squares method in 0.06s:

$$\begin{aligned} &z = 0.418 + 0.061\,\rho - 0.003\,x + 0.005\,y && \textit{A multiple linear regression of z} \\ &\rho \in \{1, 2, \ldots, 14\} && \textit{The row number from } 1 \textit{ to } 14 \\ &(R^2 = 0.9673) && \textit{A satisfactory regression} \end{aligned} \tag{7}$$

It can be found in Eq. (7) that z had the highest correlation with the row number ρ. In addition, the coefficients +0.005 of y and -0.003 of x in Eq. (7) suggested that the repetition formations of chairs had a 1:200 slope over the y-axis and a 1:330 slope over the x-axis, while the R^2 = 0.9673 confirmed the confidences of the two slopes. As a result, the projections of the 14 curves on the x-z plane were in "⊃" shapes. The reason could be that the datum in the Stanford dataset was not perfectly calibrated, or that there were heavy asymmetric errors in the data measurement, or that the lecture hall itself has such an as-designed feature (possibly for drainage) or as-built error.

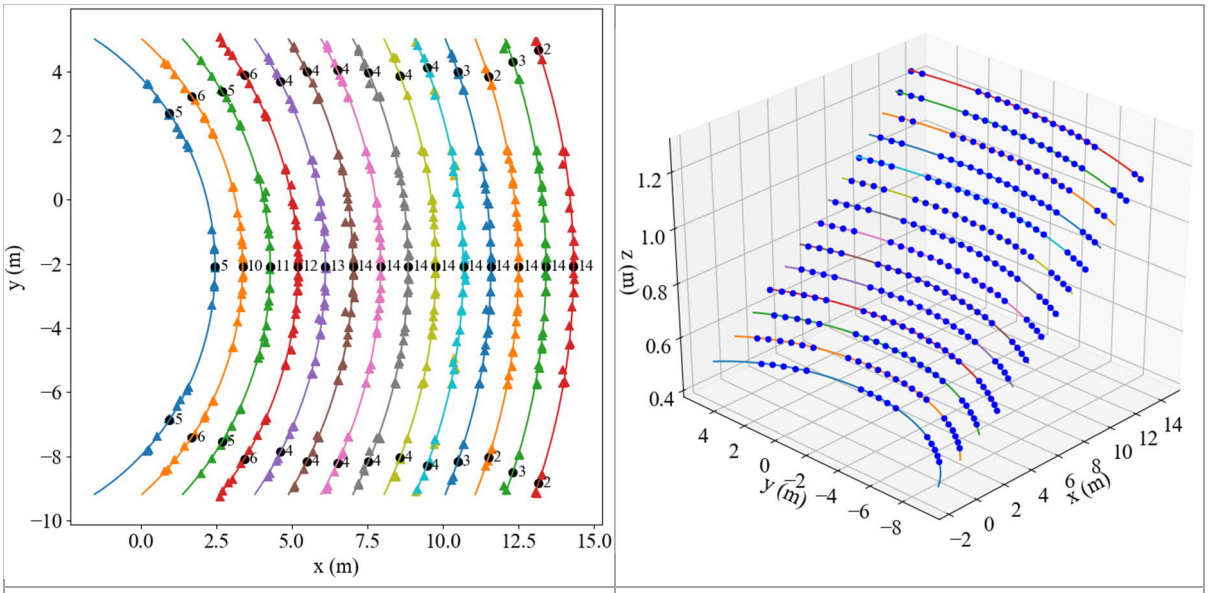


(a) Top view of the 14 rows of repetitive chairs and the 14 concentric circles, where color indicates clusters

(b) 3D view and projections of the multiple linear regression of z (Eq. 7); note that the "⊃" shapes on the left projection indicate slopes along the y-axis

Figure 7. Visualization of regression equations of the repetitive BIM components

Both the symmetric regularization and isometric regularization were set for the BIM chairs. The symmetric regularization first mirrored all the labeled chairs using the axis $y = -2.091$ (see Eq. 6) on the x-y plane, and merge the transformed positions to the original ones, as shown in Figure 8.a. So that most of the missing chairs, i.e., the wrong gaps in the red boxes in Figure 6.b, were filled as shown. In the isometric regularization, the chairs in each extended cluster were first sorted by a clockwise order of their angles to the center (–5.832, –2.091) and grouped using a maximum gap (i.e., the minimum aisle width) set at twice the median chair distance (2 × 55.26cm). Then, the 14 clusters of chairs were split into 42 groups. The center and the estimated number of isometric chairs in each group are shown in Figure 8.a, where the sum of the estimated number was 289. Then, 289 new chair positions with isometric distances were generated on the x-y plane for the 42 groups. The z values were computed by Eq. (7) and the heading directions were set to the center of the concentric circles, as shown in Figure 8.b.



(a) Top view of the centers (black dots) and estimated number of isometric chairs in the 42 symmetric groups

(b) 3D centroids of 289 regularized chair positions using the detected repetition equations (Eqs. 6, 7)

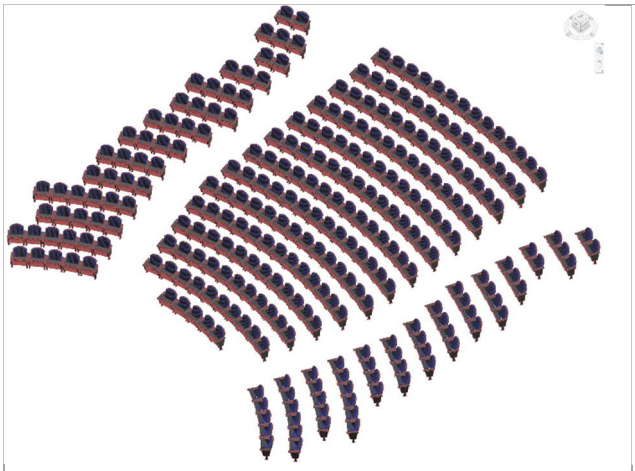
Figure 8. Results of the symmetric regularization and isometric regularization based on repetition in architecture

The 289 regularized chair positions were used to generate a list of semantics of the regularized BIM components, as shown in Table 1. Beside the variables involved in Eq. 4, the cluster (row number), group (theater section), sequence number within group (seat number from right to left), and neighbors of a BIM component. Examples of neighboring components included the left, the right, and the one symmetric to the whole plan.

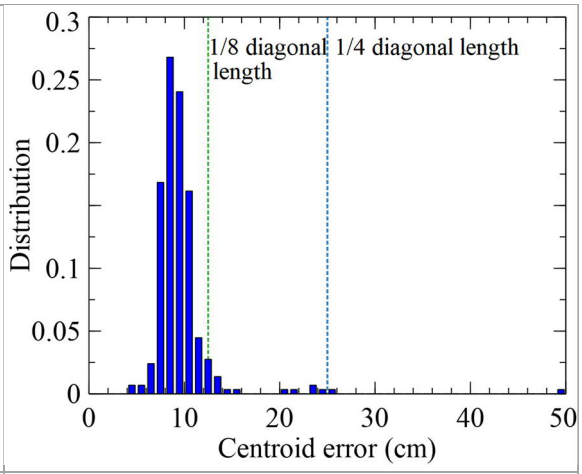
Table 1. List of semantics and topological relations of the 289 regularized components

Id	Type	Variables in Eq. (4)				Cluster	Group	Seq. in	Parent	Neighbors' Ids		
		t_x	t_y	t_z	r_z	(Row)		group		Left	Right	Symmetric
1	1	0.255	3.549	0.497	2.318	1	A	1	Ground	2		15
2	1	0.617	3.131	0.493	2.251	1	A	2	Ground	3	1	14
3	1	0.950	2.690	0.490	2.185	1	A	3	Ground	4	2	13
4	1	1.254	2.229	0.487	2.118	1	A	4	Ground	5	3	12
5	1	1.525	1.747	0.484	2.052	1	A	5	Ground		4	11
6	1	2.393	-0.989	0.467	1.704	1	B	1	Ground	7		10
7	1	2.448	-1.539	0.464	1.637	1	B	2	Ground	8	6	9
8	1	2.467	-2.091	0.461	1.571	1	B	3	Ground	9	7	8
9	1	2.448	-2.644	0.458	1.504	1	B	4	Ground	10	8	7
10	1	2.393	-3.193	0.456	1.438	1	B	5	Ground		9	6
11	1	1.525	-5.930	0.444	1.090	1	C	1	Ground	12		5
12	1	1.254	-6.411	0.442	1.023	1	C	2	Ground	13	11	4
13	1	0.950	-6.873	0.440	0.957	1	C	3	Ground	14	12	3
14	1	0.617	-7.314	0.439	0.890	1	C	4	Ground	15	13	2
15	1	0.255	-7.732	0.438	0.824	1	C	5	Ground		14	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
287	1	14.015	-5.664	1.199	1.393	14	B	14	Ground		286	274
288	1	13.261	-8.582	1.185	1.243	14	C	1	Ground	289		273
289	1	13.076	-9.103	1.183	1.216	14	C	2	Ground		288	272

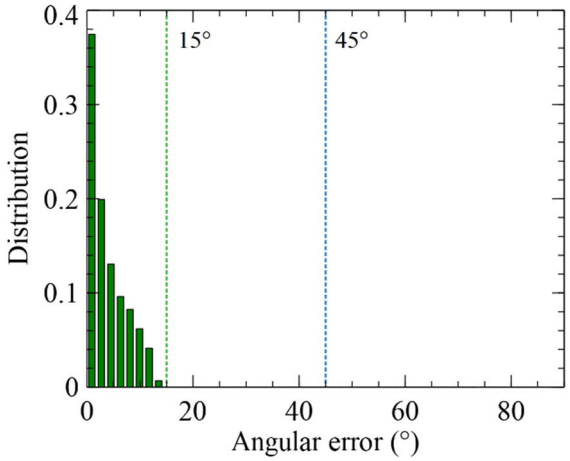
The semantics in Table 1 were then registered to form the final BIM as a 1.82 MB Autodesk Revit project (.rvt) as shown in Figure 9.a, using the COBIMG-Revit plugin in 4.6s. The overall processing time for the final BIM, including automatic reconstruction and semi-automatic regularization, was 1,155.0s (i.e., about 3 seconds per chair). The BIM regularization showed an encouraging improvement as shown in the distributions of position and angular errors in Figures 9.b and 9.c. As shown in Figure 9.d, the number of true positive chairs increased from 267 to 287, the false positive reduced from 33 to 2, and the false negative reduced from 24 to 4. It is worth noting that none of the 4 missing chairs was detected in the intermediate BIM before regularization (see Figure 6). As a result, the final BIM had a precision of 99.3%, a recall of 98.0%, and an F_1 score of 98.6%. The RMSE between the visible surface of the BIM and the input scene cloud was 8.79cm. In addition, the average distance error and the average angular error of the final BIM were 9.6cm and 3.8° regarding the ground-truth values, respectively. Because the majorities in both distributions fall below the two green dashed lines, the precision and recall rates will remain almost the same if the acceptance thresholds are tightened to the green dashed lines.



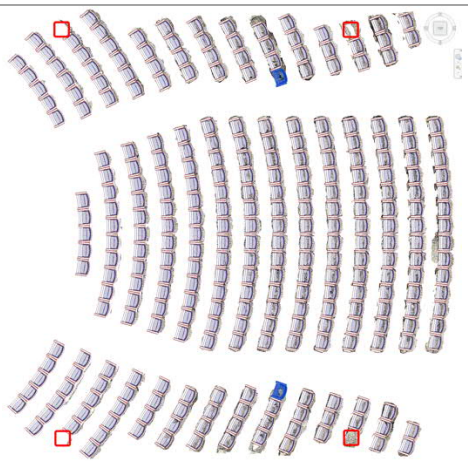
(a) Screenshot of 3D view of the BIM consisting of 289 regularized chairs



(b) Distribution of position errors of the instances, mean = 9.6cm, standard deviation = 3.5cm



(c) Distribution of angular errors of the instances, mean = 3.8°, standard deviation = 3.3°



(d) Top view of 287 true positive chairs, 2 false positives (highlighted in blue), and 4 false negatives (boxes), where the gray points are the input cloud *S*

Figure 9. The final BIM after regularization by the repetition formations

4.3 Semantics stored in the BIM

Rich semantics were also registered to the chairs in the BIM, as a characteristic of the semantic registration approach. Figure 10 shows the screenshot of the properties of the No. 1 chair listed in Table 1. The location, heading direction, and parent component, i.e., the invisible "Ground" was registered as usual. Furthermore, the semantics such as the cluster number, group number, sequence, and the "neighbor" components were also stored as properties from Table 1.

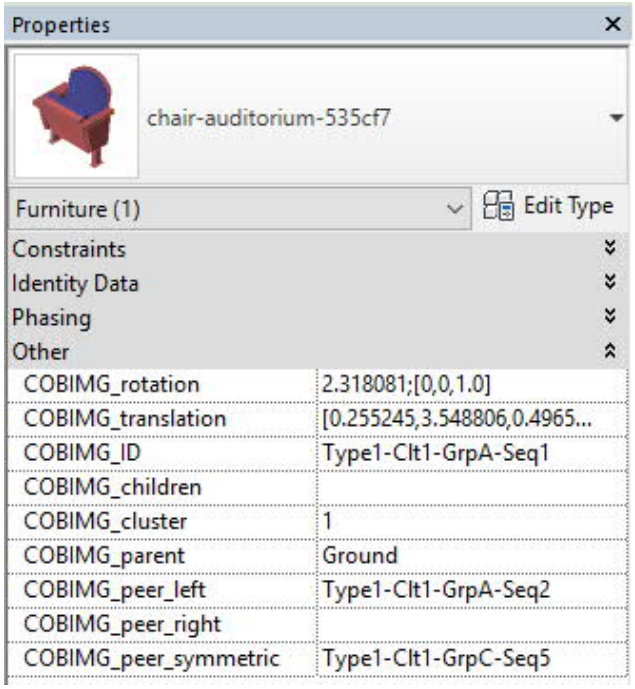


Figure 10. Screenshot of rich semantics of a chair in the final BIM

4.4 Comparison to other algorithms

The results of the proposed MMO-based approach were compared with two well-known 'unimodal' algorithms ICP (Kim et al. 2013) and CMA-ES (Xue et al. 2019b) applied to the same test case. Table 2 lists the comparison results, including the RMSE between the input point cloud and the output BIM, computational time, 3D (t_x, t_y, t_z) position error of chair centroids, angular error of r_z, precision, recall, and F_1 score, of the BIMs by the four methods on the pilot case.

Table 2. Comparison of different algorithms

Evaluation (unit)	*ICP* (Kim, et al. 2013)	*CMA-ES* (Xue, et al. 2019b)	This study	
			Intermediate BIM (automatic)	Final BIM (semi-automatic)
RMSE (cm)	**7.28**	8.10	10.38	8.97
Computational time (s)	3,702.4	1,434.2	**926.6**	1,155.0
Number of chair instances	322	288	300	**289**
Distance error (cm, mean ± stdev.)	17.9 ± 10.8	15.2 ± 7.7	13.7 ± 6.5	**9.6 ± 3.5**
Angular error (°, mean ± stdev.)	28.4 ± 42.9	17.6 ± 28.8	10.4 ± 13.6	**3.8 ± 3.3**
Precision (%)	69.3	81.9	89.0	**99.3**
Recall (%)	76.1	80.5	91.1	**98.0**
F_1 (%)	72.5	81.2	90.1	**98.6**

(Remark: Bold fonts in each row indicate the best value)

The first three data columns are the results of three fully automatic methods, i.e., ICP, CMA-ES, and NMMSO (i.e., the intermediate BIM) in this paper. In comparison to unimodal ICP and CMA-ES, the NMMSO outperformed in automatic BIM reconstruction in all aspects except for the RMSE metric defined on points. The precision by NMMSO was about 10% higher than CMA-ES and about 17% higher than ICP, with 35% time saved versus CMA-ES and 75% saved versus ICP. Thus, the pilot study preliminarily confirms the competence of the MMO-based semantic registration approach for automatic BIM reconstruction.

Furthermore, the effect of using architectural design knowledge can be summarized from the comparison of the last two columns. The Step 3, semi-automatic BIM regularization, resulted in superior results than NMMSO's intermediate BIM in all aspects except for the time cost. Specifically, the precision (99.3%) and recall (98.0%) became very satisfactory at a computational cost of 228.4s to apply architectural design knowledge. The overall processing time of the proposed approach was 1,155.0s and was less than the two unimodal methods reported in literature.

5 Discussion

The experiment reported in this paper confirms the power of a multimodal optimization (MMO) approach to reconstructing semantically rich Building Information Models (BIM). In the methodological sphere, the mathematical concept of multimodality and the practical design rules complement each other to become an efficient (i.e., using less time) and effective (i.e., more accurate) approach. In the practical sphere, the proposed approach can be easily embedded in mainstream BIM platforms to enable various value-added BIM applications such as architectural design, construction management, heritage conservation, and urban digital twin.

This study goes beyond existing unimodal 'semantic registration' algorithms (e.g., ICP and CMA-ES) and introduces a generic MMO approach so that multiple semantics in measurement

552 data (e.g., 3D point clouds) can be found and registered in a more efficient and effective way.
553 The precision and recall of automatic BIM reconstruction increased from about 80% using
554 unimodal CMA-ES to about 90% using the NMMSO (niching migratory multi-swarm
555 optimizer), and the time cost reduced by about 30%. The experiment further confirms the
556 power of architectural design knowledge (e.g., repetition rules) in improving the efficiency and
557 effectiveness of BIM reconstruction. As-designed or as-built building components need to
558 follow certain rules and the rules can reduce the search space in a formulated BIM
559 reconstruction problem. The repetition rules are utilized as rules for 'BIM regularization'. After
560 a human-in-the-loop BIM regularization, the precision and recall increased to over 98%, which
561 were close to the ideal values of 100%, while the overall time cost was still less than ICP and
562 CMA-ES.
563
564 Our approach is not weakness free. For example, its scalability needs to be verified in more
565 complicated scenes with different types of repetitive objects. For various scenes, the different
566 sets of repetition formations and algorithm parameter configuration should be comprehensively
567 investigated. In addition, some other limitations of this study should be clarified for future
568 research:
569 1. Utilizing architectural domain knowledge for BIM regularization still requires manual
570 intervention for labeling and clustering BIM components. In the reported experiment, the
571 labeling and clustering of chairs cost 223.7 seconds. For more complicated scenes, the
572 increase in the number of different types of repetitive objects will proportionally increases
573 the time needed for manual intervention. In such situations, end-users face a trade-off
574 between accuracy and speed when adopting automatic and semi-automatic approaches to
575 BIM reconstruction. Therefore, how to further automate labeling and clustering is of great
576 importance to further improve the performance of our proposed method.
577 2. The method is confirmed to be more robust than existing methods, but false positives and
578 false negatives were still witnessed in the results. In order to further improve accuracy
579 without undue penalty in computation time, additional information and architectural domain
580 knowledge such as the function of the building and local regulations could be used in the
581 BIM reconstruction.
582 3. Although the regression and regularization can improve the BIM reconstruction in terms of
583 object recognition and semantics discovery (see Figure 10 and Table 2), the idea of applying
584 the as-designed patterns to a cluster can impose the risk of "regularizing" some as-built
585 errors caused by poor craftsmanship or deformation). In this case, the modeler can simply
586 undo the regularized components back to their as-built status.
587
588 The legitimacy of BIM reconstruction and semantic enrichment needs to be justified in terms
589 of cost-benefit. BIM researchers have developed various reconstruction methods, including
590 'semantic segmentation' and 'semantic registration'. Our study recognizes the strengths and
591 weaknesses of the semantic registration approach and further improves it by introducing

additional algorithms from other disciplines. It demonstrates the power of architectural domain knowledge in improving the effectiveness and efficiency of such approaches. In this sense, the research opens up a new avenue for exploring semantic registration algorithms enhanced by architectural design knowledge.

6 Conclusion

This study advances the realm of semantically rich BIM reconstruction by addressing the widespread challenge of dealing with complicated scenes (e.g., indoor environments with repetitive, irregular-shaped objects, and noisy measurement data as input). An MMO-based semantic registration approach for BIM reconstruction was proposed. The approach consists of three steps: multimodal problem formulation based on repetition in architecture; automatic BIM reconstruction based on MMO; and semi-automatic BIM regularization based on repetition formations. The proposed approach was prototyped and tested in an experiment by following a series of rigorous processes. The experimental results showed that the proposed approach can reconstruct an indoor scene of 293 theater chairs from 1.9 million noisy points with satisfactory accuracy (99.3% precision and 98.0% recall) and less modeling time than previously published algorithms. Our study confirms that the MMO approach, by finding all identical or repetitive objects in one go, is more effective and efficient than traditional 'unimodal' problem solving in BIM reconstruction. The research also confirms that architectural domain knowledge, particularly 'repetition', can further augment an MMO approach to improve efficiency and effectiveness of BIM reconstruction.

The paper thus makes significant contributions to the methodology and practice of advanced BIM technologies. An original methodological contribution is to translate the reconstruction of BIM with repetitive objects into a MMO problem, thereby allowing a number of well-established MMO algorithms to be applied to the problem. The paper endorses and extends the paradigmatic shift from 'semantic segmentation' to 'semantic registration'; advances the approach by introducing MMO. Particularly, another novelty of the research is to make use of architectural design knowledge hidden within a point cloud (i.e., symmetry, repetition, and structure regularity), which can be used to eliminate noise or errors of measurement data and to reduce search space of the formulated problem. Augmenting MMO algorithms with architectural domain knowledge is considered a novel philosophy for semantically rich BIM reconstruction. Practically, our approach is suitable for scaling up and embedding in mainstream BIM platforms, for example, to enable value-added applications such as creating BIMs of architectural design, construction management projects, heritage conservation sites and so on, which require accurate mapping of domain-specific semantics to geometric components.

Future research can be conducted in three directions. First, the effectiveness of the proposed approach should be tested on other complicated cases with less obvious repetitions. Secondly,

more advanced computer vision methods can be developed to improve the semi-automatic or manual labeling and clustering deployed in current approach. Thirdly, other types of architectural domain knowledge, including symmetry, architectural styles, historical building materials and technology, local standards and regulations, and parametric curved surfaces, can be tested in BIM reconstruction.

Acknowledgements

This study was supported in part by two grants from the Hong Kong Research Grant Council (Nos. 17201717, 17200218) and three grants from The University of Hong Kong (Nos. 201711159016, 102009741, 20300083). We wish to express our gratitude to the anonymous reviewers for their constructive comments to improve the quality of the paper.

References

Ahrari, A., Deb, K. & Preuss, M. (2017). Multimodal optimization by Covariance Matrix Self-Adaptation Evolution Strategy with repelling subpopulations. *Evolutionary Computation, 25*(3), 439-471. doi:10.1162/evco_a_00182

Andreopoulos, A. & Tsotsos, J. K. (2013). 50 years of object recognition: Directions forward. *Computer vision and image understanding, 117*(8), 827-891. doi:10.1016/j.cviu.2013.04.005

Armeni, I., Sax, A., Zamir, A. R. & Savarese, S. (2017). *Joint 2D-3D-Semantic Data for Indoor Scene Understanding.* Stanford, CA, USA: ArXiv e-prints.

Babacan, K., Chen, L. & Sohn, G. (2017). Semantic segmentation of indoor point clouds using Convolutional Neural Network. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-4*(W4), 101-108. doi:10.5194/isprs-annals-IV-4-W4-101-2017

Barazzetti, L. (2016). Parametric as-built model generation of complex shapes from point clouds. *Advanced Engineering Informatics, 30*(3), 298-311. doi:10.1016/j.aei.2016.03.005

Bassier, M., Van Genechten, B. & Vergauwen, M. (2019). Classification of sensor independent point cloud data of building objects using random forests. *Journal of Building Engineering, 21*, 468-477. doi:10.1016/j.jobe.2018.04.027

Belsky, M., Sacks, R. & Brilakis, I. (2016). Semantic enrichment for building information modeling. *Computer-Aided Civil and Infrastructure Engineering, 31*(4), 261-274.

Bosché, F., Guillemet, A., Turkan, Y., Haas, C. T. & Haas, R. (2013). Tracking the built status of MEP works: Assessing the value of a Scan-vs-BIM system. *Journal of computing in civil engineering, 28*(4), 05014004. doi:10.1061/(ASCE)CP.1943-5487.0000343

Bueno, M., Bosché, F., González-Jorge, H., Martínez-Sánchez, J. & Arias, P. (2018). 4-Plane congruent sets for automatic registration of as-is 3D point clouds with 3D BIM models. *Automation in Construction, 89*, 120-134.

Chen, K., Lu, W., Peng, Y., Rowlinson, S. & Huang, G. Q. (2015). Bridging BIM and building: From a literature review to an integrated conceptual framework. *International Journal of Project Management, 33*(6), 1405-1416. doi:10.1016/j.ijproman.2015.03.006

Chen, K., Lu, W., Xue, F., Tang, P. & Li, L. H. (2018). Automatic building information model reconstruction in high-density urban areas: Augmenting multi-source data with architectural knowledge. *Automation in Construction, 93*, 22-34. doi:10.1016/j.autcon.2018.05.009

Chen, M. R., Li, X., Zhang, X. & Lu, Y. Z. (2010). A novel particle swarm optimizer hybridized with extremal optimization. *Applied Soft Computing, 10*(2), 367-373. doi:10.1016/j.asoc.2009.08.014

Ching, F. D. (2007). *Architecture: Form, Space, and Order* (3rd ed.). John Wiley & Sons.

Czerniawski, T., Sankaran, B., Nahangi, M., Haas, C. & Leite, F. (2018). 6D DBSCAN-based segmentation of building point clouds for planar object classification. *Automation in Construction, 88*, 44-58. doi:10.1016/j.autcon.2017.12.029

Das, S., Maity, S., Qu, B. Y. & Suganthan, P. N. (2011). Real-parameter evolutionary multimodal optimization—A survey of the state-of-the-art. *Swarm and Evolutionary Computation, 1*(2), 71-88. doi:10.1016/j.swevo.2011.05.005

De Luca, L., Véron, P. & Florenzano, M. (2006). Reverse engineering of architectural buildings based on a hybrid modeling approach. *Computers & Graphics, 30*(2), 160-176. doi:10.1016/j.cag.2006.01.020

Dore, C. & Murphy, M. (2014). Semi-automatic generation of as-built BIM façade geometry from laser and image data. *Journal of Information Technology in Construction, 19*(2), 20-46.

Du, W. B., Gao, Y., Liu, C., Zheng, Z. & Wang, Z. (2015). Adequate is better: particle swarm optimization with limited-information. *Applied Mathematics and Computation, 268*, 832-838. doi:10.1016/j.amc.2015.06.062

Fan, H., Zipf, A. & Wu, H. (2017). Detecting repetitive structures on building footprints for the purposes of 3D modeling and reconstruction. *International Journal of Digital Earth, 10*(8), 785-797.

Fieldsend, J. E. (2014). Running up those hills: Multi-modal search with the niching migratory multi-swarm optimiser. *Proceedings of 2014 IEEE Congress on Evolutionary Computation* (pp. 2593-2600). IEEE. doi:10.1109/CEC.2014.6900309

Fisher, R. (2003). Solving architectural modelling problems using knowledge. *Proceedings of the Fourth International Conference on 3-D Digital Imaging and Modeling.* Banff, Alta., Canada: IEEE. doi:10.1109/IM.2003.1240268

Forrester, A. & Keane, A. (2008). *Engineering design via surrogate modelling: A practical guide.* John Wiley & Sons.

Hidaka, N., Michikawa, T., Motamedi, A., Yabuki, N. & Fukuda, T. (2018). Polygonization of point clouds of repetitive components in civil infrastructure based on geometric similarities. *Automation in Construction, 86*, 99-117. doi:10.1016/j.autcon.2017.10.014

Huber, D., Akinci, B., Oliver, A. A., Anil, E., Okorn, B. E. & Xiong, X. (2011). Methods for automatically modeling and representing as-built building information models. *Proceedings of the NSF CMMI Research Innovation Conference.* Retrieved September 18, 2018, from https://ri.cmu.edu/pub_files/2011/1/2011-huber-cmmi-nsf-v4.pdf

Jung, J., Hong, S., Yoon, S., Kim, J. & Heo, J. (2016). 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, 30*(4), 04015074.

Jung, J., Stachniss, C., Ju, S. & Heo, J. (2018). Automated 3D volumetric reconstruction of multiple-room building interiors for as-built BIM. *Advanced Engineering Informatics, 38*, 811-825. doi:10.1016/j.aei.2018.10.007

Kim, C., Son, H. & Kim, C. (2013). Automated construction progress measurement using a 4D building information model and 3D data. *Automation in Construction, 31*, 75-82. doi:10.1016/j.autcon.2012.11.041

Li, X., Engelbrecht, A. & Epitropakis, M. G. (2013). *Benchmark functions for CEC'2013 special session and competition on niching methods for multimodal function optimization.* Melbourne: Evolutionary Computation and Machine Learning Group, RMIT University. Retrieved from https://titan.csit.rmit.edu.au/~e46507/cec13-niching/competition/cec2013-niching-benchmark-tech-report.pdf

Liu, J. & Wu, Z. (2016). Rule-based generation of ancient Chinese architecture from the Song Dynasty. *Journal on Computing and Cultural Heritage, 9*(2), Article 7. doi:10.1145/2835495

Lu, W., Lai, C. C. & Tse, T. (2019). *BIM and Big Data for Construction Cost Management.* Abingdon, Oxon: Routledge.

Mehta, M., Johnson, J. & Rocafort, J. (1999). *Architectural acoustics: Principles and design.* Englewood Cliffs, NJ United States: Prentice Hall.

Nguyen, C. H. & Choi, Y. (2018). Comparison of point cloud data and 3D CAD data for on-site dimensional inspection of industrial plant piping systems. *Automation in Construction, 91*, 44-52. doi:10.1016/j.autcon.2018.03.008

NIBS. (2015). *National BIM Standard - United States Version 3.* Retrieved September 24, 2018, from http://www.nationalbimstandard.org/nbims-us

Pătrăucean, V., Armeni, I., Nahangi, M., Yeung, J., Brilakis, I. & Haas, C. (2015). State of research in automatic as-built modelling. *Advanced Engineering Informatics, 29*(2), 162-171. doi:10.1016/j.aei.2015.01.001

Previtali, M., Díaz-Vilariño, L. & Scaioni, M. (2018). Towards automatic reconstruction of indoor scenes from incomplete point clouds: Door and window detection and regularization. *ISPRS TC-4 Mid-term Symposium 2018* (pp. 507-514). ISPRS. doi:10.5194/isprs-archives-XLII-4-507-2018

Qu, B. Y., Liang, J. J. & Suganthan, P. N. (2012). Niching particle swarm optimization with local search for multi-modal optimization. *Information Sciences, 197*, 131-143. doi:10.1016/j.ins.2012.02.011

Schlueter, A. & Thesseling, F. (2009). Building information model based energy/exergy performance assessment in early design stages. *Automation in Construction, 18*(2), 153-163.

Shamir, A. (2008). A survey on mesh segmentation techniques. *Computer Graphics Forum, 27*(6), 1539-1556. doi:10.1111/j.1467-8659.2007.01103.x

Son, H. & Kim, C. (2017). Semantic as-built 3D modeling of structural elements of buildings based on local concavity and convexity. *Advanced Engineering Informatics, 34*, 114-124.

760 Tang, P., Huber, D., Akinci, B., Lipman, R. & Lytle, A. (2010). Automatic reconstruction of
761 as-built building information models from laser-scanned point clouds: A review of
762 related techniques. *Automation in Construction, 19*(7), 829-843.
763 doi:10.1016/j.autcon.2010.06.007

764 Thomson, C. & Boehm, J. (2015). Automatic geometry generation from point clouds for BIM.
765 *Remote Sensing, 7*(9), 11753-11775. doi:10.3390/rs70911753

766 Tran, H., Khoshelham, K., Kealy, A. & Díaz-Vilariño, L. (2019). Shape Grammar Approach
767 to 3D Modeling of Indoor Environments Using Point Clouds. *Journal of Computing in*
768 *Civil Engineering, 33*(1), 04018055.

769 Valero, E., Adán, A. & Bosché, F. (2016). Semantic 3D reconstruction of furnished interiors
770 using laser scanning and RFID technology. *Journal of Computing in Civil Engineering,*
771 *30*(4), 04015053.

772 Valero, E., Adán, A. & Cerrada, C. (2012). Automatic method for building indoor boundary
773 models from dense point clouds collected by laser scanners. *Sensors, 12*(12), 16099-
774 16115. doi:10.3390/s121216099

775 Van Kaick, O., Zhang, H., Hamarneh, G. & Cohen-Or, D. (2011). A survey on shape
776 correspondence. *Computer Graphics Forum, 30*(6), 1681-1707. doi:10.1111/j.1467-
777 8659.2011.01884.x

778 Volk, R., Stengel, J. & Schultmann, F. (2014). Building Information Modeling (BIM) for
779 existing buildings—Literature review and future needs. *Automation in construction, 38*,
780 109-127. doi:10.1016/j.autcon.2013.10.023

781 Wang, J., Wu, Q., Remil, O., Yi, C., Guo, Y. & Wei, M. (2018). Modeling indoor scenes with
782 repetitions from 3D raw point data. *Computer-Aided Design, 94*, 1-15.
783 doi:10.1016/j.cad.2017.09.001

784 Wong, K. C., Leung, K. S. & Wong, M. H. (2010). Protein structure prediction on a lattice
785 model via multimodal optimization techniques. *Proceedings of the 12th annual*
786 *conference on Genetic and evolutionary computation* (pp. 155-162). ACM.
787 doi:10.1145/1830483.1830513

788 Xiong, X., Adan, A., Akinci, B. & Huber, D. (2013). Automatic creation of semantically rich
789 3D building models from laser scanner data. *Automation in Construction, 31*, 325-337.
790 doi:10.1016/j.autcon.2012.10.006

791 Xue, F., Lu, W. & Chen, K. (2018). Automatic generation of semantically rich as-built building
792 information models using 2D images: A derivative-free optimization approach.
793 *Computer-Aided Civil and Infrastructure Engineering, 33*(11), 926-942.
794 doi:10.1111/mice.12378

795 Xue, F., Lu, W., Chen, K. & Zetkulic, A. (2019b). From 'semantic segmentation' to 'semantic
796 registration': A derivative-free optimization-based approach for automatic generation
797 of semantically rich as-built building information models (BIMs) from 3D point clouds.
798 *Journal of Computing in Civil Engineering, Accepted, in press, 33*(4), 04019024.
799 doi:10.1061/(ASCE)CP.1943-5487.0000839

800 Xue, F., Lu, W., Webster, C. J. & Chen, K. (2019a). A derivative-free optimization-based
801 approach for detecting architectural symmetries from 3D point clouds. *ISPRS Journal*
802 *of Photogrammetry and Remote Sensing, 148*, 32-40.
803 doi:10.1016/j.isprsjprs.2018.12.005

Zou, C., Colburn, A., Shan, Q. & Hoiem, D. (2018). LayoutNet: Reconstructing the 3D room layout from a single RGB image. *2018 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2051-2059). Salt Lake City, USA: IEEE.