Proceedings of the 2013 Winter Simulation Conference
R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, eds

SIMULATION FOR CHARACTERIZING A PROGRESSIVE REGISTRATION ALGORITHM ALIGNING AS-BUILT 3D POINT CLOUDS AGAINST AS-DESIGNED MODELS

Pingbo Tang

School of Sustainable Engineering and the Built Environment
Arizona State University
Tempe, AZ 85287-0204, USA

Syed Hammad Rasheed

Department of Civil and Construction Engineering
Western Michigan University
Kalamazoo, MI 49008-5316, USA

ABSTRACT

Construction engineers compare as-built data against as-designed models for monitoring construction defects or changes. As laser scanners can collect 3D point clouds as as-built data in a few minutes, engineers start to compare point clouds against the as-designed model. Such comparison requires a reliable data-model registration that precisely distinguishes data-model differences (e.g., displacements) from the well-matched parts. Previously developed registration methods have limitations on aligning two geometries with geometric differences. Target-based registration methods pose challenges of installing targets and ensuring their visibilities on job sites. Feature-based registration algorithms need engineers to manually set proper parameters to precisely reject data-model differences. Through the simulation of a progressive data-model registration process, this study characterizes a progressive 3D registration approach that can precisely reject data-model differences. Sensitivity analysis results of this approach in a case study show that this approach outperforms previous methods in terms of precision without losing substantial computational efficiency.

1 INTRODUCTION

Construction engineers and project managers need to compare as-built data against as-designed models for construction project control. Project engineers need to compare the number of constructed components against the expected number of erected components according to the schedule for progress monitoring (Turkan et al. 2012). It is necessary to evaluate the deviations of as-built conditions from the as-designed model for construction defect detection, quality control, and workspace arrangements (Akinci et al. 2006; Cho et al. 2011). Differences between as-built data and as-designed models also help engineers to analyze which construction operations designed based on as-designed models need to be updated for ensuring safe and productive construction operations (Gong and Caldas 2010; Gong and Caldas 2008).

Comparison of the as-built data and as-designed model needs to firstly align the data against the model, and then identify the deviations of data points from the model (P. Tang et al. 2011). Such data-model alignment is known as *data-model registration,* which transforms the data and model to a common coordinate system (Bosche et al. 2009). Analyzing the data-model deviations reveals various differences between as-built and as-designed conditions, such as dislocations of components. This step is defined as *deviation analysis* in this paper. Without knowing apriori the data-model differences, engineers need a data-model registration algorithm that can precisely distinguish highly-matched parts of the data and model from parts that have substantial differences.

Previous studies reveal the challenges of registering two geometries having significant differences. Some studies use surveying equipment to geo-reference 3D data sets and as-designed models, but require installing targets to be surveyed for such geo-referencing. Target-based registration methods, therefore, pose challenges of installing targets and ensuring their visibilities on job sites (P. Tang et al. 2011). Some studies focus on developing feature-based methods, which extract standard surveying targets or salient building features in the data and model to establish data-model correspondences and derive a least-square best-fit based on corresponding feature pairs.

Feature extraction and matching can be manual (F. Bosché 2011) or automatic (Huber and Hebert 2003). These the registration accuracy of feature-based approaches are sensitive to the layouts and localization accuracies of the used features (Becerik-Gerber et al. 2011). Possible data-model differences can also result in unreliable feature pairs that are "*correspondence outliers*" capable of significantly biasing the registration results. For example, the dislocation of a column on the job site may cause the registration algorithm using the corners of it as data-model correspondences produce imprecise registration. In other words, features on changed components are misleading correspondence information.

Some registration algorithms use all points in point clouds rather than small number of features for computing the best-fit between the data and model. These algorithms are mostly based on the ICP (*Iterative Closest Point*) algorithm, which was developed for registering multiple 3D point clouds (Chetverikov et al. 2002; Besl and McKay 1992; Pomerleau et al. 2010; Zhang and Choi 2011; Rusinkiewicz and Levoy 2001). ICP-based algorithms iteratively update the transformation (rotation, translation) needed to minimize the deviations between two point clouds. The algorithm firstly associates points from one point cloud with their nearest neighbors in the other (*data association*), and then derives the transformation that minimizes the distances between all pairs of associated points. The ICP algorithm iterates these two steps until the sum of the squares of distances between associated points decreases below a user-defined threshold. When adopted for data-model registration, ICP-based algorithms associate data points with the nearest surface in the as-designed model (*data-model association*), and then iteratively minimizes the sum of the squares of point-surface distances. This data-model association step needs users to specify a "maximum distance" to avoid associating points with surfaces far from them, and thus reject possible correspondence outliers. Compared with feature-based approaches, ICP-based approaches adaptively update the data-model correspondences in the data-model association step, and automatically remove correspondence outliers based on the "maximum distance" threshold for improving the registration accuracy.

On the other hand, different data-model deviation cases need different "maximum distance" setting of ICP-based approaches for precisely eliminating correspondence outliers. As will be detailed in section 2, a rough data-model alignment has relatively large data-model deviations for most parts of the point cloud, and has more ambiguities about which parts contain outliers. Such cases need a larger "maximum distance" for exploring more possibilities of data-model matches. In contrast, a refined data-model alignment has substantial parts of point clouds aligning well with the model, and the rest has large deviations. Such case needs a smaller "maximum distance" value for further refining the registration rather than a broad search. ICP-based algorithms using a fixed "maximum distance" throughout the registration process ignore the needs of adjusting the "maximum distance" based on the present alignment conditions, and thus lose registration accuracy due to imprecise outlier rejections.

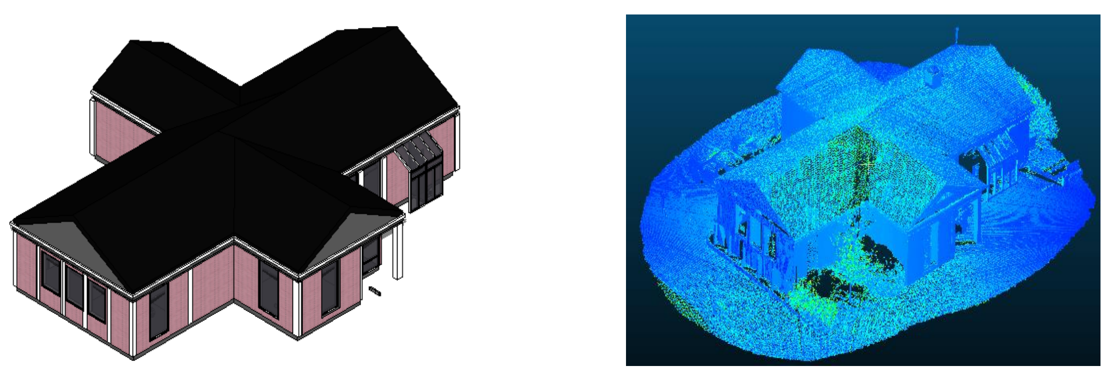
Some researchers developed *robust ICP algorithms* that adaptively adjust the "maximum distance" thresholds during the ICP iterations based on the data-model deviation values of interim registration results. Robust ICP algorithms can achieve more reliable outlier rejections, but still require engineers to configure the initial value of "maximum distance" and the sub-sampling rate of the point clouds. Generating data-model deviations for each loop in the ICP process also consumes large amounts of computing time and resources. On the other hand, quantitative knowledge about how various parameters of robust ICP algorithms influence the registration precision and computational efficiency are still limited. As a result, engineers tend to configure ICP algorithms in an ad-hoc manner without quantitative awareness on the time and registration accuracy implications of settings of the algorithm. Section 2 details that proper parameter setting of robust ICP is critical for ensuring the precision and efficiency of the data-model registration.

This paper presents a framework for quantitatively evaluating the performances of data-model registration algorithms and characterizing how the performance of a robust ICP algorithm designed by the authors varies with the values of its critical parameters. This framework uses three metrics for measuring the performance of data-model registration algorithms: 1) computational efficiency, which indicates the execution time of the algorithm; 2) accuracy, which indicates the capability of an algorithm of finding the best fit while correctly rejecting outliers; 3) repeatability, which measures the variations of the registration results in multiple runs. The robust ICP algorithm designed by the authors repeatedly executes the classical ICP algorithm, and each execution uses a "maximum distance" value halving the "maximum distance" used in the last execution. This algorithm is called "*Progressive Registration*" hereafter due to this progressive approach of reducing the "maximum distance" threshold. The first ICP execution of this algorithm uses a user-defined maximum distance (1 m in this paper). This algorithm then uses a predefined sequence of "maximum distance" values for a series of ICP executions, while does not vary the "maximum distance" in the internal loops of each ICP execution. It eliminates the data-model deviation computations, while sacrificing some computational efficiency by repetitive ICP executions. Another critical parameter of this algorithm is the sub-sampling rate of the registered point cloud: more sub-sampling will reduce the computing time while compromising the registration accuracy. The evaluation of the performance of this algorithm, therefore, focuses on understanding the impacts of the series of the predefined series of "maximum distance" values, and sub-sampling rate on the three metrics mentioned above.

The following sections will firstly introduce a motivating case showing the necessity of exploring and characterizing robust ICP methods (section 2). After that, section 3 presents the design of the progressive registration algorithm and the experiment design for evaluating it. Section 4 discusses the characterization results of the progressive registration algorithm on the data of two campus buildings of significantly different sizes. Section 5 finally concludes with some observations and possible future research directions.

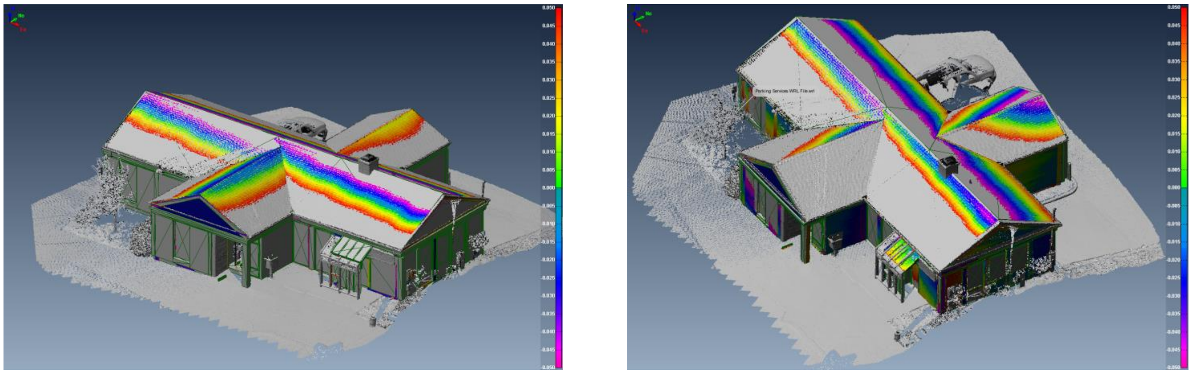
2 MOTIVATING CASE

The Facilities Management (FM) at Western Michigan University (WMU) worked jointly with the authors on evaluating the qualities of a number of as-is Building Information Models (as-is BIM) created based on as-built drawings of a number of WMU campus buildings. This effort is part of the WMU Bronco BIM initiative on active uses of BIM for campus facility management. WMU FM plans to use these as-is BIMs for planning renovations, analyzing energy performances of facilities, and other applications related to the life-cycle management of campus facilities. These applications have a variety of requirements about the geometric accuracies of as-is BIMs for supporting the spatial analysis and decision making. WMU would like to identify significant modeling errors in these as-is BIMs. For example, planning the exterior renovation of a campus building (Figure 1a) may require the deviations of its as-is BIMs from the actual conditions to be less than 5 cm according to the U.S. GSA (General Services Administration) BIM guide (General Services Administration 2009). WMU FM, therefore, needs to know whether all parts of the as-is BIM of this building is within 5 cm from the actual physical conditions.



(a) (b)

Figure 1 As-is BIM (a) and dense 3D point clouds (b) collected for identifying the as-is BIM modeling errors



(a) (b)

Figure 2 (a) The alignment result of ICP algorithm using a maximum distance of 1m, in which the color scale is from -5 cm (magenta) to +5cm (red), and the points beyond these limits are of the default color, grey; (b) The alignment result of ICP algorithm using a maximum distance of 5 cm.

To assist WMU FM in this case, the authors compared the dense point clouds collected by a Time-of-Flight terrestrial laser scanner (TOF scanner). This TOF scanner collected point clouds capturing the geometries of this building with 5 mm accuracy for each point. These point clouds are also dense enough to have all building features larger than 1 cm documented. We then registered these point clouds collected at several locations around this building into one point cloud capturing the comprehensive exterior geometries of this building. After that, we registered the integrated point cloud and the as-is BIM into the same coordinate system for identifying differences between the as-is BIM and the physical conditions. More details of the data collection and processing of this case study are in (Pingbo Tang and Alaswad 2012) and (Alaswad 2011). First, we roughly aligned the point cloud and the BIM based on the corner points along the top sides of the walls (twelve top corner points of walls for this "crux" shape building). After this rough alignment, the authors refined the alignment using the ICP algorithm with 1 m as the "maximum distance," and found that the best-fit found by this ICP is sub-optimal due to the significant modeling errors in the as-is BIM. Figure 2(a) shows this 1m ICP result. In this figure, we can see that only parts of the roof are within 5 cm from the nearest 3D points in the point cloud (color points are points within 5 cm from the BIM surfaces), while most parts of the walls show no color as they are not align well with the point cloud such that no points are within 5 cm from them. Close inspection of this alignment results showed that such misalignment was caused by the significant modeling errors in the roof slope: the actual roof slopes are much smaller than the ones in the BIM. Significant differences in roof slopes cause the ICP algorithm to shift the model away from a good alignment for walls, while trying to minimize the distances between points on roofs and the incorrectly modeled roof. On the other hand, we found that using 5 cm as the "maximum distance," the alignment results seem to be much better than the 1m ICP result as most of the walls start to have colors (Figure 2b). This is because that the roughly aligned point clouds are based on the top corners of walls, and the maximum distance of 5 cm ignores significant parts of the roof because many parts of it have deviations between points and BIM larger than 5 cm.

In this case study, we also found that repeating the ICP algorithm with progressively reducing maximum distances can gradually reject outliers and improve the registration results. Unfortunately, we also found that the registration quality and computing time was sensitive to the values of the "maximum distance" and the sub-sampling rates used in by each ICP execution in the progressive registration process. For evaluating the quality of a given data-model registration, we computed the shortest distance from each point in the point cloud to the as-is BIM, and identified the percentage of the point cloud lying within 5 cm from the as-is BIM. That "5 cm" is the tolerance of positioning accuracy for exterior walls according to GSA BIM guideline (U.S. GSA 2009). We call that percentage as "*5 cm fit percentage*" hereafter. A higher

5 cm fit percentage generally indicates a higher level of geometric agreement between the data and as-is BIM, and thus more precise registration ignoring significant differences between data and the as-is BIM.

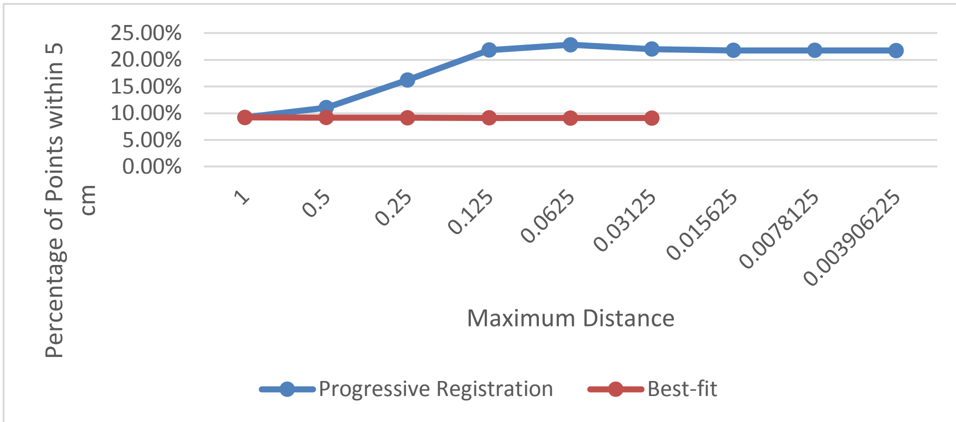


Figure 3 The change of the "5 cm fit percentage" along a progressive registration process, and the corresponding results for using ICP algorithm adopting a single "maximum distance" value (the curve of "best fit," the unit of "maximum distance" is meter)

Figure 3 shows that how the percentage of points in the integrated point cloud that are within 5 cm of the as-is BIM varies with a progressive registration process. Overall, the larger this percentage, the more points in the registered point cloud align well with the as-is BIM (within 5 cm), and thus the registration is more likely to be a good-fit agreed on by more building parts. The horizontal axis shows the sequence of maximum distance values used for executing the ICP process. The progressive registration algorithm executes the ICP process using the first value, and then repeats ICP using the second value, and so on. This figure shows how the percentage of agreeing points vary along the progressive registration process for two sequence of "maximum" distance values: the first one gradually reduce the maximum distance value by halving the maximum distance value for each execution with the initial maximum distance set as 1 m (1m, 0.5 m, ….), while the second one use 1 m maximum distance to execute the ICP process, and then use 3.125 cm to execute the second ICP process. This figure show that the former can gradually improve the percentage of points agree with the as-is BIM, while latter seem not be able to achieve that. Such results show the sensitivity of the registration result to the selection of a sequence of "maximum distances" for gradually rejecting outliers and achieving a reliable registration rejecting outliers.

In this case study, we also found that the computing time can significantly vary with the subsampling rates of point clouds. All above results are generated based on 1/1 sampling rate (no-sub-sampling). Subsampling the point cloud can reduce the registration time but cause the registration results less reliable, as different registration process uses different set of randomly sampled points for ICP, and executing the algorithm multiple times can produce slightly different registration results. For example, we found that using a sampling rate of 1/64 only takes 5 seconds to complete the registration based on progressive series 1, while need 140 seconds if the sampling rate is 1/1. As detailed later in section "Results," if the user repeat the progressive registration process multiple times, the registration results will have some variations due to the randomness of the subsampling process. 1/64 subsampling rate will generate registration results of larger variations across multiple executions of the progressive registration process compared with denser subsampling rates (e.g., 1/4, 1/16).

All above discussions show the limitations of the ICP algorithm and the sensitivity of the progressive registration algorithm to the parameter settings. The ICP algorithm uses all data points, while substantial differences between the as-built and modeling conditions can bias the registration results so that the alignment result may not correctly align the data and model for disregarding errors. The progressive registration algorithm can address the limitation of ICP only when the parameters are appropriately configured. In addition, the sub-sampling rate of registered point cloud can significantly influence the progressive registration time as well as the reliability of the registration results. These observations motivate the development of a framework for systematically characterizing data-model registration algorithms and quantify the relationships among parameter settings, data set characteristics, the execution time, and the accuracy and reliability of the registration results. With such efforts, engineers will be able to correctly use automated data-model registration approaches with the awareness of the impacts of data collection options (sub-sampling rate, variation of the maximum distance of data-model association).

3 METHODOLOGY

The subsections below first describe the progressive registration algorithm, and then present a framework for evaluating the performance of data-model registration algorithms. That evaluation framework is composed of a number of algorithm performance metrics, testing data sets and parameter settings assessed in this performance evaluation.

3.1 Progressive Registration Algorithm

The progressive registration algorithm (PRA) explored in this study execute ICP algorithm multiple times until a data-model fitness measure is maximized.

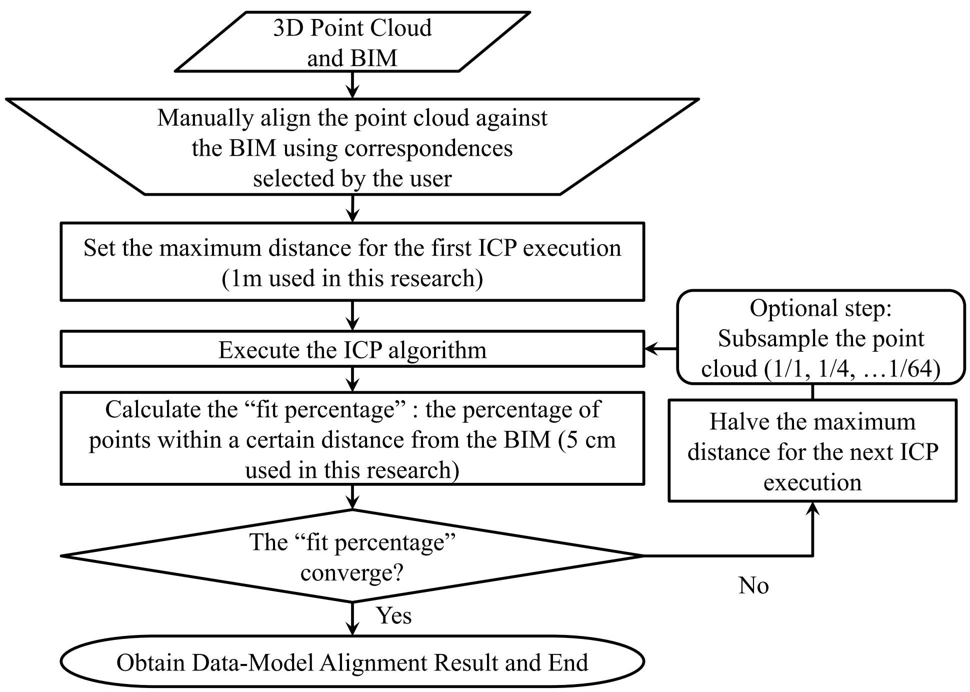


Figure 4 Flowchart of the progressive registration algorithm

Figure 4 shows a flowchart of this algorithm. Overall, given a user-defined “maximum-distance” value (1m used in this research) and manually aligned point cloud and BIM, this algorithm completes the first execution of ICP, and then each subsequent execution of the ICP algorithm uses a maximum-distance value that is half of the maximum-distance value used in the last ICP execution. As a result, a series of ICP registrations using a series of progressively halved “maximum distance” values form the progressive registration process. After each execution of ICP, the algorithm evaluates the data-model fitness using a predefined measure. The data-model fitness measure used in this study is “5 cm fit percentage” described in section 2: the percentage of the point cloud falling within 5 cm from the as-designed model. This 5 cm threshold is based on the tolerance of BIM accuracy in this case study, and can be adjusted based on specific domain requirements if other projects had different tolerances. Large differences between the data and the BIM can reduce this percentage, and maximizing it can likely find the maximum agreement between the data and model having possible inconsistencies. As each subsequent registration is a refinement of the results obtained by its preceding execution, this "5 cm fit percentage" will gradually increase during the progressive registration process, but will not increase anymore after such refinements continue to a certain point, as shown in section 2. The repetitive executions of ICP will stop when the "5 cm fit percentage" does not vary significantly. In this research, if the "5 cm fit percentage" of the current ICP execution is within ±0.5% from the "5cm fit percentage" from the last ICP execution, then the algorithm will regard the "fit percentage" converges and stop the progressive ICP iterations. As the first ICP execution uses 1 m as the maximum distance, the tested sequence of maximum distance values in this research use the sequence of (1 m, 0.5 m, 0.25 m, …), and in most test cases, the progressive registration stops when the maximum distance value being halved achieves 0.015625 m (1 m / 64, which is 2^6) in the 6^{th} execution of the ICP process. Figure 3 shows that for the motivating case, the slope of the curve of "progressive registration" results is nearly zero (very close to a horizontal line) while progressively reaching 0.15625m.

An optional step in the progressive registration process is the subsampling of the point cloud. For each execution of ICP, the algorithm can randomly subsample the point cloud for reducing the computing time. On the other hand, random subsampling of point clouds introduce uncertainties, as using different subsampled data point sets would lead to different best-fit ICP result in every iteration. Sparer/higher sub-sampling rate will bring higher level of uncertainties and thus larger variations of the registration results. In this research, we tested the following subsampling rates: 1 out of 1 (no subsampling), 1 out of 4, 1 out of 16, and 1 out of 64 (subsample 1 point out of each 8 pixel by 8 pixel region of a 3D image).

3.2 A Framework for Evaluating the Progressive Registration Algorithm

The framework for evaluating the progressive registration algorithm is composed of two aspects: 1) metrics for evaluating the performance of the progressive registration algorithm (PRA); 2) the experiment design for this performance evaluation. The metrics indicate the domain requirements and guide the experiment design. The experiment design includes the preparation of the testing data sets, and the parameter settings to be evaluated in this research.

3.2.1 Metrics for evaluating the performance of PRA

Based on extensive literature review and experiments with PRA, the authors found that the performance of PRA can be quantified along the following dimensions: 1) efficiency, which indicates the computational complexities of algorithms for estimating computing time of them on certain computing platforms; 2) reliability, which indicates whether the registration results vary significantly if we execute the registration algorithms on the same data-model pairs for multiple times using slightly different initial manual data-BIM alignments (repeatability).

Above discussions in section 3.1 indicate that three parameters of PRA significantly influence its efficiency: 1) the "maximum distance" value determined by the user and used by the first execution of the ICP algorithm; 2) the definition of a measure for indicating the data-BIM fitness; 3) the "sub-sampling rate." This research focuses on understanding the impacts of "the subsampling rate" on the computational efficiency of the progressive registration algorithm. For measuring the efficiency of PRA, the authors measured the durations of PRA executions using different subsampling rates, having the "maximum distance" set as 1 m, and the "5 cm fit percentage" as the data-BIM fitness measure.

This research defines the "reliability" as the variation of data-BIM alignment results regardless of the initial manual data-BIM alignment and the sub-sampling value used. The parameters of a 3D registration result include three translational (translations along X, Y, and Z axes) and rotational parameters (α, β, and γ indicating the rotational angles around X, Y, and Z axes). Random subsampling of the point cloud in the PRA process will cause uncertainties and result in slightly different 3D registration parameters when the user execute PRA multiple times on a given point cloud. Another uncertainty comes from the initial data-BIM alignment used by PRA. The subjective manual correspondence selections influence the manual data-BIM alignment result, which is needed as the input of PRA. According to the authors' extensive experiments, PRA sessions using different manual data-BIM registration generate different registration results and execution time. It is therefore necessary to test whether slightly different manual data-BIM registration significantly influence the reliability and efficiency of PRA.

In this research, the authors manually aligned the point cloud against the BIM for five times, and executed PRA for 30 times for each manual alignment. Each manual alignment used twenty pairs of manually selected correspondences between the point cloud and BIM for ensuring a certain level reliability of the alignment. Because 30 is the minimum number for valid statistical assessments (Diez 2009), the authors chose 30 as the number of repeating PRA for acquiring statistically significant performance assessment results. For all six 3D registration parameters, the authors generate boxplots of them to assess the variations of them across different PRA execution sessions using different subsampling rates. For 30 results of a given 3D registration parameter, the box of a box plot will show three lines at the lower quartile, median, and upper quartile values (Diez 2009). A box plot will also include Whiskers extend from each end of the box to the largest/smallest value lying within 1.5 times the box height (distance between the lower quartile and the upper quartile values).

3.2.2 Experiment Design

The experiment design includes two main aspects: 1) the data sets used for performance evaluation of the progressive data-model registration algorithm; 2) parameter settings to be tested for performance characterization. This research used the point cloud collected from a campus building shown in Figure 1. As discussed above, the authors focus on understanding the impacts of the subsampling rate on the performance PRA, because this parameter showed more significant impacts on the registration time and accuracy compared with the other two parameters, which are the initial "maximum distance" value and the threshold value for computing the "fit percentage." The authors fixed 1 m as the initial maximum distance and 5 cm as the threshold value for computing a "5 cm fit percentage" in this paper. Currently, the authors are collecting data on larger campus buildings and using them to conduct more assessments of PRA, but due to the space limit of this paper, these larger test cases are not introduced in this paper.

4 RESULTS AND DISCUSSIONS

4.1 Efficiency Analysis of Progressive Registration Algorithm

Table 1 below shows the efficiency analysis results of PRA. The authors executed PRA for 30 times on each manual data-BIM alignment. The table shows the average "5cm fit percentage" and execution time of the 30 executions for each manual alignment using a specific subsampling rate. These results show that the subsampling rate can significantly influence the computing time without seriously influence the "5cm fit percentage," which to some extent is a measure of the overall agreement between the point cloud and BIM. This observation is important because it means that the users can save significant amounts of time while still keeping certain level of registration accuracy. For example, for manual alignment 1, using 1/64 as the subsampling rate will only use 14.78% of the computing time needed for the test case of 1/1 (no subsampling), while the fit percentage almost does not change (from 22.88% to 22.89%).

The authors noticed that the initial manual alignment do have some impacts on the computing time and the 5cm fit percentage. Such impacts are insignificant for dense sampling rates 1/1 and 1/4, because the results have very slight variations across different manual alignments while using 1/1 and 1/4 as the subsampling rate. On the other hand, cases using sparse subsampling rates show some significant variations in the results across multiple manual alignment results.

Table 1 The impacts of subsampling rates on the computing time of progressive registration: each "fit percent(age)" and computing time shown in this table is the average of 30 executions, because the authors repeated the PRA for 30 times on each manual data-BIM alignment

Sub-sampling Rate	Manual Alignment 1		Manual Alignment 2		Manual Alignment 3		Manual Alignment 4		Manual Alignment 5	
	Fit Percent	Time (sec)	Fit Percent	Time (sec)	Fit Percent	Time (sec)	Fit Percent	Time (sec)	Fit Percent	Time (sec)
1 by 1	22.88%	203	22.88%	211	22.88%	222	22.80%	209	22.88%	201
1 by 4	22.96%	69	22.95%	60	22.95%	54	22.97%	66	22.96%	64
1 by 16	22.93%	32	22.93%	47	22.94%	28	22.93%	73	22.94%	23
1 by 64	22.89%	30	22.99%	13	22.94%	11	23.01%	12	23.01%	12

4.2 Reliability Analysis of Progressive Registration Algorithm

Figure 5 below shows the box plots visualizing the variations of the PRA executions using different subsampling rates. Using each tested subsampling rate, the authors executed the PRA for 30 times on each manual data-BIM alignment. This research tested five manual data-BIM alignments, so that each box plot shown in the figures below indicate the distribution of 150 PRA results (details explained in the caption of Figure 5). These figures shows how subsampling rates influence the variations of six 3D registration parameters: three translational parameters along X, Y, and Z axes, and three rotational angles around X, Y, and Z axes.

These figures show that relatively larger variations exist in the cases using sparser subsampling rates (1/16 and 1/64). Such observation is obvious in the figures of the translational parameter along Z direction, and the rotational angle around the Y axis (β), as these two parameters shows larger variations compared with other parameters. The authors are still investigating why the Z translation and β show larger variations compared with other parameters. Considering the Z translation, one possible explanation is that vertical walls provide less clues for identifying the precise Z translation for aligning the point cloud against the BIM. Considering the β value, the authors found that the Y axis is along two flat walls with very limited features, while the X axis is along the facades having two dislocated walls, which may provide more salient features for the registration algorithm to pinpoint the precise rotation angle around X axis. Further investigations are needed to understand the complexities of geometric features and the registration accuracy. Observing the results here and the efficiency results above, the authors found that the subsampling rate of 1/4 can save data processing time without observable losses of reliability.

Another interesting observation is that results from all cases using 1/1 subsampling rate (no subsampling) have several outliers. This is abnormal: the randomness of selecting points should not exist if the algorithm did not subsample the point cloud, and all registration results should converge to the same translational and rotational parameter values. These outliers indicate that some random factors other than the randomness of subsampling the point cloud and the manual data-BIM alignments exist. Future research needs to further examine these outliers and identify those unknown random factors in PRA.

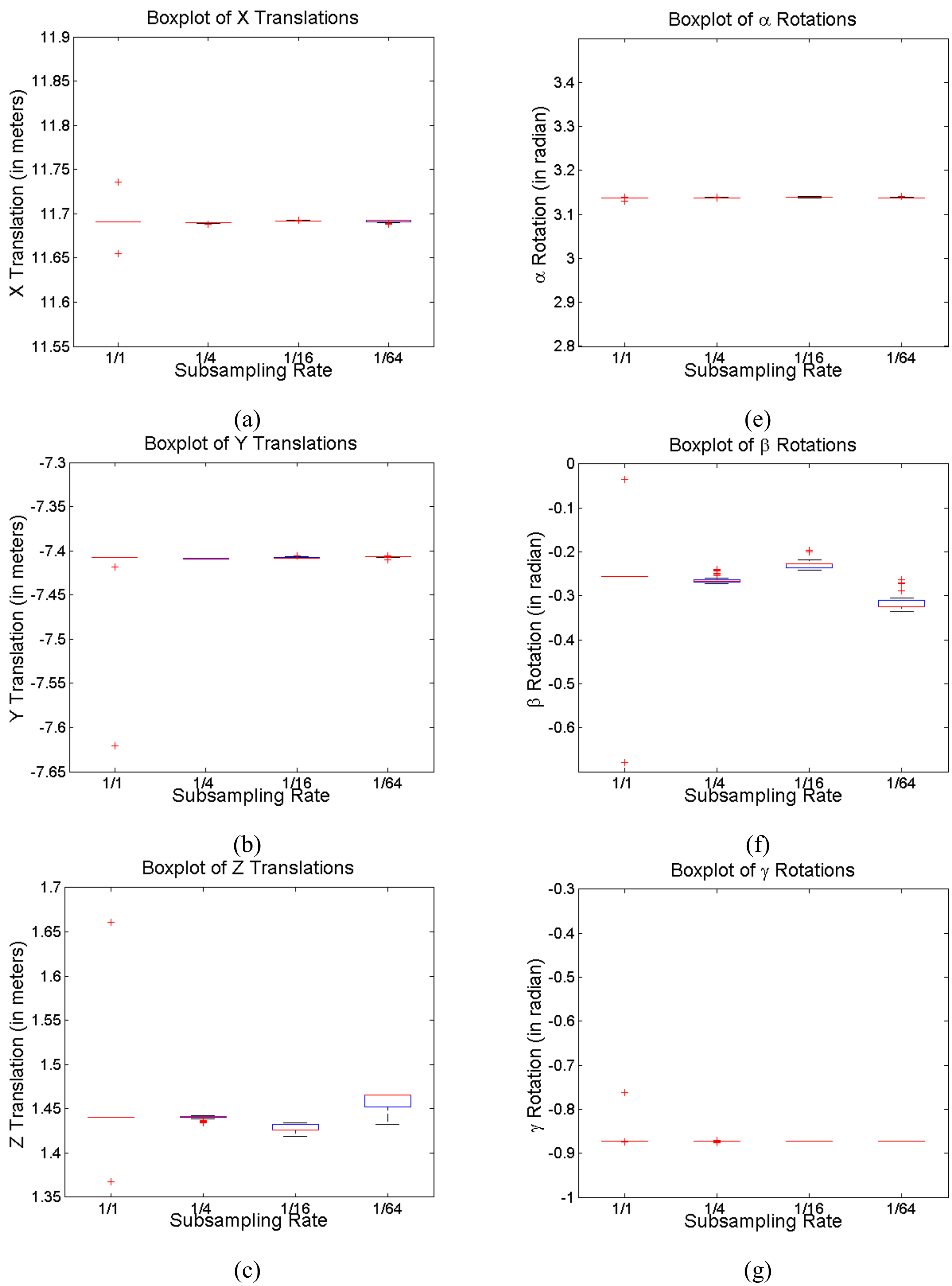


Figure 5 The impacts of subsampling rates on the reliability of the progressive registration results on five different manual data-BIM alignments: each box plot in this figure shows the distribution of 150 registration results (30 repetitions of PRA on each manual alignment × 5 manual alignments).

5 CONCLUSIONS AND FUTURE RESEARCH

This research proposes a progressive 3D registration algorithm to overcome the challenges of aligning 3D laser scanning point clouds against as-designed models having significant geometric differences from the physical conditions. The major contribution is the development of a framework for characterizing the performance of this progressive registration algorithm (PRA), which is composed of two performance metrics for measuring the computational efficiency and reliability of the PRA, and an experiment design for carrying out the tests using this framework. The results shows that subsampling the point cloud (sample 1 point out of 64 points) will substantially reduce the data processing time without significantly compromising a measure indicating the percentage of agreement between the point cloud and the as-designed model (the percentage of points within 5 cm from the as-designed model). This finding will help engineers to save data processing time without worrying about significant losses of registration accuracy. The reliability testing results show that sparser subsampling rates do cause larger variations in the registration results when the users execute the PRA multiple times. Using 1/4 as the subsampling rate can save data processing time without noticeable losses in the efficiency and reliability.

Future research will conduct more tests on larger buildings and further analyze how geometric complexities of buildings influence the performance of PRA, and why a few registration results of the cases using 1/1 subsampling rate deviate significantly from most registration results (more details in section 4.2). In addition, the authors will explore how to integrate PRA with target-based approach for combining their merits: improving the accuracy of registration while reducing the number of targets needed on job sites.

REFERENCES

Akinci, B. et al. 2006. «A formalism for utilization of sensor systems and integrated project models for active construction quality control». *Automation in Construction* 15 (2) (martxoak): 124–138. doi:10.1016/j.autcon.2005.01.008.

Alaswad, Fahd. 2011. «Sensor Modeling and Cost Benefit Of Using Laser Scanning Technology in Architecture/Engineering/Conctruction». Western Michigan University.

Becerik-Gerber, Burcin et al. 2011. «Assessment of target types and layouts in 3D laser scanning for registration accuracy». *Automation in Construction* (otsailak). doi:10.1016/j.autcon.2010.12.008. http://dx.doi.org/10.1016/j.autcon.2010.12.008.

Besl, P.J. and McKay, N.D. 1992. «A method for registration of 3-D shapes». *IEEE Transactions on pattern analysis and machine intelligence* 14 (2): 239–256. http://dl.acm.org/citation.cfm?id=132022.

Bosche, F. et al. 2009. «Automated recognition of 3D CAD objects in site laser scans for project 3D status visualization and performance control». *Journal of Computing in Civil Engineering* 23: 311. http://link.aip.org/link/?JCCEE5/23/311/1.

Bosché, F. 2011. «Plane-based registration of construction laser scans with 3D/4D building models». *Advanced Engineering Informatics*.

Bosché, Frédéric. 2012. «Plane-based registration of construction laser scans with 3D/4D building models». *Advanced Engineering Informatics* 26 (1) (urtarrilak): 90–102. doi:10.1016/j.aei.2011.08.009. http://dx.doi.org/10.1016/j.aei.2011.08.009.

Chetverikov, D et al. 2002. «The Trimmed Iterative Closest Point algorithm». *Object recognition supported by user interaction for service robots* 3 (c): 545–548. doi:10.1109/ICPR.2002.1047997.

Cho, Yong K. et al. 2011. «Target-Focused Local Workspace Modeling for Construction Automation Applications». *Journal of Computing in Civil Engineering* 1 (1) (irailak): 116. doi:10.1061/(ASCE)CP.1943-5487.0000166.

Diez, David. 2009. *Intro to Spatial Statistics in R*. Springer Berlin / Heidelberg.

FRANASZEK, M; CHEOK, G and WITZGALL, C. 2009. «Fast automatic registration of range images from 3D imaging systems using sphere targets». *Automation in Construction* 18 (3) (maiatzak): 265–274. doi:10.1016/j.autcon.2008.08.003. http://dx.doi.org/10.1016/j.autcon.2008.08.003.

General Services Administration. 2009. «GSA Building Information Modeling Guide Series: 03 – GSA BIM Guide for 3D Imaging». Washington, D.C. http://www.gsa.gov/portal/content/102282.

Gong, J. and Caldas, C.H. 2008. «Data processing for real-time construction site spatial modeling». *Automation in Construction* 17 (5) (uztailak): 526–535. doi:10.1016/j.autcon.2007.09.002. http://linkinghub.elsevier.com/retrieve/pii/S0926580507001136.

———. 2010. «Computer vision-based video interpretation model for automated productivity analysis of construction operations». *Journal of Computing in Civil Engineering* 24: 252. http://link.aip.org/link/?JCCEE5/24/252/1.

Huber, D.F. Daniel F and Hebert, Martial. 2003. «Fully automatic registration of multiple 3D data sets». *Image and Vision Computing* 21 (7): 637–650. http://www.sciencedirect.com/science/article/pii/S026288560300060X.

Pomerleau, F. et al. 2010. «Relative motion threshold for rejection in ICP registration» in . *Field and Service Robotics*, 229–238. Springer. http://www.springerlink.com/index/V12R40266321277K.pdf.

Rusinkiewicz, S and Levoy, M. 2001. «Efficient variants of the ICP algorithm». *3-D Digital Imaging and Modeling,*. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=924423.

Tang, P. et al. 2011. «Efficient and Effective Quality Assessment of As-Is Building Information Models and 3D Laser-Scanned Data» in Zhu, Yimin and Issa, Raymond (arg.). *Computing in Civil Engineering (2011)*, 486–493. Miami, FL: ASCE. http://link.aip.org/link/doi/10.1061/41182(416)60/html.

Tang, Pingbo and Alaswad, Fahd Saleh. 2012. «Sensor Modeling of Laser Scanners for Automated Scan Planning on Construction Jobsites» in Amr Kandil and Cai, Hubo (arg.). *Construction Research Congress 2012*. West Lafayette, IN: ASCE.

Turkan, Yelda et al. 2012. «Automated progress tracking using 4D schedule and 3D sensing technologies». *Automation in Construction* 22 (0) (martxoak): 414–421. doi:10.1016/j.autcon.2011.10.003. http://www.sciencedirect.com/science/article/pii/S0926580511001956.

U.S. GSA. 2009. «GSA BIM Guide for 3D Imaging». Washington, D.C.: U.S. General Services Administration. http://www.gsa.gov/graphics/pbs/GSA_BIM_Guide_Series_03.pdf.

Zhang, L and Choi, SI. 2011. «Robust ICP Registration using Biunique Correspondence». *3D Imaging, Modeling,*. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5955346.

AUTHOR BIOGRAPHIES

PINGBO TANG is an assistant professor in Del E. Webb School of Construction at Arizona State University. He obtained his Bachelor Degree of Civil Engineering and Master Degree of Bridge Engineering, both from Tongji University, Shanghai, China. He obtained is PhD from Carnegie Mellon University in 2009. His research explores the remote sensing and information modeling technology (e.g., BIM) to support spatial analyses needed for effective management of construction sites and civil infrastructure systems. He holds memberships of several professional organizations, which include: TRB (Committee on Bridge Management, AHD35), ASCE, IABSE, ASPRS, and ASTM International (Committee E57: 3D imaging systems). He is serving as an associate editor of ASCE Journal of Computing in Civil Engineering, as well as a reviewer of multiple top journals and conferences in the general area of IT and Computing in Civil and Construction Engineering. He won one of the best paper award recipients of Construction Research Congress, ASCE, 2009, and the best poster award of Construction Industry Institute's 2011 Annual Conference. His email address is tangpingbo@asu.edu and his web page is https://sites.google.com/site/tangpingbo/.

SYED HAMMAD RASHEED is a graduate student in the Department of Civil and Construction Engineering at Western Michigan University. His research interests lie in sensing and information technology for construction and facility management. He is a student member of ASCE. He holds a Bachelor Degree in Civil Engineering from NED University of Engineering and Technology, Pakistan. His email address is syedhammad.rasheed@wmich.edu.