Automation in Construction xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Deviation analysis method for the assessment of the quality of the as-is Building Information Models generated from point cloud data

Engin Burak Anil [a,*], Pingbo Tang [b,1], Burcu Akinci [a,2], Daniel Huber [c,3]

[a] Department of Civil and Environmental Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, USA
[b] School of Sustainable Engineering and the Built Environment, Arizona State University, P.O. Box 870204, Tempe, AZ, 85287-0204, USA
[c] Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:
Accepted 20 June 2013
Available online xxxx

Keywords:
As-is Building Information Modeling
BIM
Laser scanning
Point clouds
Quality assessment
Quality inspection

ABSTRACT

Generating three-dimensional (3D) as-is Building Information Models (BIMs), representative of the existing conditions of buildings, from point cloud data collected by laser scanners is becoming common practice. However, generation of such models currently is mostly performed manually, and errors can be introduced during data collection, pre-processing, and modeling. This paper presents a method for assessing the quality of as-is BIMs generated from point cloud data by analyzing the patterns of geometric deviations between the model and the point cloud data. The fundamental assumption is that the point cloud and the as-is BIM generated from the point cloud should corroborate in the depiction of the components and their spatial attributes. Major geometric deviations between as-is models and point clouds can indicate potential errors introduced during data collection, processing and/or model generation. The research described in this paper provides a taxonomy for patterns of deviations and sources of errors and demonstrates that it is possible to identify the source, magnitude, and nature of errors by analyzing the deviation patterns. The method is validated through a comparison with the currently adopted physical measurement method in a case study. The results show that the deviation analysis method is capable of identifying almost six times more errors with more than 40% time savings compared to the physical measurement method.

1. Introduction

Generating Building Information Models (BIMs) from point cloud data obtained by laser scanners is becoming a common practice as many government agencies and private owners are demanding accurate and up-to-date information related to the current conditions of their facilities. As-is BIM generation is composed of several steps, including data collection, pre-processing, and modeling [1]. Inadvertent errors can happen in any step of this process, affecting the quality of the final product. Therefore, the stakeholders need to conduct quality assessment (QA) of as-is BIMs to verify the completeness and accuracy of the model. In this paper, we specifically focus on the QA of as-is BIMs generated from laser scan data.

This paper introduces the deviation analysis method as a way of assessing the quality of as-is BIMs generated from the laser scanned point cloud data. This method classifies and analyzes the patterns of deviations between the data and the derived as-is BIM. Such classification and analysis provide insights into the source, type, and magnitude of errors in less time with greater coverage compared to previous methods of quality assessment of as-is BIMs.

A good quality-assessment (QA) method should exhibit several desirable properties. First, the QA method should be able to comprehensively assess the accuracy and completeness of the model and data based on project requirements [2,3]. The end uses of the BIM might dictate the required level of detail for the model and its accuracy [4]. Second, it is beneficial to obtain intermediate results as the project progresses. Errors can potentially be corrected if identified early in the process. For example, data can be re-collected with a different scanner if a calibration error is identified. Third, there is a need to identify the sources of errors so that the errors can be rectified effectively in a timely fashion. Finally, a good QA procedure should be easy to learn and efficient in time and resource requirements.

The deviation analysis method has been developed in order to satisfy the aforementioned desired properties of a good QA method. A fundamental assumption of this method is that the BIM and the point cloud should geometrically agree to within an application- or project-related tolerance. This is a reasonable assumption, since the model is generated based on the point cloud. Geometric inconsistencies beyond that tolerance point to potential errors. Similarly, individual point clouds should, within a tolerance, geometrically agree in overlapping regions. By computing the distances between a model

* Corresponding author. Tel.: +1 412 482 2013.
E-mail addresses: eanil@andrew.cmu.edu (E.B. Anil), tangpingbo@asu.edu (P. Tang), bakinci@cmu.edu (B. Akinci), huber@cs.cmu.edu (D. Huber).
[1] Tel.: +1 480 727 8105.
[2] Tel.: +1 412 268 2959.
[3] Tel.: +1 412 268 2991.

0926-5805/$ – see front matter
http://dx.doi.org/10.1016/j.autcon.2013.06.003

and the point cloud or between individual scans and by coloring these distances color-coded deviation maps are generated. Different errors result in different deviation patterns, which can be analyzed to pinpoint sources, types, and magnitudes of errors within the point cloud data or derived as-is BIMs.

The deviation analysis method procedure is composed of three main steps (Fig. 1). The first step is computing the deviations of the point cloud to the as-is BIM (Fig. 2a and b). The main challenge in this step is finding the correspondences between the point cloud and the BIM. Directly obtaining correspondences between point cloud and BIM is difficult [5]. Indirect methods, such as those based on minimum Euclidian distance between the points and the BIM are used for estimating correspondences. The second step is visualizing deviations for an analysis of deviation patterns. Deviations are color-coded for such an analysis (Fig. 2c). Several coloring methods can be used, including binary maps, continuous coloring, or unsigned or signed maps [6,22]. Customized color-coding aids in the understanding of patterns in the data [7]. The third step is analyzing the deviation maps to identify the deviation patterns. In addition to the deviation maps, raw point cloud (Fig. 2a) and site photographs can also be used. For example, if deviation maps point to a missing component in the as-is BIM, photographs can be checked to verify whether there is such a component in reality. The deviation analysis is often carried out in 3D, which is a better method of information visualization compared to 2D information visualization, such as cross-sectional analysis [7,8].

The deviation analysis method has several benefits. First, it can comprehensively evaluate the quality of every surface as long as there is corresponding data. Second, this method can produce results at any time during the model generation process. The analysis can begin as early as the data collection step to identify data collection errors (e.g., missing data, calibration errors of the sensor). Third, analyzing the deviation patterns can identify sources of errors. Fig. 2c and detailed analysis in Section “Types of Errors and Deviation Patterns” indicate visual differences between the patterns caused by different types of errors (e.g., dislocations of walls or missing objects). Finally, as we will show in the section titled “Case study for validation,” the deviation analysis method is fast, straight-forward to apply, and can potentially be further optimized by adopting computerized methods for automatically analyzing deviation patterns. Such automation is out of the scope of this paper, but is briefly described in the Conclusion section. Additionally, the deviation analysis method can be applied together with other QA methods. Section 3 discusses some situations in which using different methods and information sources in conjunction with the deviation analysis method can help in pinpointing errors and eliminating alternative explanations of the observed patterns.

In order to formalize the deviation analysis procedure for QA, we performed several case studies. Specifically, we investigated the as-is BIMs of a post-office and a courthouse using the deviation analysis method to identify error sources and related deviation patterns, and to formalize the steps and software requirements of the deviation analysis method. To validate the performance of the deviation analysis method, this paper also presents the results of the deviation analysis of a plant facility and the office rooms in that plant facility. The authors compare the QA results based on the deviation analysis against those of the physical measurement analysis conducted by another research team in terms of the number of identified errors, magnitudes of errors, and time requirements.

2. Related work

The idea of comparing point cloud and the resulting model for quality assessment has been applied previously for building components and manufactured parts [12,13]. In manufacturing, color-coding the geometric differences between point clouds and models is applied for detecting defects [13]. The idea of comparing as-built parts to as-designed models to identify defects has been investigated [24]. Distance/deviation coloring has been used as a measure for identifying construction defects, such as incorrect location or orientation of components [5]. Despite the mentioned similarities between QA of manufactured parts and QA of as-is BIMs, overall goals of the procedures are different. QA of as-is BIMs aim at identifying quality issues of the model, which was generated from the point cloud. QA of manufactured parts uses the design model as the ground truth and aims at identifying the quality issues in manufactured/constructed objects. However, in both cases point cloud is assumed to depict the geometry of reality. Identifying data collection and data pre-processing errors are additional challenges in both applications.

Current methods of QA of as-is BIMs generated from point clouds include visual inspection, clash detection, and the physical measurement method. Among the current methods, only the physical measurement method is a formal QA procedure with quantitative measures of quality and a formalized measurement collection and analysis procedure specifically targeting as-is BIMs [9].

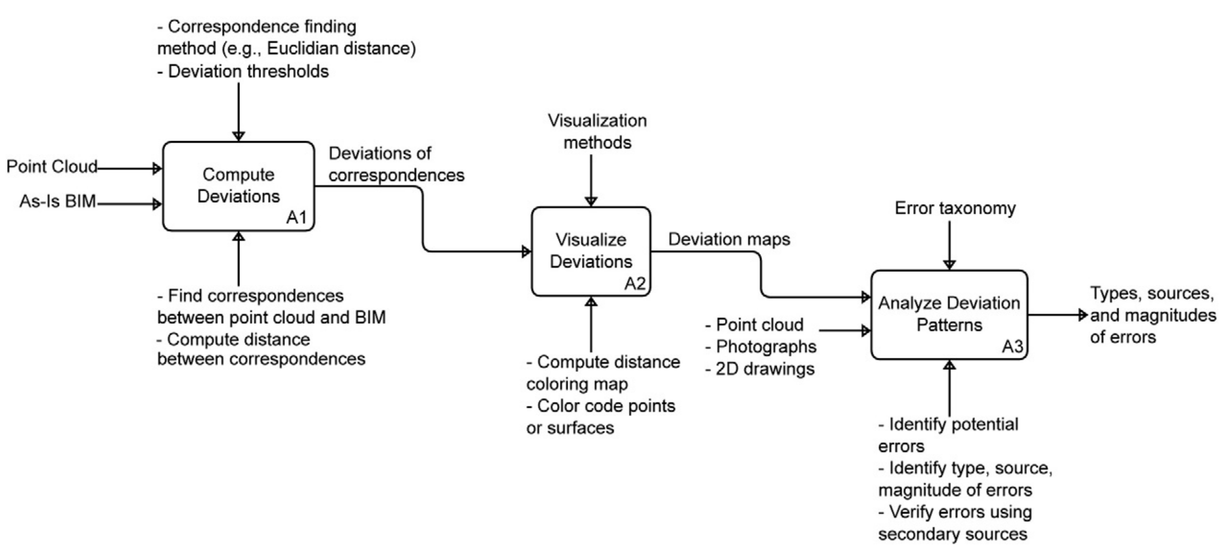

Fig. 1. The deviation analysis method is composed of three major steps. In the first step, deviations are computed by finding the correspondences between point cloud and BIM and then computing the distances between correspondences. The second step is visualization of deviation maps by color-coding deviations. In the final step, deviation maps are inspected; potential errors are identified and verified.

Please cite this article as: E.B. Anil, et al., Deviation analysis method for the assessment of the quality of the as-is Building Information Models generated from point cloud data, Automation in Construction (2013), http://dx.doi.org/10.1016/j.autcon.2013.06.003

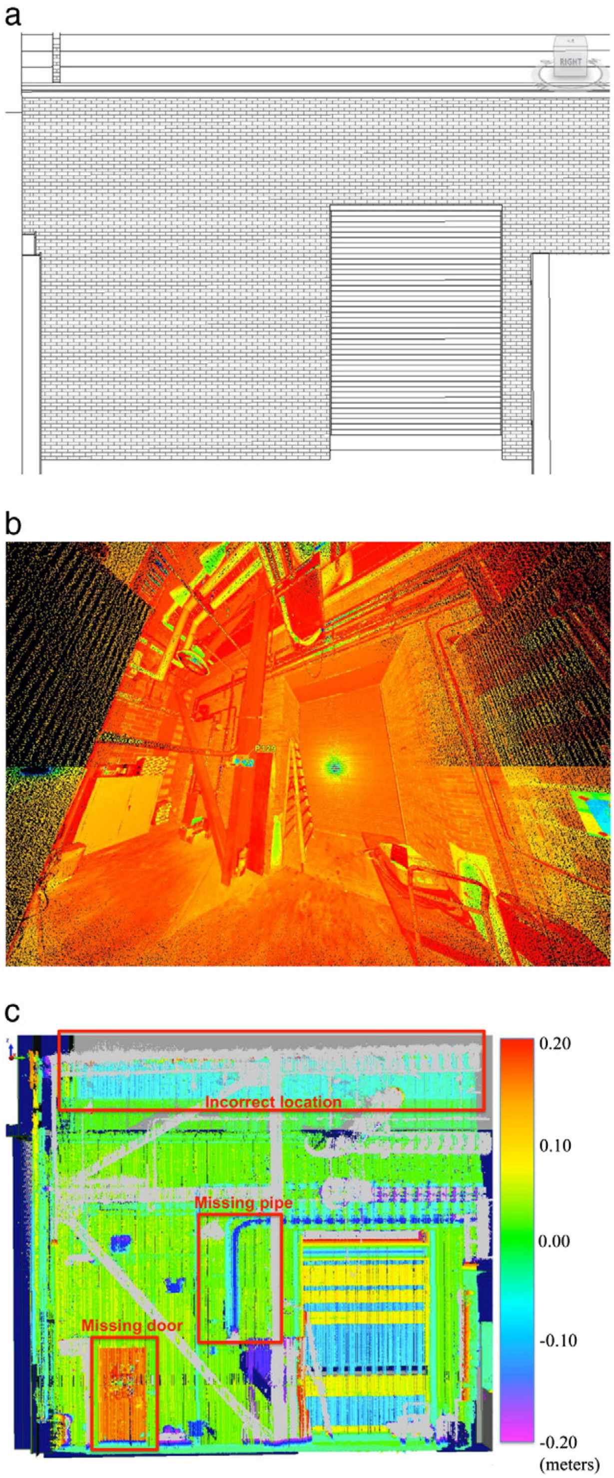

Fig. 2. Substantial deviations show that the pipes and the door are not modeled. Additionally, the top portion of the wall is at the wrong location. (a) BIM of the wall, (b) point cloud, (c) deviation map (0.2 m threshold).

The visual inspection method overlays the BIM and the point cloud, and enables visual comparison of the overlaid point cloud with the BIM for any anomalies. This method is straightforward and does not require much computation. However, visual inspection is subjective and cannot provide a quantitative assessment, such as how much a component is dislocated or deformed.

The clash detection method was originally used for checking whether a new design or modification interferes with components in the designs of collaborating engineers [10,11]. Clash detection compares the locations of the components and declares a clash if the compared components are closer than a given threshold or are occupying the same space. In a QA procedure, we are interested in the non-clashing components and the amounts of deviations rather than clashing components. Therefore, it becomes somewhat cumbersome to interpret the results of a clash detection analysis for QA.

The physical measurement method compares a set of virtual measurements on the BIM and the corresponding physical measurements taken on the actual building using, for example, tapes [9]. The compared values are then statistically analyzed to obtain a confidence value for achieving a given target quality. The physical measurement method is potentially the best method for verifying the accuracy of a specific measurement on the model, since virtual measurements are compared to real measurements. In addition, the physical measurement method can capture some types of errors that are difficult to identify by comparing the point cloud to BIM alone, such as scaling errors in the point cloud. On the other hand, the physical measurement method has some limitations. First, it does not provide full coverage of all possible measurements on the compared point cloud data and as-is BIM. Second, sources of errors cannot be directly identified, because limited number of measurements cannot typically provide sufficient information to distinguish different errors in the point clouds or as-is BIMs. Third, it is a time-consuming process, as it requires collecting large number of physical measurements. Finally, it is difficult to achieve highly accurate data with tape measures or other contact- based measuring methods, and it might be challenging to obtain some measurements due to accessibility limitations, such as ceiling heights.

3. The deviation analysis method

The deviation analysis method is composed of three main steps. In the first step, the deviations are computed. In the second step, the computed deviations are visualized using one or more visualization techniques. The goal of this step is to assist engineers in revealing deviation patterns that potentially indicate the types of errors through the combined usage of various visualization methods. In the final step, the deviation patterns are analyzed to identify the sources, types, and magnitudes of errors. The next subsections describe these steps in more detail.

3.1. Computing the deviations

This step computes the discrepancies between the model and point cloud or between individual scans. The main challenge is how to efficiently and precisely identify the correspondences between the compared objects. The distance between correspondences, when identified, can be used to describe the deviations. However, finding those correspondences can be challenging and almost impossible [5]. Therefore, we use indirect methods of associating data points with corresponding objects in the as-is BIM. A natural possibility is using the minimum Euclidian distance for associating data points with objects which are their nearest neighbors. Other methods have been proposed to find correspondences, such as projecting points onto surfaces, tracing rays to surfaces to find correspondences, or eliminating matches by certain metrics (e.g., normal direction) [14,5]. We selected minimum Euclidian distance data-object association due to its computational efficiency and effectiveness shown in previous studies [14,6].

One problem with the minimum Euclidian distance based approach is that some points, such as those corresponding to furniture, may have no correspondence in the model because they are not relevant to the problems of engineers and are purposefully not modeled. A threshold is used to limit distraction caused by irrelevant points. The selection of the threshold depends on several factors, including scene complexity, data noise level, types of error being analyzed, and accuracy requirements of the as-is model.

A threshold value equal to the maximum expected error in the scene can help in utilizing the color palette most effectively when visualizing deviations. For example, if the maximum deviation of the wall surface due to non-planarity is 5 cm, then the threshold value should be set to 5 cm. For smaller threshold values, the magnitude and exact location of the largest error cannot be identified because all regions with deviations larger than the threshold value are colored with the same color. For larger threshold values, the color spectrum will be spread over a larger deviation range, and the reduced range of colors allocated to the 0 to 5 cm range would make the visual identification of errors more difficult. However, due to the fact that a scene can contain errors of unknown magnitudes, it is best to run the analysis several times using different threshold values to maximize the identification of errors.

Smaller thresholds are more effective for visualizing detailed deviations, such as local geometrical errors. Larger thresholds are more effective for visualizing modeling errors influencing the global geometries of larger components in the facilities. A data set can be analyzed using a series of thresholds to identify different types of errors (Fig. 3). In our research, we have not identified a set of threshold values that would apply to all cases, as the threshold values can be project or even case specific. However, 5 cm, 20 cm, and 50 cm threshold values produced acceptable results in our case studies, including the comparison case study in Section 5.

3.2. Visualization of deviation maps

In the second step of the developed approach, the deviations are visualized to reveal potential deviation patterns. Deviation maps can be visualized in several ways. Major variations of deviation map visualization are:

- Signed vs. unsigned deviation maps: The directions of the surface normals are used to designate the signs of the deviations (i.e., positive or negative). The signs can either be ignored, resulting in unsigned maps, or different colorings can be used to designate the positive and negative deviations (Fig. 4a and c).
- Continuous vs. binary coloring: Deviation maps can be thresholded by assigning one color to deviations above the threshold and another color to deviations below the threshold (Fig. 4d). This coloring is called binary coloring. Such thresholding can facilitate the identification of deviations that exceed project-specific accuracy requirements. Continuous coloring assigns a color to every deviation value according to a gradient coloring scheme.
- Coloring points vs. coloring surfaces: Instead of coloring points, surfaces can be colored to designate their deviations from the points. Two ways of achieving this are: 1) coloring projections of the scan points on the surface (Fig. 4b); or 2) averaging deviations over a region of the surface, such as a triangle in the mesh representing the surface [6]. Generally, averaging over the surface is a less desirable method because it can obscure small, but important, deviation details.

3.3. Analysis of deviation patterns

The third and final step of the deviation analysis method is the analysis of the deviation maps to identify deviation patterns that point to sources and types of errors. This step is currently performed manually. It requires a professional to visually inspect the deviation maps, identify potential problems, and verify the source, type, and magnitude of the errors. This step may be aided with additional information sources, such as site photographs, field notes, or older blueprints of the building. Additionally, the professional may use other QA methods, such as the physical measurement method, in parallel with the deviation analysis method for a more effective QA.

By analyzing the deviation patterns, the QA professional tries to identify patterns of deviations, which point to potential errors. Different error types result in different deviation patterns. The types of errors and corresponding patterns of deviations are explained in the following section.

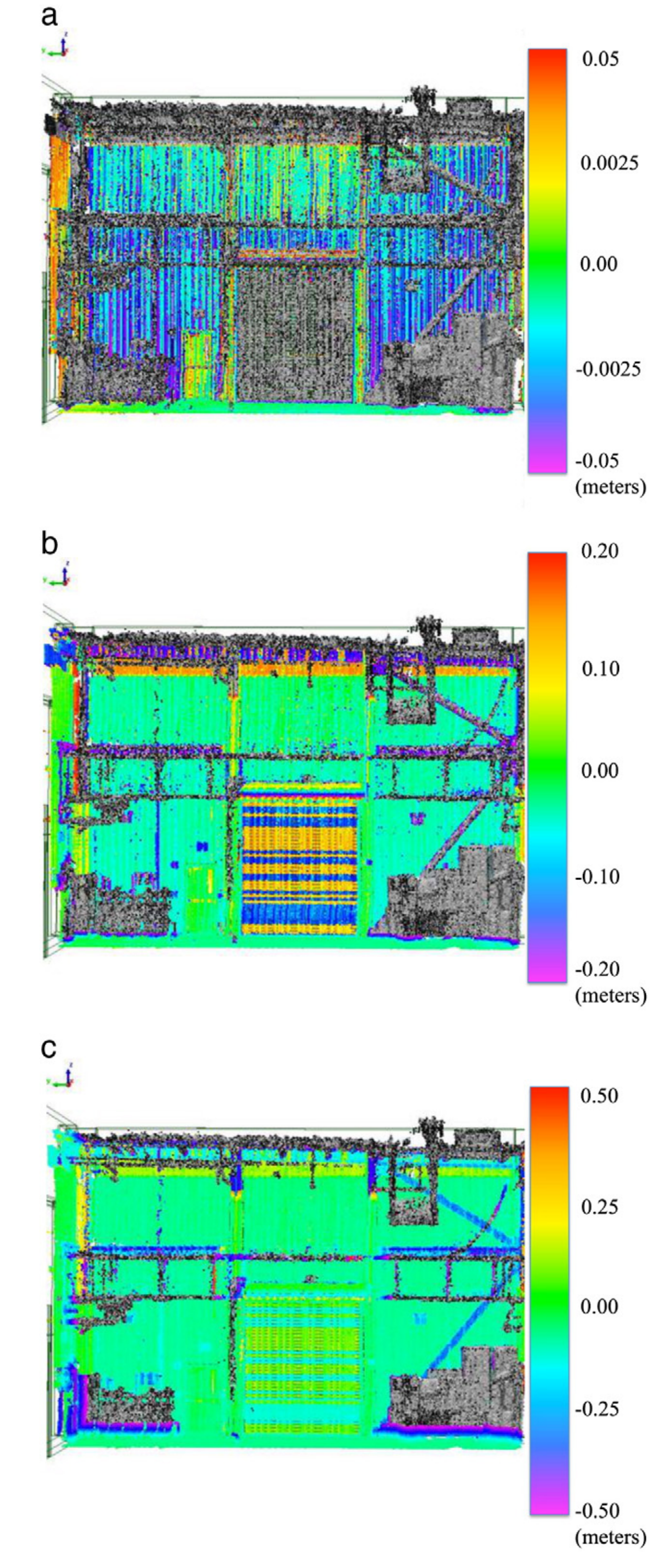

Fig. 3. Deviation maps with three different maximum deviation values. (a) At the 5 cm scale, an increasing deviation from the top right corner to the bottom left can be observed, which suggests an orientation error. Also, light blue lines are non-modeled pipes on the wall. (b) At the 20 cm scale, the diagonal deviation pattern and non-modeled pipes are difficult to see. (c) At the 50 cm scale, missing steel members that were not captured in the 20 cm scale are partially colored. Additional missing components beyond the 50 cm limit are also visible. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Please cite this article as: E.B. Anil, et al., Deviation analysis method for the assessment of the quality of the as-is Building Information Models generated from point cloud data, Automation in Construction (2013), http://dx.doi.org/10.1016/j.autcon.2013.06.003

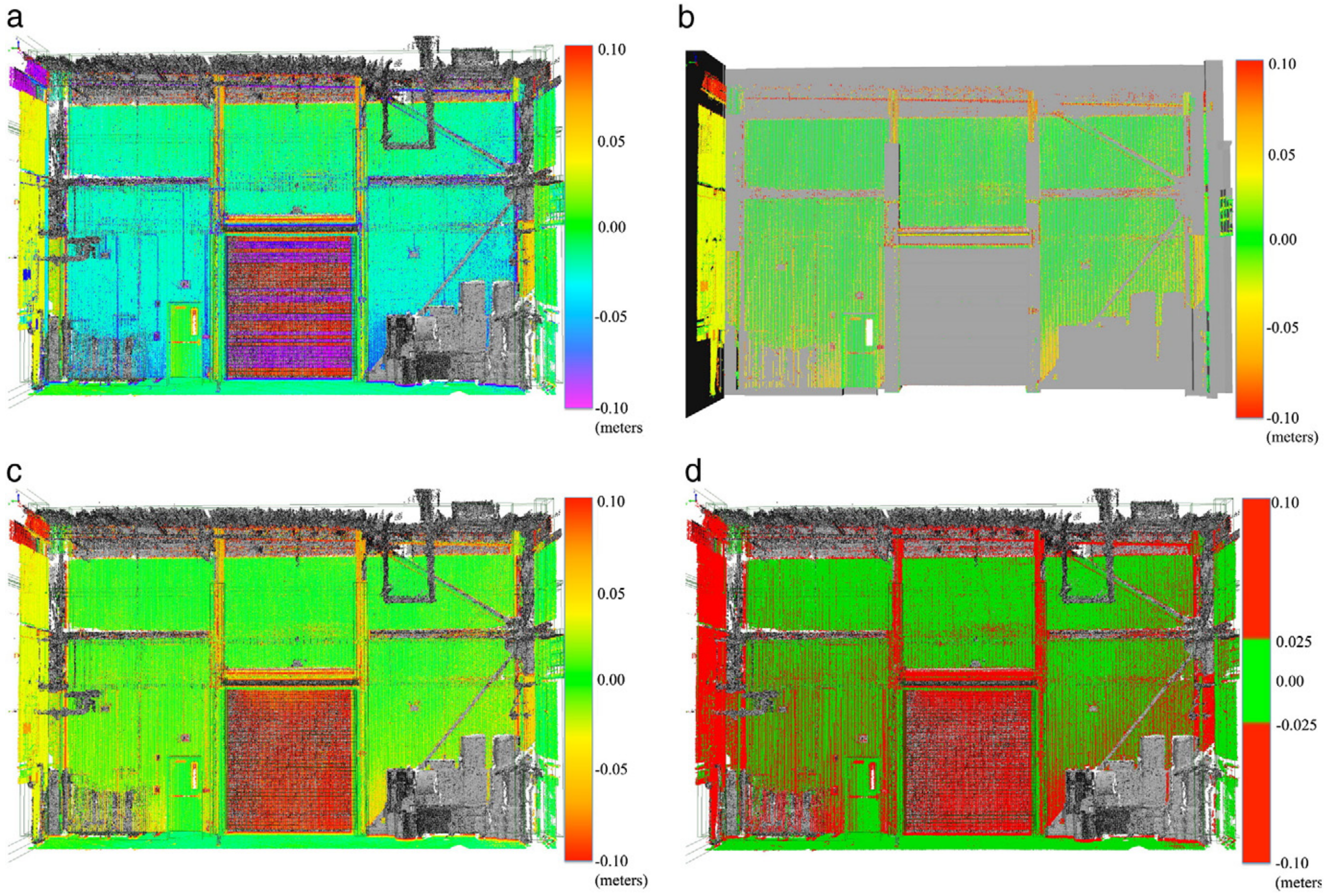

Fig. 4. Deviation maps can be colored in various ways for an effective analysis. There are four variations of coloring. Point coloring vs. surface coloring ((a) & (b)); signed vs. unsigned maps ((a) & (c)); continuous vs. binary coloring ((a) & (d)). (a) Signed continuous deviation map. Point coloring. (b) Unsigned continuous deviation map. Surface coloring. (c) Unsigned continuous deviation map. Point coloring. (d) Binary coloring. Unsigned. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

4. Types of errors and deviation patterns

At any stage of the BIM generation process, various types of errors may occur. The deviation patterns of errors occurring in each step are different. By analyzing these patterns, we can identify the sources of errors as well as the nature of these errors.

A typical BIM generation process has the following steps: 1) data collection; 2) data preprocessing, which involves the registration of the scans, removal of unnecessary clutter, and cleaning of data artifacts; and 3) modeling of objects, which involves determining the geometry of the relevant objects, establishing the relationships between objects (i.e., topology), and assigning categories and material properties to objects [1]. This section discusses the types of errors that can occur in each step, their corresponding deviation patterns, and strategies to identify the errors based on deviation patterns.

4.1. Data collection errors

Common error sources in the data collection step are calibration errors and data artifacts, such as mixed pixels and reflections [15,19]. The incidence angle and range of the laser beam hitting the surface influence the accuracies of points and occurrences of artifacts [19]. Data collection errors are also affected by the shape, color, and surface conditions (e.g., reflectivity) [15]. In addition, when observed from another scan location, the deviation pattern identified using the point cloud data from the initial scan location may look different.

4.1.1. Calibration errors

For various reasons, such as improper operation or unnoticed physical damage, laser scanners may become uncalibrated if they are not regularly maintained [16,17]. The patterns of calibration errors are sensor-specific, but they typically affect the perceived shape in the data set at a large scale. For example, generally multiple components, which were observed by the scanner, contain erroneous data and deviation patterns are continuous across components.

There are common properties of calibration errors that lead to patterns of deviations correlating with them. First, calibration errors generally lead to repetitive or cyclic patterns of deviations (Fig. 5a and b). For example, the trunnion axis error results in a sinusoidal pattern in the angular position as a function of the horizontal position [17]. Second, calibration errors are independent of the sensor location in the building and geometry of objects. The same deviation pattern will reoccur at the same relative location from the scanner in a different scan, regardless of the absolute location of the scanner.

These two observations lead to a practical rule for identifying calibration errors through analyzing the deviation patterns. The analysis of individual scans from different locations will show the same patterns of deviations, and the overlapping regions of individual scans having calibration errors should exhibit inconsistent deviation patterns. If the deviation patterns are due to a conflict in the geometry of the environment and the as-is BIM, the scans of the same geometry from different locations should show the same deviation patterns. This fact can help engineers in eliminating calibration errors as a potential cause of deviations.

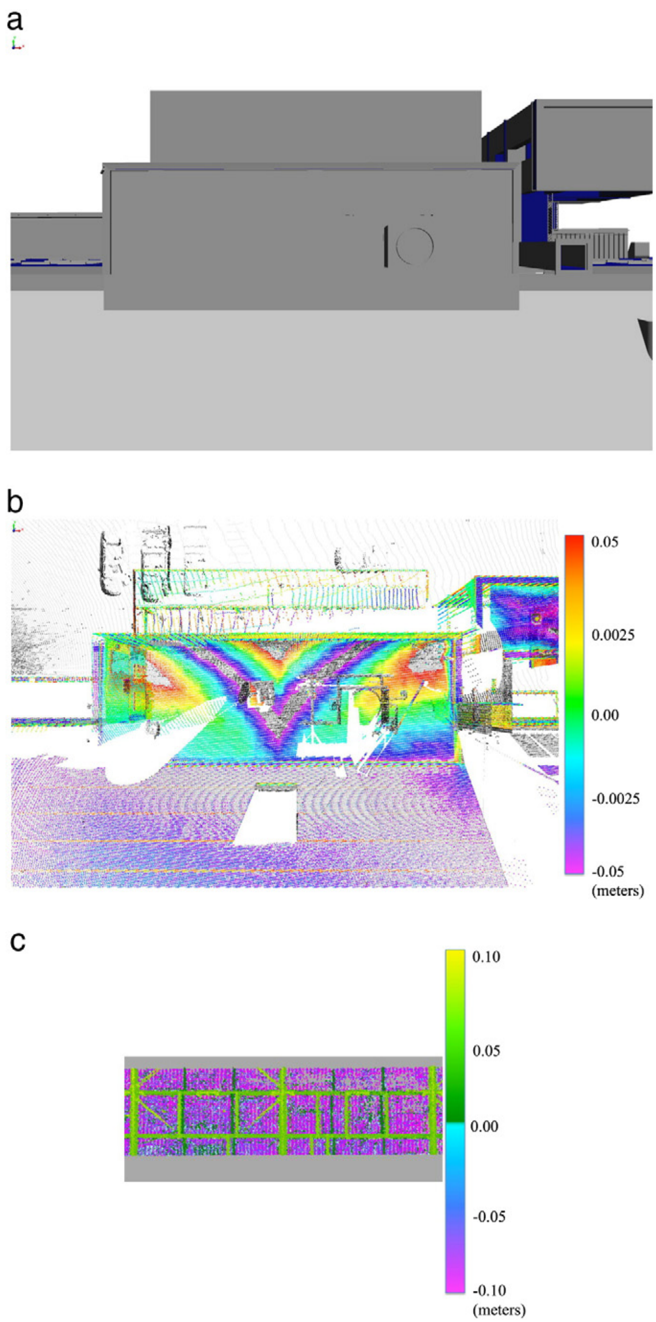

Fig. 5. Calibration errors result in repetitive circular or wavy patterns at scanner locations (b). The plan view of the BIM of the roof is shown in (a). Surface geometry and reflective properties can cause data artifacts (c). Shiny steel surfaces or corrugations on the surface are examples of such noise-prone surfaces, such as in the figure. Please also note that steel girders are missing in the model, and hence they are the reason for the green deviations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

4.1.2. Data artifacts

Behavior of the laser beams in particular environments can cause data artifacts, such as mixed pixels or reflections [15,18]. The data artifacts can be sensor and scan location specific. For example, the distribution of the points that are mixed-pixels is a function of the mixed pixel behavior of a scanner [19]. Similarly, reflective surface properties and surface geometry influence occurrence of data artifacts (Fig. 5c). Therefore, knowing the technology and the properties of the scanner can help in identifying and distinguishing data artifacts.

4.2. Registration errors

Registration errors can be identified by analyzing the deviations in the overlapping regions of individual scans or by analyzing the deviations of the registered point cloud from the model surfaces (Fig. 6). A feature of deviation maps of registration errors is that they are constrained to overlapping regions of scans.

4.3. Modeling errors

Deviation maps of modeling errors normally exhibit relevant patterns only on incorrectly modeled components. Exceptions, however, do exist. Sometimes, smaller components on relatively large components are modeled relative to larger components. For example, if window locations are fixed relative to walls that are modeled at incorrect locations, the windows will also be at incorrect locations even though they are located correctly relative to the walls. In these cases, the smaller components can also show similar patterns of deviations as the reference component (e.g., walls). In such cases, fixing the errors of the reference component fixes the errors of the associated components, such as components that contain them.

Modeling errors can be classified into four categories:

- **Missing components** show up as deviating regions in the shape of the unmodeled object and the deviation maps can also depict the surface shape of the missing component. The deviation pattern should be the same across different scans.
- **Incorrect geometry** of the components can be identified by inspecting local deviations of individual components. The deviation map shows the difference between the actual shape and the modeled shape of the component. If the deviations are relatively constant over the entire component, they are more likely "incorrect location" errors.
- **Incorrect location** errors can be identified by the constant deviation values over the component. The dislocation of the component is equal to the deviation value (Fig. 7).
- **Incorrect orientation** related deviations diminish at an axis, which coincides with the axis of the rotation. For example, Fig. 8 shows a wall that was assumed to be orthogonal with the other walls in the room, when in reality it was not. The deviations are small at the left side of the wall, which is the axis of rotation of the disorientation of this component.

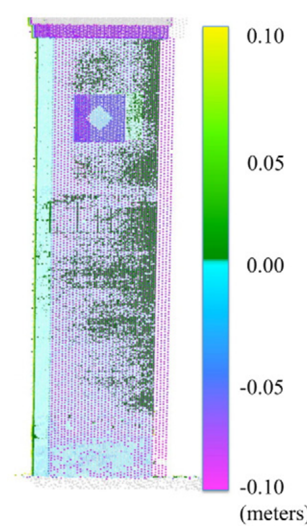

Fig. 6. A registration error revealed at the overlapping regions of three scans. Two scans are compared to a third scan, which is not shown in the figure. Notice the purple points which belong to one scan, and the light and dark green points belong to another. The second scan fits the model well, leading to low deviation values. The registration error of the first scan is about 5 cm, hence it has the purple color. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Please cite this article as: E.B. Anil, et al., Deviation analysis method for the assessment of the quality of the as-is Building Information Models generated from point cloud data, Automation in Construction (2013), http://dx.doi.org/10.1016/j.autcon.2013.06.003

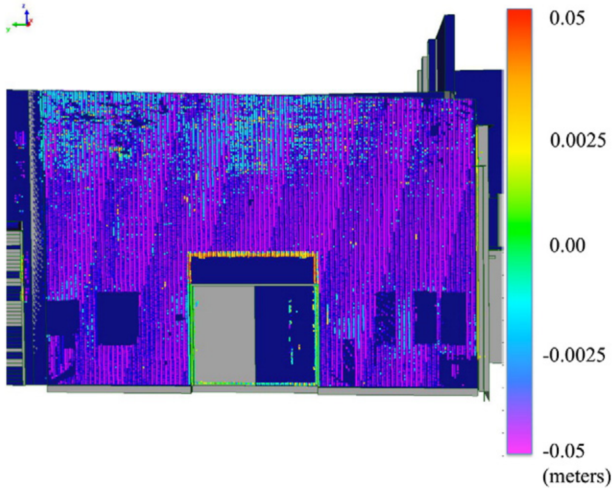

Fig. 7. A deviation map containing a component with incorrectly modeled geometry. Large deviations around the door in the middle point to the incorrect dimensions of the door. Measuring the dimensions of the door on the point cloud and on the model verifies the error. Additionally, the entire wall is modeled 5 cm away from its correct location, hence the reason for the predominant purple color. In this particular example, the wavy pattern on the wall is a visualization artifact. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Modeling errors can be introduced due to human error, or they can be artifacts of modeling process. Approximating surfaces using planes, assuming perfectly vertical surfaces, and assuming orthogonal object intersections (e.g., intersection of a wall to another wall or floor) are common modeling assumptions [1]. However, in reality, there are almost always irregularities and construction errors. Deviation maps reveal these irregularities from idealizations.

5. Case study for validation

This section discusses the application of the deviation analysis method on a professionally generated BIM of a plant facility and its office building attachment. The model is the same one used in [6]. The study described in this paper builds on the discussion provided in [6,22], but significantly extends it by providing more comprehensive and detailed discussions in terms of the analysis of the error

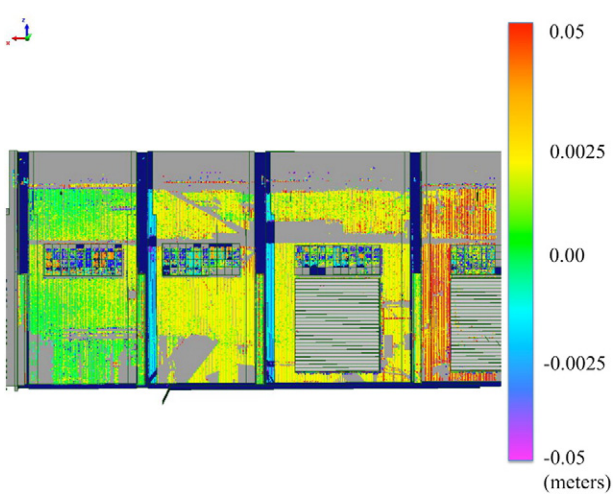

Fig. 8. Incorrect orientation of the entire wall is revealed through the gradient in the deviation map. Notice how the green color (0 cm deviation) on the left of the wall blends into yellow and then into orange (5 cm deviation) toward the right. This deviation pattern suggests that the wall is incorrectly oriented about its vertical axis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

types, breakdown of the errors, and the inclusion of additional portions of the facility located in the office building attachment.

The case study facility has two parts: the plant part, which contains the equipment, pipes, boilers, etc.; and the office building attachment, which contains workspaces for the employees. A professional service provider performed the laser scanning and the modeling of this facility. The as-is model created has been analyzed previously using the physical measurement method by the National Institute of Standards and Technology (NIST) [9]. Most of the measurements that were analyzed using the physical measurement method are room dimensions, such as distances between two walls and object dimensions, such as window size. The targeted accuracy of the model is 2.5 cm. The deviation analysis presented below uses the same accuracy level.

Following sub-sections present the results of the deviation analysis of the facility. Our specific implementation of the deviation analysis procedure for the case study is discussed. Then, the results of the deviation analysis method and the physical measurement method are compared in terms of the number of errors identified, agreement on common measurements, and time requirements.

5.1. Implementation of the deviation analysis procedure

The steps of the deviation analysis method, as discussed in Section 3, are generic and can be achieved differently using specific functionalities of various 3D data processing software systems. In this section, we describe our implementation of the deviation analysis method and discuss potential problems in this implementation and resolution strategies.

We performed the deviation analysis in Polyworks v9, a commercially available 3D reverse engineering environment [23]. The BIM of the facility is originally a Revit model [11]. We imported the BIM into Polyworks as a CAD model. Similarly, we imported the 3D point cloud through a format compatible with Polyworks.

Since the as-is model is derived from the point cloud, it should already be in the same coordinate system as the point cloud. In some cases, the model may be transformed purposefully, such as to match a certain coordinate system (e.g., actual geographic coordinates of the building). In practice, the transformation may not be recorded. In our case studies, whenever this occurred, we used automatic registration algorithms offered by the 3D modeling and analysis software to align point clouds with the model. When the automatic registration algorithms failed to achieve a satisfactory alignment (i.e., if it produced alignment errors larger than the accuracy tolerance of the model), we used a manual registration method that — matches manually selected distinctive features in the model and the point cloud (e.g., corners of walls). In these steps, any software capable of importing and aligning multiple 3D geometries into a common coordinate system can be utilized. For the computation of the deviations, we used the built-in deviation analysis functionality of Polyworks, which identifies corresponding objects of given data points through identifying the nearest neighbors in terms of the Euclidian distance.

In order to understand the capabilities and limitations of various 3D modeling and analysis environments for conducting deviation analysis, we conducted a survey of the relevant software systems [6] Several potential issues regarding the implementation of the deviation analysis method were found. These issues can be grouped into three major categories:

- Interoperability: Conversion of the BIM to a CAD model may result in the loss of critical information. We found that certain software systems divide component surfaces into triangles and assign surface normals in arbitrary directions. Since, the surface normal directions are used to compute the direction of the deviations, jumps in deviation maps can be observed at common edges of model surfaces with inconsistent surface normals. Confusion can be avoided if the front and back surfaces are colored distinctively during the inspection of the as-is model, or if the normal directions are flipped manually prior to the analysis. An example can be found in [6].

- Results visualization: We observed that some software systems, when colorizing surfaces rather than points, average deviation values over surfaces and assign the average value of these deviations to the entire surface. Such representation can indicate that an error exists on the surface, but ultimately hides the local details of the deviation patterns. Additionally, we observed that these software systems perform better when working on the exterior compared to inside the buildings [6]. In many cases, analysis results in the interiors of buildings were difficult to visualize [6].
- Computational capacity: We observed that some software systems have limited capacity for importing and processing 3D data and information. This may constitute a problem from the QA perspective depending on the size of the data and as-is model being analyzed. In order to work around this issue for large models, we divided the model into smaller regions called "interest regions." An interest region can be a component, such as a wall or a slab, or a collection of components, such as an entire floor, of the size that can be handled by the software tool and the computer hardware. This step can be performed before importing the model and the point cloud or the model and the point cloud can be analyzed piece-by-piece.

In our case studies, we used threshold values of 5, 20, and 50 cm for the deviation computation. The 5 cm threshold enabled the observation of small local deviations, such as non-planarity of walls. The intermediate 20 cm value allowed visualization of missing objects mounted on walls, such as circuit breaker boxes or wall-mounted pipes. The 50 cm threshold enabled observation of large modeling errors, such as missing steel columns that are not close to any model surface.

Following the computation of the deviations, the deviation maps were visualized using the visualization techniques described in Section 4. Observed patterns of deviations were then inspected to identify the error types, sources, and magnitudes. This step often required using the raw point cloud data and photographs to verify the errors.

5.1.1. Comparison method

In order to compare the deviation analysis method to the physical measurement method, we created equivalent relative measurements with the deviation analysis method. This was deemed necessary because the deviation analysis method gives absolute values of errors, whereas the physical measurement method gives the relative errors in the dimensions. We converted absolute deviations to their equivalent relative measurements by analyzing opposite components for deviations (i.e., opposing walls) and subtracting the absolute deviation values at the locations used for physical measurements. The surface normal directions of the opposite components were used to determine sign of the deviation values.

5.1.2. Analysis results

The steps of the deviation analysis method were repeated for all regions of interest in the facility. In total, 30 regions were analyzed. These regions were analyzed using 5 cm, 20 cm, and 50 cm threshold values. The errors within each region were grouped according to error type, and the number of errors in each group was counted. Table 1 presents the types of errors found in each room as well as the ranges of errors for the modeling errors. Rooms 1 and 2 are in the plant part of the facility. Rooms 3, 4, and 5 are office spaces.

In addition to the results in Table 1, there were many small orientation issues and incorrect geometry issues. However, those issues were within the tolerance (i.e., magnitudes less than 2.5 cm) of this project and thus not regarded as errors in the model.

We found that the two methods agreed with each other on whether the quality of the model is acceptable or not on the sampled measurements. Among these measurements, all measurements agreed within 1 cm except for one. For the disagreeing measurement, we found a significant error (45 mm), whereas the physical measurement method found 5 mm error. This may be because we measured the deviations at a slightly different location than the physical measurement location.

The results show that the deviation analysis method was able to identify 79 errors along with the types and sources of these errors. On the other hand, the physical measurement method was able to identify only 14 errors. This shows that the deviation analysis method alone identified 5.6 times more errors than the physical measurement method.

We verified all of the 79 errors by inspecting the raw scan data, visually comparing the scan to the model and referencing the site photographs and 2D drawings. The site photographs and 2D drawings were especially useful for verifying missing components in the model. Location, geometry, and orientation issues were verified primarily using the raw scans and by visually comparing the scan to the BIM.

5.1.3. Time comparison

We compared the two methods in terms of the approximate time needed for the analysis. First, we estimated the expected time to perform physical measurement analysis, and then we did the same for the deviation analysis.

The physical measurement method involves several steps, outlined in Table 2. Preparation time, travel time to the site, and site setup are highly variable, and depend primarily on the distance to the site. We assumed this step would require between 5 and 10 h at each end of the site visit, depending on the distance. Assuming a site has 30 rooms to measure, this time cost, averaged across all rooms, would be 10–20 min per room.

The next step is to plan the measurement locations. This can be a complex process and must be based on the actual geometry of the building, which may be different from the plans available in advance. For this particular case study, NIST researchers reported that the planning took approximately half a day.

During the NIST data collection, the time required to make the physical measurements was documented. Each location was measured up to 10 times, and the results were averaged. Although the site contained a number of different sizes and types of rooms, a typical large room was assessed to be approximately 6×10 m and

Table 1
Numbers and ranges of errors identified by the deviation analysis method and the physical measurement method in the plant facility model. The values in parenthesis depict the range of errors for the modeling errors. The model was divided into interest regions. The results are grouped into rooms.

Location	Data collection		Preprocessing	Modeling				Total by the deviation analysis method	Total by the physical measurement method
	Calibration	Artifacts	Registration	Missing components	Incorrect geometry	Incorrect location	Incorrect orientation		
Room 1	0	8	0	27	0	10 (5 cm–22.5 cm)	1 (5 cm)	46	4
Room 2	0	0	0	4	1 (5 cm)	3 (5 cm–3.1 m)	0	8	5
Room 3	0	0	0	3	0	6 (2.5 cm–23 cm)	0	10	2
Room 4	0	0	0	2	2	4 (3 cm–5 cm)	0	8	1
Room 5	0	0	0	1	3 (3 cm–5 cm)	3 (2.5 cm–5 cm)	0	7	2
Total	0	8	0	37	6	26	1	79	14

Please cite this article as: E.B. Anil, et al., Deviation analysis method for the assessment of the quality of the as-is Building Information Models generated from point cloud data, Automation in Construction (2013), http://dx.doi.org/10.1016/j.autcon.2013.06.003

Table 2
Estimated timing for the physical measurement method for the testbed.

Processing step	Estimated time (minutes)
Amortized travel and set up time	10–20
Measurement planning	8
Physical measurements	120
Virtual measurements	90–180
Write up	60
Total	288–398

contained 10–20 measurements. Assuming an average of 15 measurements per room and 8 repetitions per measurement, a total of 120 measurements must be taken in a typical room. The approximate time to conduct one measurement during the NIST data collection session was 1 min. Hence, we can expect that data collection using the physical measurement method requires approximately 2 h per room.

It is also necessary to extract corresponding virtual measurements in the as-is BIM for each physical measurement. This process involves navigating to the appropriate location, cross referencing the position of the measurement in the documentation for the physical measurements, and then using a virtual measurement tool to extract and record the measurements. Based on our experience, we estimate that this process would require 6 to 12 min per measurement (depending on whether the multiple repetitions at different locations are replicated in the virtual measurements). For a room with 15 measurements, the total time for the extraction of the virtual measurements would be 90 to 180 min.

The final step is to perform the comparison between the corresponding virtual and physical measurements, and write up the results. Based on our experiences in this case study, it would take approximately 1 h per room to document the results and write up an analysis report.

This estimate of the time requirements does not consider the fact that normally two or more people are required to travel to the site and obtain the measurements. The physical measurements would typically be obtained by teams of two people working together. Such teaming allows one person to operate the measuring device while another person verifies the location at the other end of the device (for measuring tapes) and documents the measurements. The addition of a second person would add an estimated 150–180 min per room (combining added travel time and physical measurement time), making the total time 445–575 min per room. Plus, we did not consider the monetary cost of traveling to the site or accommodations during the visit in this analysis to have a conservative estimate for the physical measurement method.

The deviation analysis method involves the steps listed in Table 3. We estimate the time for each step based on the time required for conducting the analysis in this case study. We estimated the analysis time based on 5 regions of interest per room. These interest regions correspond to parts of rooms, which include the locations of physical measurements. Point cloud data extraction and BIM data extraction required an average of 2 min each per region of interest. Importing data and model required about 1 min. We generated 3 deviation maps and took screen shots, which took about 5 min in total. The maps were then visually assessed (5 min) at a global scale and each abnormal deviation was analyzed in more detail (5 min per deviation). Finally, the results were written up in a report, which was the most time-consuming aspect of the analysis. If the write up of the errors in an interest region requires 12 min (on average) for 5 interest regions per room, it would take 60 min for the write up per room.

Using the physical measurement method, the total time to analyze a large room is estimated to be between 5 to 7 h. By comparison, the estimated time requirement for analyzing a room using the deviation analysis method is just over 3 h. This represents a time saving of 40 to 57% with respect to the physical measurement method. The main contributor to this time saving is the long time required to obtain numerous physical and virtual measurements. The analogous measurements in the deviation analysis approach are all computed automatically.

The results of the comparison of the deviation analysis method to the physical measurement method show that the deviation analysis method agrees with the physical measurement method in terms of the magnitude of the errors. In addition, the deviation analysis method found 5.6 times more errors and distinguished between the sources of the errors in 60% of the time required by the physical measurement method.

6. Conclusions

This paper introduced the deviation analysis method and error classification method for quality assessment of as-is Building Information Models obtained from laser scanned point clouds, and evaluated the method through a case study.

The deviation analysis method agrees with the conventional physical measurement method and is reliable in detecting the errors and error sources in the modeling process. Different deviation patterns correlate with different types of errors. Distinguishing between these different deviation patterns facilitates identifying sources of the errors. The deviation analysis method is an efficient and effective method of quality assessment of as-is BIMs. Results show that the deviation analysis method can achieve higher coverage and identify more errors while reducing the time requirements by 40%.

The deviation analysis method has a low barrier for adoption in practice, since some of the software packages used in the industry have readily available deviation analysis functionalities. Often it is necessary to make use of other information sources, such as photographs, raw scan data, field notes, and building plans to verify the errors identified through the deviation analysis method.

The deviation analysis method can be applied in conjunction with other methods (e.g., physical measurement method) for a more thorough analysis. For example, the deviation analysis method is not capable of identifying scaling issues in the scan data or the model. This can be worked around by using the physical measurement method in conjunction with the deviation analysis method and comparing actual measurements to virtual measurements.

We observed some limitations of the deviation analysis method. The deviation maps can be complex due to the errors and clutter. Numerous potential errors can be identified. This poses a limitation because the operator has to deal with all of potential errors and determine which deviations can be considered as errors according to the project specifications. Often, the observed deviations can be irrelevant to the QA task. For example, in some applications only the major building components are modeled. In these cases, all the other objects in the building (e.g., furniture, electrical equipment, items attached to walls, etc.) will cause clutter in the deviation maps. Additionally, multiple deviation patterns can overlap. In such cases, the operator has to manually process the deviation patterns to identify different types of errors.

Table 3
Estimated timing for the deviation analysis method.

Processing step	Estimated time (minutes)
Point cloud data extraction and export	10
BIM data extraction and export	10
Data import into analysis program	5
Deviation map generation and screen shots	25
Initial assessment	25
Detailed assessment	50
Write up	60
Total	185

Identifying sources of errors requires knowledge of the technology, sensor behavior, and unique deviation patterns caused by the environmental conditions or sensing parameters. Hence, for an effective application of the deviation analysis method, the operators need to be trained to a certain degree.

7. Future research

The deviation analysis method can be improved in several ways through further research. One immediate research direction is automating the classification of 3D deviation patterns for identifying sources and types of errors. Potentially, 3D pattern classification methods can be explored for this purpose [1]. In the domains of 3D computer vision and remote sensing, supervised or semi-supervised pattern classification methods have shown the potential of automatically classifying 3D data sets based on the variations of the data points across a surface (e.g., elevation changes captured in air-borne LiDAR data) [20,21]. Such a methodology could potentially be extended to classify the 3D deviations between laser scanner point clouds and as-is BIMs. In addition, research can be conducted on formalizing the usage of the deviation analysis method in conjunction with other methods, such as the physical measurement method.

Currently, only geometrical, location, and orientation issues can be identified. Methods of a more in-depth analysis of the deviation maps for identifying topological errors (i.e., relationships between components) can be researched. In addition, errors occurring in the plane of the component surfaces are hard to distinguish in the current implementation of the deviation analysis method. An example to these types of errors would be lateral location issues of components, such as the location of a door on a wall.

Potentially, the deviation analysis method can be applied at any stage of the model generation process for analyzing the intermediate results of modeling, providing opportunities for proactive quality control of the as-is BIM generation process. Similarly, it can be expected to achieve better results when the model is analyzed in the order it was generated, since errors in one step can lead to other errors in preceding steps. In other words, first data collection errors should be sought, followed by, identification of possible registration and alignment errors. Finally, the model should be analyzed for modeling errors. However, further research is required to validate these claims and formalize the relations between errors occurring in different stages of the model generation.

A deviation map involving multiple types of errors and complex geometries can be difficult to interpret through deviation analysis. The number of possible combinations of the deviations caused by multiple types of errors can be exponentially large. It is challenging even for experienced engineers to manually decompose such combined deviation maps for identifying all types of errors producing them. As a future research direction, decomposition of complex deviation maps containing several overlapping deviation maps will be studied.

References

[1] P. Tang, D. Huber, B. Akinci, Automatic reconstruction of as-built building information models from laser-scanned point clouds: a review of related techniques, Automation in Construction 19 (7) (2010) 829–843.
[2] L. Boudet, N. Paparoditis, F. Jung, G. Martinoty, M. Pierrot-Deseilligny, Supervised classification approach towards quality self-diagnosis of 3D building model using aeiral imagery, in: W. Förstner, R. Steffen (Eds.), Symposium of ISPRS Commission III Photogrammetric Computer Vision PCV '06, XXXVI, ISPRS, 2006.
[3] I. Sargent, J. Harding, M. Freeman, Data Quality in 3D: Gauging Quality Measures from Users' Perspectives, International Symposium on Spatial Quality, 2007.
[4] General Services Administration (GSA), BIM Guide for 3D Imaging, General Services Administration, 2009.
[5] F. Bosche, C.T. Haas, B. Akinci, Atuomated recognition of 3D CAD objects in site laser scans for project 3D status visualization and performance control, Journal of Computing in Civil Engineering 23 (6) (2009) 311–318.
[6] E.B. Anil, P. Tang, B. Akinci, D. Huber, Assessment of quality of as-is building information models generated from point clouds using deviation analysis, Society for Imaging Science and Technology (IS&T) and SPIE Electronics Imaging Science and Technology, SPIE, San Jose, California, USA, 2011.
[7] G. Atasoy, P. Tang, J. Zhang, B. Akinci, Visualizing laser scanner data for bridge inspection, 12th International Symposium on Automation and Robotics in Construction (ISARC), 2010.
[8] R. Fabio, From point cloud to surface: the modeling and visualization problem, International Workshop on Visualization and Animation of Reality-based 3D Models, 2003, pp. 24–28.
[9] G. Cheok, M. Franaszek, Phase III: Evaluation of an Acceptance Sampling Method for 2D/3D Buildling Plans, National Institute of Standards and Technology (NIST). Building and Fire Research Laboratory, Gaithersburg, MD, 2009.
[10] Autodesk. (2012). Autodesk Navisworks Products. Retrieved 2012, from http://usa.autodesk.com/navisworks/.
[11] Autodesk. (2012). Autodesk Revit Products. Retrieved 2012, from Revit: http://usa.autodesk.com/revit/architectural-design-software/.
[12] K. Yue, D. Huber, B. Akinci, R. Lipman, R. Krishnamurti, The ASDMCon project: the challenge of detecting defects on construction sites, International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT'06), 2006, pp. 1048–1055.
[13] Y. Li, P. Gu, Free-form surface inspection techniques: state of the art review, Computer-Aided Design 36 (13) (2004) 1395–1417.
[14] S. Rusinkiewicz, M. Levoy, Efficient variants of the ICP algorithm, Third International Conference on 3-D Digital Imaging and Modeling, 2001, pp. 145–152.
[15] M. Hebert, E. Krotkov, 3D measurements from imaging laser radars: how good they are? Intelligent Robots and Systems '91, Intelligence for Mechanical Systems Proceedings (IROS), 1991.
[16] T.P. Kersten, K. Mechelke, M. Lindstaedt, H. Sternberg, Methods for geometric accuracy investigations of terrestrial laser scanning systems, PFG — Photogrammetrie Fernerkundug Geoinformation, 2009. 301–314.
[17] H. Ingensand, A. Ryf, T. Schulz, Performances and experiences in terrestrial laser scanning, in: Kahmen Grun (Ed.), Optical 3D Measurement Techniques, IV, 2003, pp. 236–243.
[18] M. Adams, P. Probert, The interpretation of phase and intensity data from AMCW light detection sensors for reliable ranging, International Journal of Robotics Research 15 (5) (1996) 441–458.
[19] P. Tang, B. Akinci, D. Huber, Quantification of edge loss of laser scanned data at spatial discontinuities, Automation in Construction 18 (8) (2009) 1070–1083.
[20] A.P. Charaniya, R. Manduchi, S.K. Lodha, Supervised parametric classification of aerial LiDAR data, Computer Vision and Pattern Recognition Workshop, 2004, 2004, http://dx.doi.org/10.1109/CVPR.2004.172, (CVPRW'04. Conference on).
[21] E.H. Lim, D. Suter, 3D terrestrial LIDAR classifications with super-voxels and multi-scale conditional random fields, Computer-Aided Design 41 (10) (2009) 701–710, http://dx.doi.org/10.1016/j.cad.2009.02.010.
[22] P. Tang, E.B. Anil, B. Akinci, D. Huber, Efficient Quality Assessment of As-is Building Information Models and 3D Laser-scanned Data, ASCE, ASCE Computing Workshop, 2011.
[23] Polyworks. Innovmetric Software Inc. www.innovmetric.com.
[24] T.A. Clarke, S. Robson, J. Chen, Comparison of three methods for the 3D measurement of turbine blades, Proc. SPIE 2101, Measurement Technology and Intelligent Instruments, 1993.