Automated Retrieval of 3D CAD Model Objects in Construction Range Images

F. Bosche [a,*] C.T. Haas [a]

[a] *Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, ON, N2L 3G1, Canada*

Abstract

Automated and robust retrieval of three-dimensional (3D) Computer-Aided Design (CAD) objects from laser scanned data would have many potentially valuable applications in construction engineering and management. For example, it would enable automated progress assessment for effortless productivity tracking, automated 3D image database searching for forensic and legal analysis, and real-time local modeling for automated equipment control and safety. After reviewing and analyzing previous research in the field of automated object recognition, this paper presents a new approach for robust automated recognition/retrieval of 3D CAD objects in range point clouds in the Architectural/Engineering/Construction & Facility Management (AEC-FM) context. This approach is validated in laboratory experiments. A first experiment demonstrates that this new approach can efficiently and robustly automatically retrieve 3D CAD model objects in construction laser scanned data. A second experiment demonstrates how this approach can be used for efficiently assessing construction progress. The results presented here are preliminary but conclusive for proof of concept. More extensive field experiments in this and other application areas will follow to characterize performance trade-offs in practice.

Key words: Laser scanner, Range point cloud, Computer aided design, Data referencing, Automated object recognition.
PACS:

* Corresponding Author: tel: +1 (519) 888-4567 ext. 33872, fax: +1 (519) 888-4300
Email addresses: `fbosche@engmail.uwaterloo.ca` (F. Bosche), `chaas@civmail.uwaterloo.ca` (C.T. Haas).
URLs: `http://www.eng.uwaterloo.ca/~fbosche/` (F. Bosche), `http://www.civil.uwaterloo.ca/chaas/` (C.T. Haas).

1 Introduction

The Architectural/Engineering/Construction - Facility Management (AEC-FM) industry constantly needs to assess project performance with as much precision as possible and as fast as possible. Performance is tracked using metrics that meaningfully and efficiently estimate it. For instance, construction progress and productivity tracking requires assessing progress in terms of quantities and elements put in place, tests conducted, etc. Construction quality assessment requires, among other aspects, assessing the three-dimensional (3D) similarity between as-built and as-planned 3D objects. Similarly, construction dispute resolution and forensic analysis may in the future require exhaustive searches of range point cloud databases to acquire incontrovertible evidence of facts on the ground. In all these examples quantities and structural elements can be described in design documents and tracked as 3D shapes. Tracking quantities, elements, and quality automatically with the aid of automated recognition/retrieval of 3D Computer-Aided Design (CAD) objects from construction range point clouds would thus be beneficial and is possible with the method described in this paper. For brevity, the authors focus primarily in this paper on the application to construction progress tracking.

Traditional practice for construction progress assessment relies on intensive manual data collection and processing. This is labor intensive, expensive, and generally results in partial and sometimes erroneous information. As a result, it is difficult to make appropriate and timely management decisions ([1]; [2]; [3]).

The recent and rapid development of laser scanners, also referred to as LAser Detection And Ranging (LADAR), allows precise and comprehensive acquisition of range point clouds. Laser range point clouds are often referred to as *range images* or $2\frac{1}{2}D$ *data* because they contains 3D information about visible surfaces only. In the specific context of construction progress assessment, laser scanners can be used to acquire range point clouds from an asset in construction at any time. Acquired range point clouds can be analyzed to identify the presence of 3D project objects, so that the quantity of work that has been performed up to that specific time can be estimated. The advantage of using laser scanning data for assessing construction progress is that it directly identifies in-place quantities. It is thus potentially more robust than and at least complimentary to other approaches that indirectly calculate work progress — e.g. by recording in real-time the location of construction resources for inferring production quantities ([3]; [4]). However, industry managers could benefit from laser scanning technologies for effortless construction progress tracking only if they can be used to obtain reliable and high-value information, rapidly and, if possible, automatically [5].

A new approach is presented in this paper that allows robust automated retrieval of 3D CAD objects from range images. Sections 2 and 3 review existing approaches for automated object recognition in sensed data, and analyze their applicability and expected efficiency and robustness in the investigated context. This analysis leads to the formulation of a new approach described in Section 4. Section 5 presents two laboratory experiments, conducted in the Centre for Pavement And Transportation Technologies (CPATT) at the University of Waterloo, that validate this new approach and demonstrate its applicability to automated construction progress tracking. Section 6 then discusses the impact of measurement uncertainties on the proposed approach and suggests methods to take them into account. Finally, Section 7 discusses the estimations of the different parameters used in the proposed approach and how these could be automated.

2 Automated Recognition of 3D Objects in Range Images

2.1 Common Approaches to the Object Recognition Problem

The automated recognition of objects in sensed data, also referred to as *object recognition* is not a new problem and previous research in this field has been extensive, especially for application in robotics. In [6], Arman & Aggarwal propose a definition of the *object recognition problem* as *"locating a desired object in a scene and determining its exact location and orientation"*. In this definition, the combination of the location and orientation of an object is also generally referred to as its *pose*. Systems performing object recognition must have some *a priori* knowledge of the search object(s) (e.g. shape, color, temperature). This *a priori* knowledge is generally contained in an object *model*. As a result, such systems are generally referred to as *model-based object recognition systems* and they generally follow the following process:

(1) A data representation is chosen to meaningfully describe the object model,
(2) Features are extracted from the object model described using the chosen data representation,
(3) Features are extracted from the sensed data described using the same data representation,
(4) Object features are matched to sensed data features in order to infer the recognition of the object,
(5) The poses of recognized objects are estimated.

The choice of the data *representation* determines the recognition strategy and thus has a significant impact on the efficiency and robustness of the recognition system. An adequate representation is: unambiguous, unique, not sensitive, and convenient to use [6]. A review of most common strategies for object recognition can be found in [6] and some examples of systems for automated recognition of 3D objects in range images can be found in [7], [8] and [9].

The main challenge faced by typical model-based object recognition systems is that they are based on the extraction of features from both the search objects' models and the sensed data. These systems can be referred to as *feature-based* model-based object recognition systems. The level of difficulty in the extraction of features increases with the "complexity" of the search context, and this "complexity" is related to the following factors:

Unknown pose of each object. Object recognition systems generally assume that the pose of each object is *a priori* unknown. This assumption is genuine in most general search cases when the only *a priori* knowledge is the set of search object models.

Unknown relative pose of search objects . Similarly, object recognition systems generally assume that the relative pose of two search objects is *a priori* unknown. This assumption is also genuine in most general search cases.

Number of search objects. Object recognition systems generally search for objects one at a time in the scanned data. As a result, their computational complexity is proportional to the total number of search objects.

Occluded and cluttered scenes. Most object recognition systems genuinely assume that scanned scenes may include data about any object, searched or not searched. This however makes efficient and robust automated feature extraction very difficult.

2.1.1 Spin-Image Approach

In [8], Johnson & Hebert present another *model-based* approach that is based on 2D data representations called *spin images*. This approach is interesting because it is not *feature-based* as spin-images of the entire range data are directly compared to the spin-images of the search objects' models. In this approach, recognition is achieved as follows:

(1) All search objects are represented as polygonal surface meshes,
(2) A spin image is calculated for each vertex of the mesh representation of each object,
(3) The scanned data is represented as a polygonal surface mesh,
(4) Random vertices are identified in the sensed data mesh and spin images are calculated for each of them,
(5) Each spin image obtained from the sensed data is matched with all spin images of the search objects,
(6) For each object, if several spin-image correspondences are found, this

object is considered recognized and its pose is estimated.

The main advantage of this approach is that it is not feature-based and thus does not suffer from the limitations of feature extraction algorithms. Additionally, this approach appears fairly efficient with occluded and cluttered scenes (in the experiments, objects up to 68% occluded were systematically retrieved). Nonetheless, this method also presents some limitations:

- The scanned scene is approximated with a polygon tessellation, which results in a loss of information originally contained in the range image.
- Not all vertices in the scanned data mesh are investigated (20 to 50%), meaning that small or very occluded objects are likely to be missed. This could be avoided by investigating all vertices in the scanned data mesh, but would result in a computational complexity proportional to the number of vertices in the scanned scene mesh, which can be very high.
- Computational complexity is proportional to the number of objects and the number of spin images for each object. In [8], Johnson & Hebert nonetheless show that, for each object, Principal Component Analysis can be used to at least reduce the search domain constituted by all its spin images.
- The pose of objects presenting symmetries cannot be ensured since the spin image of a symmetrical object in one pose is exactly the same as the one in its symmetrical pose.
- Although this method is reasonably robust with object occlusions, it could be argued that it would be interesting to be able to retrieve objects more than 70% occluded. Recognition of more highly occluded objects could probably be achieved here if all vertices in the scanned data were investigated, but, as explained above, this would result in higher computational complexity.
- Finally, this approach recognizes objects by matching 2D object characteristics (spin images). This implies that some information contained in the $2\frac{1}{2}$D range data is not only lost while performing the data tessellation, but also while calculating each spin image.

2.2 Application to the Investigated Problem

The investigated problem of automatically retrieving all construction project objects present in a construction site range image has the following characteristics:

- The number of objects that should be searched in the scan is the number of 3D construction objects constituting the project model, which can be very large. Additionally, the shape of search objects can be very complex.
- Construction site scenes are generally very occluded and cluttered. Also,

many project elements might be scanned in partial construction status (e.g. partially built walls and columns).

As a result, if feature-based object recognition approaches were to be used in this specific context, they would generally be too computationally complex and would result in limited recognition results as construction scenes are too complex for efficient and robust 3D feature extraction. This feature extraction complexity is also increased by the fact that it is not possible to recognize all the features of a given model in one range point cloud due to occlusions and the fact that range information is only $2\frac{1}{2}$D. Previous works in civil engineering investigating the use of feature-based object recognition approaches to this problem acknowledge these limitations ([10], [11],[12]).

Similarly, if the spin-image approach was used, it would generally be too computationally complex due to the number of search objects, the number of spin images for each object, and the number of scanned points. Also, it could suffer from the highly cluttered and occluded characteristic of construction scenes. Nonetheless, the spin-image approach would likely be more robust than feature-based object recognition approaches. The spin-image approach is thus further investigated and feature-based approaches are discarded for the remaining of this analysis.

3 The Context: New AEC-FM Technologies

3.1 Project 3D CAD Models

In recent decades, the AEC-FM industry has been experiencing a rapid increase in the use of project 3D/4D CAD models. Project 3D CAD engines allow for the development of exact and comprehensive project designs in the form of 3D models. Project 4D CAD models enhance project 3D CAD models with schedule information. Project 3D/4D models undeniably increase design quality, management and communication among stakeholders, and decrease the number and impact of changes occurring during the project life cycle [13]. Additionally, they are now used as the central components of more complex AEC-FM management models such as Building Information Models (BIM).

Project 3D/4D CAD models do not constitute a basic library, but a spatially organized library of the project 3D objects. The relative pose of each pair of 3D project objects is thus expected to be the same in the 3D CAD model as in reality once they are built. Consequently, by using project 3D CAD models in 3D object recognition systems, the recognition of one object would provide *a priori* information on where to search for all the other objects. Or, from another perspective, the entire project 3D CAD model could be searched simultaneously.

Project 3D CAD models present another interesting advantage, regarding occlusions. From a given project 3D view point, all occlusions to a project 3D object due to other project 3D objects are expected to occur similarly in reality and in the project 3D CAD model. Such information, if efficiently incorporated in 3D object recognition systems, could significantly improve their robustness, especially when dealing with potentially very occluded scenes such as construction sites.

3.2 (Geo-) Referencing

Along with 3D CAD engines, global positioning technologies (i.e. GPS for location and digital compasses for orientation) are being used more in the AEC-FM industry since their accuracy and precision have become acceptable. Regarding location estimation, while Differential GPS (DGPS) can achieve sub-feet accuracy, Relative Kinematic Positioning (RKP) GPS technology can improve location estimation accuracy up to a couple of inches. Further, GPS technologies remain a major area of research and it is not unrealistic to imagine sub-inch accuracy systems in the near future. Similar conclusions can be made for orientation estimation systems such as digital compasses that typically achieve accuracies of half a degree.

Both field data and 3D CAD models can be geo-referenced. Therefore, field data can be typically registered into the coordinate frame of the model. In the AEC-FM industry, global positioning technologies are thus already used to enable management to track position of any type of important resource in real-time on project sites for applications as diverse as productivity tracking, lay-down yard management or safety.

In the problem investigated here, using (geo-) referencing technologies would simplify the search of the project 3D CAD model in the scanned data as the position of each search object in the scanned data would be *a priori* known (at least estimated). The authors acknowledge the limited accuracies of current (geo-) referencing technologies. Nonetheless, these technologies can be used to at least provide good pose estimations, and Section 6 discusses how, in the investigated problem, the pose of 3D CAD model in the scanned data could be optimized once a good estimation is obtained.

3.3 Impact on the Investigated Problem and Solution

The technologies above — that are already being used on construction projects but in other applications — could be leveraged in the investigated problem. Used with the spin-image approach, it seems that its major limitation — its computational complexity due to the number of search objects and the number of vertices in the scanned scene mesh — could be significantly reduced. Indeed, the project model could be searched all at once, and for each scanned scene mesh vertex, the project model mesh vertex for which the spin-image matching should provide the best result can be known *a priori*. Finally, thanks to the 3D referencing, the limitation of this method with symmetrical objects is also overcome.

However, it must be noted that the spin-image approach provides results regarding the overall recognition of each search object, but it is not suited to provide detailed recognition results of parts of the search object. The recognition of each individual project object is important in the investigated problem. Therefore, each object must thus be searched individually, not the entire project 3D model simultaneously, and the complexity of the spin-image approach remains proportional to the number of search objects. Additionally, this method is based on the approximation of the sensed data by polygon tessellation, which results in a loss of information contained in the original data. Finally, the data matching is based on a 2D data representation (spin-image). The representation of the $2\frac{1}{2}$D range data using spin-images thus further reduces the amount of information available for the matching process. As a result, the spin image approach cannot achieve optimum object recognition as it considers only part of the information contained in the acquired range data.

Despite these limitations, the authors acknowledge the apparent robustness of the spin-image -based 3D object recognition approach. A new model-based 3D object recognition approach is nonetheless presented here. This approach uses the sensed data (scanned point cloud) in its raw format, it is not feature-based, and its complexity is not proportional to the number of search objects as the entire project model is searched simultaneously. As a result, this approach is expected to be both efficient and robust for the automated recognition of project 3D CAD objects in construction range images.

4 New Approach

The proposed new approach is based on the idea that, since the performance of any approach for automated recognition of 3D object in range images is constrained by the sensed data, the best recognition approach can only be obtained if the sensed data is used in its natural representation, here the range point cloud. As a result, the authors propose an approach that uses the range point cloud as the common 3D object data representation. This implies that the project 3D CAD model must be represented as an equivalent range point cloud. To do this, (geo)-referencing information is used to reference the project 3D CAD model in the laser scanner's spherical coordinate frame. Then, for each as-built range point, a corresponding range point is calculated using the project 3D CAD model as a virtual world. This virtual world can also be referred to as the *expected world* or *as-planned world* and the point cloud resulting from the virtual scan conducted in this virtual world can be referred to as the *as-planned point cloud* (by comparison to the real *as-built point cloud*). As-built point features include at least three spatial coordinates, that are sometimes enhanced with reflectivity and color information. Similarly, as-planned point features include three spatial coordinates as well as any additional information that can be extracted from the project 3D CAD model when calculating the as-planned point cloud. These features may include object color and object reflectivity. But more importantly, one additional as-planned point feature that can systematically be extracted from the project 3D CAD model is the *"ID/name"* of the object from which each as-planned range point is obtained.

The challenge of this approach consequently lies on the calculation of the *as-planned point cloud*. A method for this calculation is presented in Section 4.2. Then, Section 4.3 presents the two metrics that are used for automatically comparing as-built and as-planned point clouds in order to infer the retrieval/recognition of all project 3D model objects.

4.1 Project 3D CAD Model Format

Full access to the information contained in the project 3D CAD model is necessary in order to practically calculate the as-planned point cloud. However, project 3D/4D models generally present the project 3D as-planned data in a proprietary 3D CAD engine format (e.g. DXF, DWG, DGN, etc.). Since these proprietary formats are protected, the as-planned point cloud calculation requires the project 3D CAD model be converted into an open-source 3D format. This open-source format must be chosen so that the conversion results in as little loss of 3D information as possible.

In [14] the authors identify one good candidate format that meets this information preservation requirement : the STereoLithography (STL) format. Detailed information about this format that approximates volume envelopes by tessellations of triangles can be found in [15]. It might be argued that, if access to proprietary formats is granted, this conversion would not be necessary anymore. However, it will be shown in the next section that polygon tessellation-based formats such as the STL format present an additional advantage over native CAD engine formats with respect to the proposed approach.

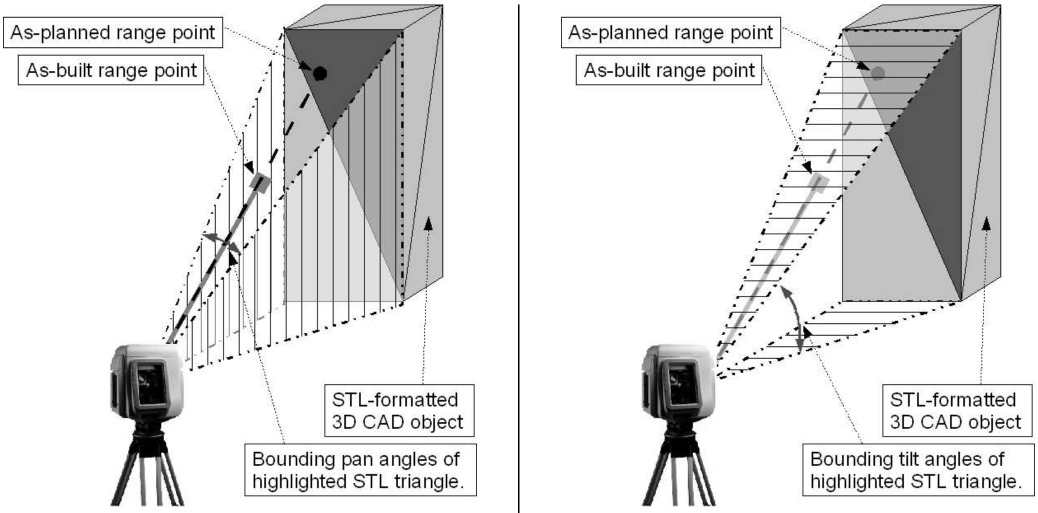
4.2 Calculation of the As-planned Point Cloud

The as-planned range point cloud can now be calculated as follows:

(1) Using the (geo-) reference information, the STL-formatted project 3D CAD model is referenced in the laser scanner's spherical frame. In this coordinate frame, the coordinates of each STL triangle composing the envelop of each object of the project model can be expressed using spherical coordinates (instead of natural Cartesian coordinates).
(2) For each as-built range point, the corresponding as-planned range point is assigned the same pan and tilt angles. Then, its range is calculated by finding the closest STL triangle intersected by the "ray" traced in the direction defined by these pan and angle angles.

The identification of the closest STL triangle intersected by a ray is a constrained version of the calculation of the projection of a point on a plane in a given direction. This problem is fairly straight-forward so that the solution won't be detailed here. Instead, the authors want to emphasize the fact that the combination of the project 3D CAD model being referenced in the laser scanner's spherical frame and the project 3D CAD model being converted into the STL format presents an opportunity for significant reduction in the computational complexity of the identification of the closest STL triangle intersected by a ray and thus of the calculation of each as-planned range point. Indeed, in this spherical frame, all the vertices of all the STL triangles are expressed with spherical coordinates: pan, tilt and range. As a result, the bounding pan and tilt values of each STL triangle can be identified. Then, as illustrated in Figure 1, it can be noted that the intersection of a ray defined by the two angles pan_0 and $tilt_0$ can only intersect a STL triangle whose bounding pan and tilt angles actually surround the pan_0 and $tilt_0$ values. This implies that the closest intersected STL triangle can be rapidly identified by analyzing only those STL triangles whose bounding pan and tilt angles surround pan_0 and $tilt_0$. Compared to the spin image approach, the complexity of this object recognition approach is thus not proportional to the number of search objects.

It must be emphasized that this complexity reduction is possible because it is fairly simple to calculate the bounding angles of a STL triangle and the intersection of a line with a STL triangle. If the project 3D CAD model was not expressed using a polygon tessellation-based format, but using a native CAD format — where each CAD object is represented as the intersection of primitive forms, these calculations would become much more complex.



(a) Bounding pan angles (b) Bounding tilt angles

Figure 1. Illustration of the selection of STL triangles based on their bounding pan and tilt angles for identifying the closest STL triangle intersected by a given “ray”.

4.3 The Range Point Matching And Object Recognition Metrics

Once the as-planned range point cloud has been calculated, it is possible to sort the as–planned range points by their object *“ID/name”* feature (the object from which each of them was obtained). This results in an as-planned range point cloud for each object constituting the project 3D model (note that each object for which no as-planned range point was obtained is simply assigned an empty point cloud). Then, for each object as-planned range point cloud, each as-planned point can be directly matched to its corresponding as-built point. This requires a *range point matching metric*. After matching each point of the object as-planned range point cloud, the recognition of the object can finally be inferred. This requires a second metric, the *object recognition metric* (or *object retrieval metric*).

4.3.1 Range Point Matching Metric

Each as-planned range point corresponds to exactly one as-built range point, and these two points have the same pan and tilt angles. Their matching can thus only be estimated based on the only remaining common feature, the range coordinate (although if additional common features exist, they should certainly be used). A range point matching metric can thus simply consider the difference in their ranges and compare it to a given threshold. For instance, an as-planned range point can be considered positively matched to its corresponding as-built point if the absolute difference in their ranges, $\Delta Range$, is lower than the distance threshold, $\Delta Range_{min}$.

In Section 7, the authors discuss a method to automatically define an adequate $\Delta Range_{min}$ threshold that takes into account context-specific factors. In the experiments presented in Section 5, a manually *a priori* estimated threshold is however used.

4.3.2 Object Recognition Metric

For each project object, once the matching of all as-planned range points with their corresponding as-built range points has been assessed, the recognition of the object can be inferred. For this, a straight-forward and commonly used object recognition/retrieval metric is used: the calculation of the object as-planned point cloud *retrieval rate*, $R_{\%}$, which is the ratio of the number of retrieved as-planned range points to the total number of as-planned range points. $R_{\%}$ can be compared to a threshold $R_{\%min}$ to infer the object recognition/retrieval. It is not however obvious what value $R_{\%min}$ should take. In fact, whatever the value of $R_{\%min}$, this metric, as is, will not be robust in the following two cases:

Object as-planned point cloud containing only a few points. For instance, if an object as-planned point cloud contains two points and if one point is recognized, then 50% of the as-planned point cloud is retrieved. Clearly, such a situation — that can occur when the object is far or very occluded, or when the range point cloud density is low — should not lead to the recognition of the object, despite the high point cloud retrieval rate.

Object occluded by non-CAD objects. This may result in objects having unreasonably low retrieval rates although many points are actually retrieved. For instance, in the case where 5% of an as-planned point cloud containing 2000 points is retrieved, the retrieval rate is very low, but there are still 100 retrieved points and it could be argued that the object should be considered retrieved.

The first situation can be handled by adding to the retrieval metric the condition that an object can only be considered for retrieval if its as-planned range point cloud contains a minimum number of points, defined by a threshold P_{nmin}. The second situation can be handled by adding to the retrieval metric the condition that, if the number of recognized as-planned points is higher than a given threshold R_{nmin}, this is sufficient to consider the object retrieved (no need to calculate the as-planned cloud retrieval rate).

Like for the point matching metric, the authors discuss in Section 7 methods to automatically estimate adequate P_{nmin}, R_{nmin} and $R_{\%min}$ threshold values by taking into consideration the context-specific factors such as: the scan point density and distance between the scanner and each search object. However, in the experiments presented in Section 5 these thresholds are manually *a priori* estimated.

This final CAD object as-planned point cloud retrieval metric is summarized in Figure 2. The pseudo-code of the overall proposed approach is presented in Figure 3.

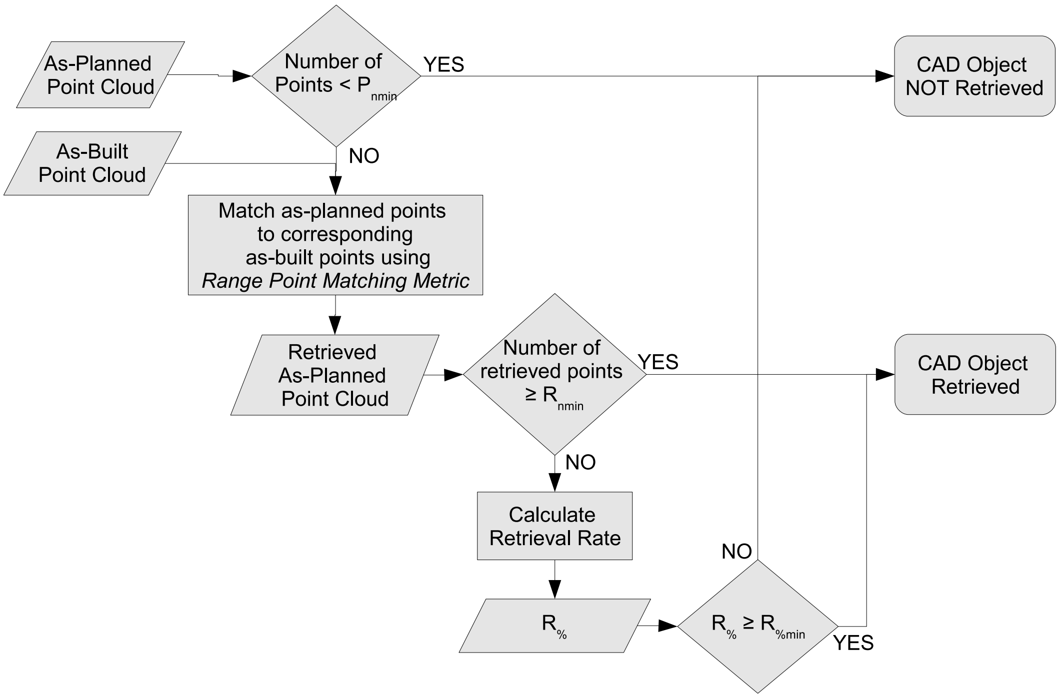


Figure 2. Object recognition/retrieval metric.

5 Experimental Results

In order to test the proposed approach, two indoor experiments have been conducted using a simple structure made of four columns and one board simulating a column-slab structure, a Trimble^{TM} GX3D laser scanner — the characteristics of which are presented in Table 1, and the 3D CAD engine Bentley^{TM} Microstation^{TM}. The first experiment aims at validating the approach. The second experiment aims at demonstrating how this approach could be successfully used for automated construction progress assessment.

It must be noted that, in these experiments, referencing is not performed using global positioning sensors but is simply performed manually, and referencing uncertainties are not considered. Also, as mentioned earlier, the thresholds used in the two metrics are manually *a priori* estimated.

```
Build As-Planned Point Cloud:
for each as-built point P(pan;tilt;range) do
    CREATE as-planned point P'(pan' = pan;tilt' = tilt;range' = +∞; ID' = NaN)
    for each STL-formatted object do
        for each STL triangle do
            if pan' ≥ pan_STLmin & pan' ≤ pan_STLmax & tilt' ≥ tilt_STLmin & tilt' ≤ tilt_STLmax then
                if ray(pan',tilt') intersects STL triangle & intersection point is closer
                    UPDATE range' ;
                    UPDATE ID' ;
                end
            end
        end
    end
end
SORT as-planned point cloud by ID' ;

For each CAD object, retrieve each individual point, deduce
retrieval metric values, and finally infer its retrieval:
for each CAD object do
    COMPUTE number of as-planned points, P_n;
    if P_n < P_nmin then
        CAD object is NOT RETRIEVED;
    else
        for each as-planned cloud point do
            COMPUTE Range Difference, ΔRange;
            if ΔRange ≤ ΔRange_max then
                As-planned point is retrieved ;
            else
                As-planned point is not retrieved ;
            end
        end
        COMPUTE Number of retrieved as-planned points, R_n ;
        if R_n ≥ R_nmin then
            CAD object is RETRIEVED ;
        else
            COMPUTE Retrieval rate, R_%;
            if R_% ≥ R_%min then
                CAD object is RETRIEVED ;
            else
                CAD object is RETRIEVED ;
            end
        end
    end
end
```

Figure 3. Algorithm for automated recognition/retrieval of STL-formatted project 3D CAD model objects in range point clouds.

5.1 Experiment 1: Approach Validation

5.1.1 Setup

In this first experiment, a 3D CAD model of the column-slab structure is initially developed using the 3D CAD engine and converted into STL format.

Table 1
Specifications of the Trimble GX3D Scanner

Model		GX3D
Laser type		Pulsed; 532nm; green
Distance	Range	2m to 200m
	Accuracy	1.5mm at 50m; 7mm at 100m
Angle	Range	Pan: 360deg Tilt: 60deg
	Accuracy	Pan: 60μrad Tilt: 70μrad

This model is composed of five CAD objects called: *column_1*, *column_2*, *column_3*, *column_4*, and *slab* (Figure 4). Then, the structure is manually built with as much precision as possible with respect to the 3D CAD model. Next, the entire scene is scanned with the laser scanner and the STL-formatted project 3D CAD model is manually referenced in the laser scanner's coordinate frame. Finally, the developed algorithm is run to automatically retrieve the STL-formatted 3D objects in the range data. Figure 5 shows the laboratory experimental setup with the column-slab structure and the laser scanner. Figure 6 displays the scene scan containing $206,360$ points, the size of each being proportional to its associated reflectivity. The following algorithm input parameters are used:

$\boldsymbol{\Delta Range_{min}}$**.** An as-planned cloud point is considered retrieved if the difference between its range and the range of the corresponding as-built point is less than 30 mm ($\Delta Range_{min}$). Construction generally aims at achieving dimensional accuracy within 10-20mm at most. Therefore, the authors consider that this threshold value is sufficiently high so that objects will not be missed due to some low construction dimensional quality, without creating false positive matches.

$\boldsymbol{P_{nmin}}$**.** The retrieval of a CAD object is performed only if its as-planned point cloud contains more than 100 points. This value is set somewhat arbitrarily and, as will be seen in the results, does not have an effect in this experiment.

$\boldsymbol{R_{nmin}}$**.** A CAD object is considered detected if at least 500 points of its as-planned point cloud are retrieved. Here also, this value is defined somewhat arbitrarily and its value does not have any specific impact in the context of this experiment.

$\boldsymbol{R_{\%min}}$**.** If less than 500 points (R_{nmin}) of a CAD object as-planned point cloud are retrieved, the object is considered retrieved only if its as-planned range point cloud retrieval rate is at least 50%. As discussed earlier, it is not obvious at this point in this research what is an acceptable $R_{\%min}$ value. As a result, in the absence of any *a priori* knowledge for setting this threshold, the authors decided to choose this midpoint value.

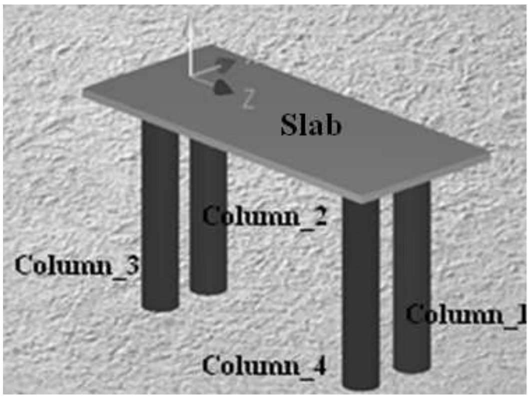


Figure 4. 3D CAD model of the column-slab structure.

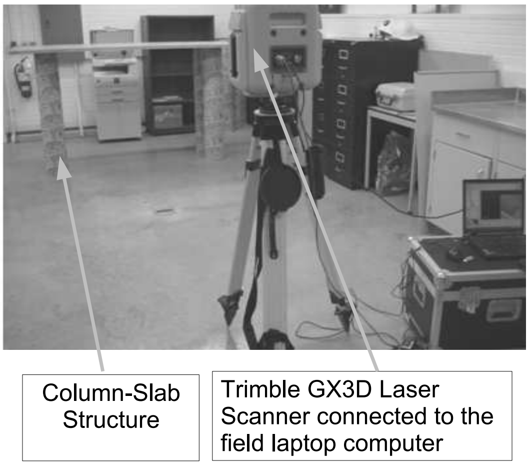


Figure 5. Indoor setup with the scanned structure and the laser scanner.

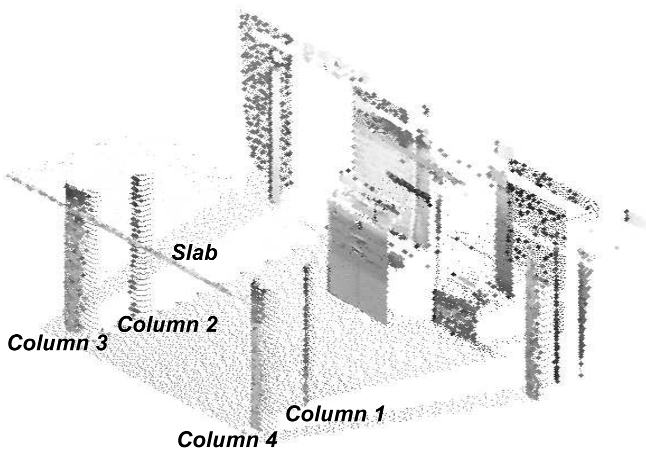


Figure 6. Experiment 1 range point cloud. The size of each point is proportional to its scanning reflectivity.

5.1.2 Results

The retrieval results are presented in Figure 7 and Table 2. Figure 7 displays the as-built, as-planned, and retrieved as-planned data. In this figure, only 1% of the total number of points of each cloud is actually displayed in order to increase picture clarity. Also, in the retrieved as-planned point cloud, retrieved as-planned points are displayed with circles and non-retrieved ones are displayed with asterisks.

Table 2 shows that all CAD objects from the 3D CAD model are retrieved. The retrieval rates of all CAD objects are high (at least 74%), including *column_1* and *column_2* despite the fact that, as can be seen in Figure 6, about 60% of their normally visible surfaces are occluded by *column_4* and *column_3* respectively. This demonstrates the robustness of this method with respect to occlusions due to other CAD objects.

It is also interesting to note that the slab is detected with a high but slightly lower retrieval rate (74%) than the other objects. A reason for this can be found in Figure 6. Remember that in this figure the size of each point is proportional to its associated reflectivity. Reflectivity can be seen as an estimator of range acquisition uncertainty, and it can be noticed that most points obtained from the slab, especially from its top surface, have a very low reflectivity. The manually set $\Delta Range_{min}$ threshold might thus have been too low to retrieve these specific points. Another reason could be error in vertical referencing. Indeed, in this example, a little error in the vertical referencing would shift the as-built slab cloud compared to the as-planned one, which would considerably alter the object retrieval results. The effect of referencing uncertainty is further discussed in Section 6.

Table 2
Retrieval results of Experiment 1.

Calculated Values	CAD Element				
	column_1	*Column_2*	*Column_3*	*Column_4*	*Slab*
Number of As-Planned Points	5,079	4,678	17,490	17,880	4,712
Number of Retrieved As-Planned Points	4,423	4,411	16,403	16,120	3,479
Retrieval Rate of As-Planned Points	87%	94%	94%	90%	74%
Detected	YES	YES	YES	YES	YES

Although these experimental results are very positive, it is acknowledged that they were obtained in a somewhat ideal indoor setup. In fact, in this experiment, all CAD objects are retrieved without considering retrieval rates (even *column_1* and *column_2*) as the total number of retrieved points are always higher than R_{nmin}. In field situations, it is likely that the number of retrieved points, the retrieval rates and the number of as-planned points would not always be so high, in which case the values of the corresponding thresh-

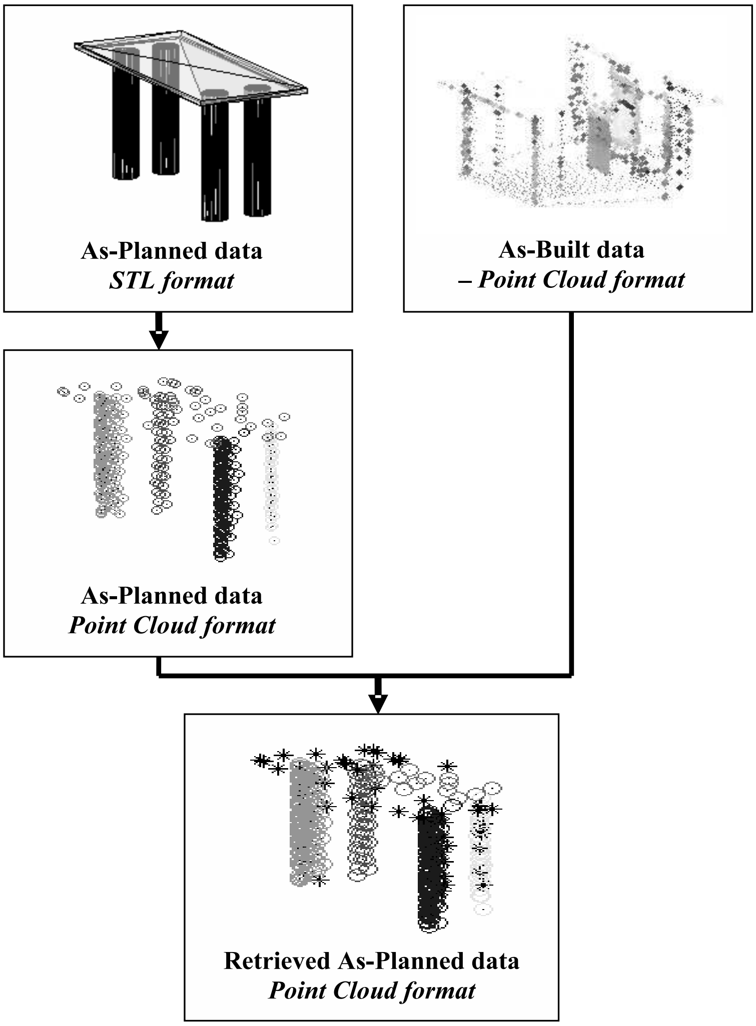


Figure 7. As-built and as-planned data at different stages of the retrieval process in Experiment 1 (only 1% of the total number of points of each cloud is displayed to increase clarity).

olds ($\Delta Range_{min}$, P_{nmin}, R_{nmin} and $R_{\%min}$) would have a higher impact on the retrieval results. More robust methods to automatically estimate these thresholds are thus suggested in Section 7.

5.2 Experiment 2: Application to Construction Progress Assessment

5.2.1 Setup

The goal of this second experiment is to demonstrate how this new approach could be applied to automated construction progress assessment. In this experiment, the same setup is used. The difference is that instead of a project 3D CAD model, a project 4D CAD model is used. It is built using the project 3D CAD model displayed in Figure 4 and the simple construction schedule, for which the bar chart is shown in Figure 8(a). The resulting as-planned project 4D CAD model is displayed in Figure 9(a). Then, the same scene as in Experiment 1 is scanned (Figure 6) and is assumed to occur on day 4 of the construction. The goal of the experiment is to retrieve all project 3D objects in the scan, and identify whether construction is on schedule, early, or late. The following input parameters are used:

Schedule Uncertainty. A one-day uncertainty in schedule is used so that work completed earlier or later by one day can be identified. This implies that the scanned data is compared with three consecutive project 3D CAD models extracted from the project 4D CAD model and centered on the day when the scan is conducted (here day 4).

$\boldsymbol{\Delta Range_{min}}$**.** Same as in Experiment 1 (30mm).

$\boldsymbol{P_{nmin}}$**.** Same as in Experiment 1 (100 points).

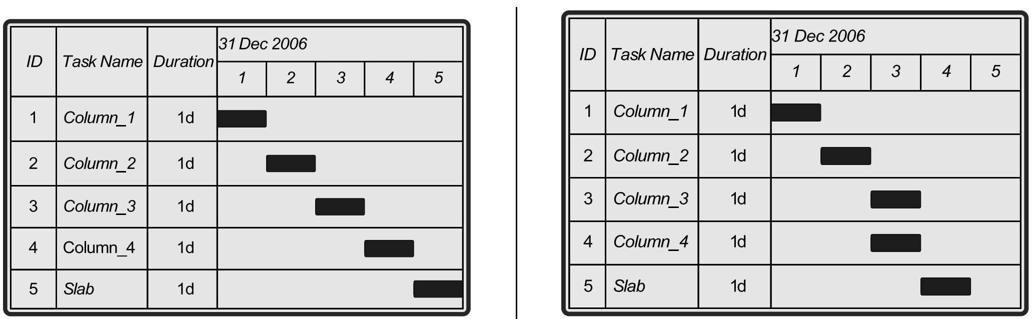
$\boldsymbol{R_{nmin}}$**.** Same as in Experiment 1 (500 points).

$\boldsymbol{R_{\%min}}$**.** Same as in Experiment 1 (50%).

5.2.2 *Results*

Table 3 summarizes the results obtained in this experiment. It shows that all 3D objects in day 5 project 3D model are retrieved in the scanned data. The retrieval of each object is made with a minimum of 4,500 retrieved as-planned points per object and very high retrieval rates. Since the scan is assumed to take place on day 4, it can be concluded that construction is one day ahead of schedule. The bar chart of a possible resulting as-built schedule is displayed in Figure 8(b) and the corresponding as-built 4D CAD model is presented in Figure 9(b).

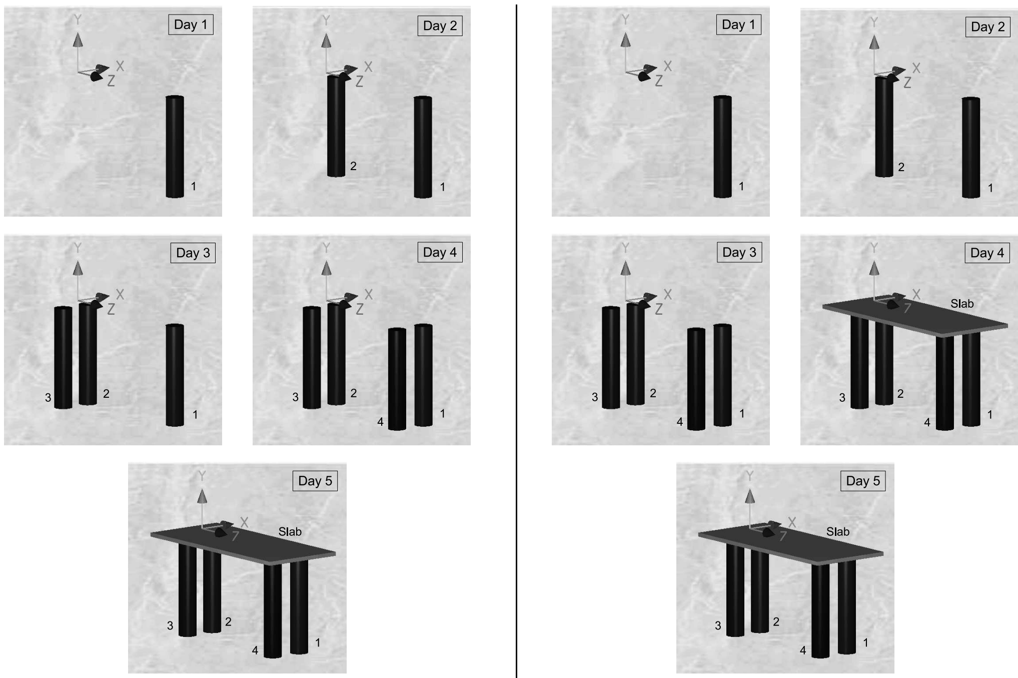
Certainly, the metric used here to identify early, on time or late construction is very basic. However, these results demonstrate that this approach has great potential for supporting automated project work progress tracking.



(a) As-planned Schedule

(b) As-built Schedule

Figure 8. As-planned and as-built schedules of the construction of the column-slab structure



(a) As-Planned 4D CAD model (b) As-Built 4D CAD model

Figure 9. As-planned and as-built 4D CAD models of the construction of the column -slab structure

Table 3
Retrieval results in Experiment 2

Day	Calculated Values	CAD Element				
		column_1	*Column_2*	*Column_3*	*Column_4*	*Slab*
3	**Number of As-Planned Points**	15,042	4,678	17,490	0	0
	Number of Retrieved As-Planned Points	4,684	4,411	16,403	0	0
	Retrieval Rate of As-Planned Points	31%	94%	94%	0%	0%
	Detected	YES	YES	YES	NO	NO
4	**Number of As-Planned Points**	5,079	4,678	17,490	17,880	0
	Number of Retrieved As-Planned Points	4,423	4,411	16,403	16,120	0
	Retrieval Rate of As-Planned Points	87%	94%	94%	90%	0%
	Detected	YES	YES	YES	YES	NO
5	**Number of As-Planned Points**	5,079	4,678	17,490	17,880	4,712
	Number of Retrieved As-Planned Points	4,423	4,411	16,403	16,120	3,479
	Retrieval Rate of As-Planned Points	87%	94%	94%	90%	74%
	Detected	YES	YES	YES	YES	YES

6 Impact of measurements uncertainties

The previous experiments were conducted with somewhat ideal conditions and all measured values were considered exact. In construction site applications, measurement uncertainty could be non negligible and should therefore be estimated and taken into account in the object retrieval process. In the investigated problem, measurement uncertainties include: *referencing uncertainties* and *laser measurement uncertainties*.

6.1 Referencing uncertainties

Referencing uncertainties refer to uncertainties in the 3D CAD model geo-referencing or/and in the range point cloud geo-referencing. These can be translated into a single set of referencing uncertainties which is the difference between the real and virtual geo-positions of the laser scanner. This referencing uncertainty includes uncertainties in location (*northing*, *easting*, *altitude*) and in orientation (*heading*, *pitch*, *roll*). *Northing*, *easting*, *altitude*, *heading*, *pitch* and *roll* can be obtained using different global positioning technologies. However, the accuracy that these technologies can currently achieve is limited to several centimeters in location and half a degree in orientation at best. These uncertainties are significant enough that their impact on object recognition systems that use these technologies can be non-negligible.

A method is suggested here for the automated correction of referencing error. This correction can be made prior to performing the actual point retrieval process. For each of the six 3D model referencing parameters (*northing*, *easting*, *altitude*, *heading*, *pitch* and *roll*), uncertainty is modeled with a discrete distribution with three values centered on the measured one. Then, for each combination of six discrete values (one discrete value for each of the six pose parameters), the retrieval of a fixed number of random range points, $n_{rpoints}$ (for instance $n_{rpoints} = 600 points$) is performed using the approach described in this paper. The likelihood of each combination being the best referencing is calculated using a mean square error estimator based on the range differences between the $n_{rpoints}$ as-built points and their corresponding as-planned points. The best referencing estimation is the one with the lowest mean square error. If a better referencing is identified for a set of six values with at least one of them different from its corresponding measured one, the measured values are correspondingly updated and this process is reiterated. This iteration occurs until the best pose is the one with the six parameters set to their measured values.

Although each pose improvement increment requires the analysis of 3^6 combinations of discrete pose values, note that the complexity of this method is fixed with respect to the number of as-built range points, as only a subset of a fixed number of points is used. Also, it is acknowledged that this method requires estimating the parameters necessary for the description of the different discrete distributions (distribution type, space between values in each distribution, n_{unc}). Previous research using likelihood estimators suggest that a value of $n_{rpoints} = 100 \cdot n_{param}$, where n_{param} is the number of uncertainty parameters (here six), is statistically sufficient. Then, the type of discrete distribution to use is not obvious. By default, it is thus suggested to consider equal probabilities for each discrete value (uniform discrete distribution). Finally, the space between values in each distribution could be set as one time or half the measurement uncertainty.

At this time, this correction approach has only been tested a couple of times, using manually defined discrete uniform distributions. While the results seemed fairly good, a comprehensive set of experiments would be required to confirm the efficiency and robustness of this approach for automated pose correction. Additionally, the adequacy of basing the mean square error estimator on range differences can be discussed. Indeed, range difference may provide different results than orthogonal projection distance which is more commonly used because more intuitive.

6.2 Laser measurement uncertainties

Laser measurement uncertainties relate to the uncertainties in the measurement of each individual point. They include uncertainties in pan, tilt and range values.

Pan and tilt uncertainties result from imperfections in the laser scanner embedded pan&tilt unit. While pan and tilt uncertainties are independent from the scanned surface, it must be noted that they are also generally considered value independent. Pan and tilt uncertainties are provided by laser scanner providers. In the case of the scanner used in this research, pan and tilt uncertainties are respectively $60\mu rad$ and $70\mu rad$ ($0°0'12''$ and $0°0'14''$). These respectively translate into $0.6mm$ and $0.7mm$ accuracy at $10m$, or $6mm$ and $7mm$ accuracy at $100m$. A common approach to take such uncertainties into account when determining a point range is to analyze the ranges of all the points neighboring the studied one. Such an approach is however inappropriate here since the pan and tilt angle uncertainties are much lower than the maximum pan and tilt point densities that the scanner can achieve. Another more computationally complex method is the calculation for each point of several “intermediate” range values obtained with different combinations of pan and tilt angles adjusted with uncertainty. All the “intermediate” ranges could then be analyzed to infer the most probable point range. This method is similar to the one proposed above for referencing correction.

Range uncertainty is related to several factors including: the scanning angle to the scanned surface, the material of the scanned surface, environmental conditions, etc. Range measurement uncertainty is generally provided by laser scanner providers for specified material reflectivity and with scanning directions perpendicular to the scanned surface. The laser scanner used in the experiments above presents the following range "best" uncertainties: $1.5mm$ at $50m$ and $7mm$ at $50m$ for 100% reflective targets. A possible method to take range measurement uncertainty into account when matching two as-built and as-planned points is presented in Section 7.2 when discussing the automated estimation of the threshold parameter $\Delta Range_{nmin}$.

Overall, it must emphasized that these laser measurement uncertainties remain negligible when compared with current geo-referencing uncertainties.

7 Thresholds Parameters Estimation

The proposed object recognition approach uses two metrics that require some input threshold parameters: $\Delta Range_{min}$, P_{nmin}, R_{nmin} and $R_{\%min}$. In the experiments presented in this paper, these thresholds were manually *a priori* estimated. But for a complete automated approach, these would have to be automatically estimated, especially since their values should be adjusted to different scanning and scene condition factors.

7.1 P_{nmin}, R_{nmin} and $R_{\%min}$

In the *object recognition metric*, P_{nmin}, R_{nmin} and $R_{\%min}$ could be estimated by taking into consideration the following factors:

Scan point density. The scan point density is the pan and tilt difference between two neighboring points. If a scene is scanned twice with two different point densities, one twice denser than the other, the as-built and resulting as-planned point clouds of each scanned object will contain twice more points in the denser scan. It is therefore possible that for a given manually a priori estimated P_{nmin} value, an object is considered for search with the denser scan and not with the less dense one. Similarly, it is possible that for a given manually *a priori* estimated R_{nmin}, the retrieval rate of an object will have to be calculated with the less dense scan, but not with the denser one. Since scan point density should not have any effect on the retrieval metrics, P_{nmin} and R_{nmin} must be adjusted to it: $P_{nmin} = f_1(d_{scan})$ and $R_{nmin} = f_2(d_{scan})$, where the functions $f_1()$ and $f_2()$ could be *a priori* experimentally estimated. Note, that $R_{\%min}$ is not impacted by the scan point density as it is expressed as a percentage of points that is invariant with this factor.

Scanner-object (or scanner-STL triangle) distance. The same argument can be made with two exactly similar objects that are at different distances from the scanner, one twice further than the other. P_{nmin} and R_{nmin} should thus be automatically adjusted for each object, and consequently for each STL triangle, by taking the as-planned scanner-STL triangle distance into account. The as-planned distance between the scanner and a STL triangle, $Range_{STL}$, can be estimated as the mean of the distance between the scanner and the three STL triangle vertices. As a result, P_{nmin} and R_{nmin} could be further customized for each STL triangle such that: $P^{STL}_{nmin} = f^{STL}_1(d_{scan}, Range_{STL})$ and $P^{STL}_{nmin} = f^{STL}_2(d_{scan}, Range_{STL})$, where the functions $f^{STL}_1()$ and $f^{STL}_1()$ could be *a priori* experimentally estimated. Note again that $R_{\%min}$ is not impacted by the scanner-STL triangle distance as it is expressed as a percentage of points that is invariant with this factor.

While methods for automating the estimation of P_{nmin} and R_{nmin} are presented here, no method is suggested for $R_{\%min}$. For $R_{\%min}$, the authors suggest, with lack of experience to use the midpoint value of 50%.

7.2 $\Delta Range_{nmin}$

In the *point matching metric*, $\Delta Range_{min}$ could be estimated by taking into consideration the following factors:

Range. As presented earlier, range measurement uncertainty depends on many factors. It is nonetheless generally provided by laser scanner providers for specified material reflectivity and with scanning directions perpendicular to the scanned surface. In Section 6.2, it can be seen in the specifications of the scanner used in this research that range uncertainty increases with range (this is true for any scanner). Therefore, the threshold parameter $\Delta Range_{min}$ should be customized for each scanned point, p : $\Delta Range^p_{min} = f^p_3(Range_p)$, where $Range_p$ is the measured range of point p, and $f^p_3()$ could be estimated *a priori* through multiple experiments.

Reflection angle. Uncertainty in range acquisition increases with the reflection angle between the point scanning direction and the scanned surface normal vector. The impact of the reflection angle on range uncertainty is illustrated in Figure 10. The as-planned reflection angle of each as-planned range point could be estimated when calculating the as-planned point. This estimation could then be used to further customize the $\Delta Range^p_{min}$ threshold: $\Delta Range^p_{min} = f^p_3(Range_p, RefAngle_{STL})$, where $RefAngle_{STL}$ is the point p as-planned reflection angle, and $f^p_3()$ could be *a priori* experimentally estimated.

Surface reflectivity. Finally, acquired range uncertainty decreases with surface reflectivity. If an estimated object surface reflectivity could be obtained from the material applied to the objects in the original project 3D CAD model, then each STL triangle could be assigned an estimated reflectivity and the function $f_3^p()$ and consequently the threshold $\Delta Range_{min}^p$ could be further customized.

Overall, while methods for automatically estimating the different input parameters used in the proposed object retrieval approach are presented here, these still require the predetermination of some functions $f_1^{STL}()$, $f_2^{STL}()$ and $f_3^p()$ through a comprehensive set of experiments. These experiments have not been conducted yet and would require a complex test bench. The need for such experiments has been expressed in previous work and the National Institute for Standards and Technology (NIST) has been working on the construction of such a facility for comprehensive LADAR performance evaluation [16].

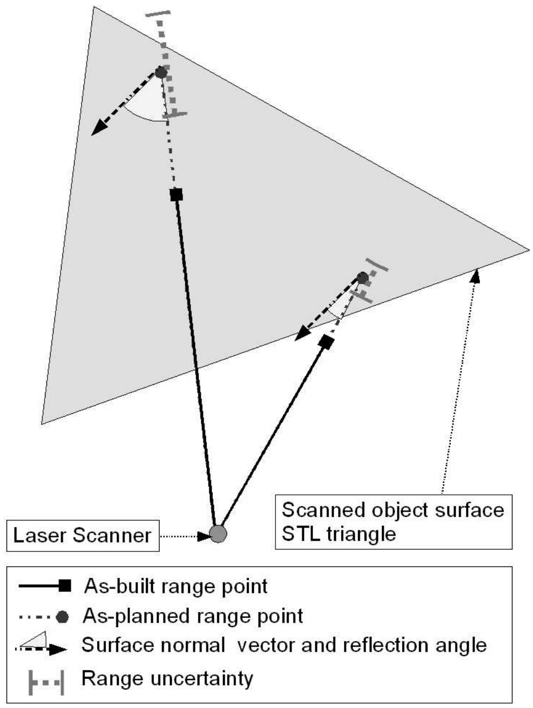


Figure 10. Impact of the reflection angle on the acquired range uncertainty.

8 Conclusion and Future Work

The cost of 3D range scanning is rapidly declining due to recent developments, and use of 3D images is increasing accordingly. In this paper, a new approach for automatically retrieving 3D CAD objects in 3D range point clouds is presented. This approach takes advantage of 3D/4D CAD models and (geo-) referencing technologies. Experimental results first demonstrate that this completely automated approach is quite robust, including in the case of occlusions due to other CAD elements. The second experiment further illustrates these strengths and demonstrates how it could robustly support applications such as automated construction progress tracking. Future work will focus on confirming these results with full-scale structures. The impact of uncertainties in (geo-) referencing values and in point measurement values will be further investigated, and methods for automating the estimation of the required threshold parameters will also be further tested.

Finally, the authors would like to re-emphasize the fact that this new approach has applications not only in automated construction work progress tracking, but also in construction quality control, in 3D image database information retrieval, and very likely in many other areas.

9 Acknowledgements

This project is partially funded by a grant from the National Science Foundation grant #0409326 and the Canada Research Chair in Construction & Management of Sustainable Infrastructure.

References

[1] Akinci, B., Boukamp, F., Gordon, C., Huber, D., Lyons, C. & Park, Kuhn, A formalism for utilization of sensor systems and integrated project models for active construction quality control, Automation in Construction, Elsevier, New York, USA, 2006, vol. 15(2), pp. 124-138.

[2] Gordon, C. & Akinci, B., Technology and process assessment of using LADAR and embedded sensing for construction quality control, ASCE Construction Research Congress (CRC), San Diego, CA, USA, 2005, pp. 557-561.

[3] Navon, R., Research in automated measurement of project performance indicators, Automation in Construction, Elsevier, New York, USA, 2007, vol. 16(2), pp.176-188.

[4] Song, J., Haas, C.T. & Caldas, C.H., Tracking the location of materials on construction job sites, ASCE Journal of Construction Engineering and Management, ASCE, New York, USA, 2006, vol. 132(9), pp. 911-918.

[5] Cheok, G.S., Stone, W.C., Lipman, R.R. & Witzgall, C., LADARs for construction assessment and update, Automation in Construction, Elsevier, New York, USA, 2000, vol. 9(5), pp. 463-477.

[6] Arman, F., & Aggarwal, J.K., Model-based object recognition in dense-range images — A review, Computing Surveys, ACM, New York, USA, 1993, vol. 25(1), pp. 5-43.

[7] Arman, F., & Aggarwal, J.K., Object recognition in dense range images using a CAD system as a model base, Proceedings of the 1990 IEEE International Conference on Robotics and Automation, 1990, vol. 3, pp. 1858-1863.

[8] Johnson, A.E. & Hebert, M., Using spin images for efficient object recognition in cluttered 3D scenes, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, vol. 21(5), pp. 433-449.

[9] Reid, I.D. & Brady, J.M., Model-based recognition and range imaging for a guided vehicle, Image and Vision Computing, Elsevier, New York, USA, 1992, vol. 10(3), pp. 197-207.

[10] Teizer, J., Caldas, C.H. & Haas, C.T., Real-time three-Dimensional modeling for the detection and tracking of construction resources, Journal of Construction Engineering and Management, ASCE, New York, USA, 2007, (in press).

[11] Brilakis I.K., Automated on-site retrieval of project information, 13th Workshop - Intelligent Computing in Engineering and Architecture, EG-ICE, Ascona, Switzerland, 2006, pp. 92-100.

[12] Brilakis I.K. & Soibelman L., Multimodal image retrieval from construction databases and model-based systems, Journal of Construction Engineering and Management, ASCE, New York, USA, 2006, vol. 132(7), pp. 777-785.

[13] Brucker, B.A. & Nachtigall, S.D., Initiating the building information model from owner and operator criteria and requirements, International Conference on Computing in Civil Engineering, ASCE, Cancun, Mexico, 2005.

[14] Bosche, F. & Haas, C.T., Investigation of data formats for the integration of 3D sensed and 3D CAD data for improved equipment operation safety, 1st International Construction Specialty Conference (ICSC), CSCE, Calgary, AB, Canada, 2006.

[15] 3D Systems Inc., *Stereolithography Interface Specification* (1989).

[16] Cheok, G.S. & Stone, W.C., Performance evaluation facility for LADARs, Proceedings of SPIE — Laser Radar Technology and Applications IX, 2007, vol. 5412, pp. 54-65.