Challenges of Identifying Steel Sections for the Generation of As-Is BIMs from Laser Scan Data

Engin Burak Anil [1]*, Raghuram Sunnam [2], and Burcu Akinci [1]

[1] *Department of Civil and Environmental Engineering, Carnegie Mellon University, USA*
[2] *School of Architecture, Carnegie Mellon University, USA*
** Corresponding author (eanil@andrew.cmu.edu)*

Purpose When a laser scan is performed and no prior information is available about the building, standard sections of components need to be identified from point cloud data in order to generate informative as-is building information models (BIMs). Currently, the standard steel sections used at a site are not automatically identified from the point cloud data. Various issues related to the laser scan data, challenge automation, such as occlusions, missing data points, angle of incidence, and imprecision of measurements on the data. **Method** The research described in this paper relates to the manual determination of steel beam sizes used in a steel worker training facility, which contained about 16 beams, 63 columns, and 12 scans collected over 4 days of construction. **Results & Discussion** We identified that occlusions and noise are the major challenges associated with recording accurate dimension measurements. The identification of correct steel sections based on such inaccurate measurements is even more challenging since the decision should be based on all defining (e.g., flange width, depth) dimensions.

Keywords: *as-is BIM, as-built BIM, point cloud data, laser scanning, steel structures*

INTRODUCTION

As-built or as-is Building Information Models are being generated for a variety of purposes, including retrofitting, architectural renovation, documenting site conditions, spatial program validation [1]. Laser scanners are getting a lot of attention for collecting 3D data regarding the geometry of the building for BIM generation with the advantages of speed, coverage, and considerably long range [5]. Identifying section dimensions from the laser scan data imposes challenges. It was reported that mixed pixels can cause sectional losses up to 5.6 cm on bridge column dimensions [3]. Furthermore, occlusions, low point density, and surface reflections can further reduce the quality of the information obtained from the laser scan data [2].

Identification of steel sections is one of the most challenging cases of the object identification problem. In steel structure identification, sub-millimeter level accuracy is generally required. The structures can be complex and so the occlusion rates can be high. Steel components can have shiny surfaces leading to reduced data accuracy. The aim of this study is twofold. First, we would like to understand the range of errors that can occur due to the unique nature (e.g., small section sizes, accuracy requirements, occlusions, complexity) of steel structures and components. Second, we would like to understand the challenges of identifying steel sections from the laser scan data.

In order to investigate the severity of the problem, We performed a case study on an objective data set. By objective we mean that the data set shouldn't be generated with special precautions to impact the results of this study. It should depict the generic problems one might face when performing a similar task on a different data set. The testbed used in this study is a training facility for steel erectors. The laser scan data was collected for automatic construction site progress and process monitoring in sequences. The BIM was generated by a professional company for potential uses such as evaluation of object recognition algorithms.

In this study, we manually recorded the all six dimensions of the I-Beams (i.e. section height, bottom and top flange widths, web thickness, and bottom and top flange thicknesses) composing the steel structure. Then, for each section, we identified the top 3 sections from the AISC Steel Sections[4] table that would fit to the dimensions identified. Assuming that no prior information is available regarding the structure, the identification stage was performed independent of the BIM. Identified sections were ordered by how well the identified sections describe the true sections. The identified sections are compared to the BIM for quantifying the identification accuracy. Potential challenges of identification of standard steel sections based on laser scan data were evaluated by observation.

Identification of steel sections from laser scan data

In order to understand potential challenges of steel section identification from the laser scan data and study the problem, we set up a testbed based on an objective data set. Using the laser scan data, we first manually identified the steel sections from the laser scan data. Then, we compared the identified steel sections with the reference BIM. Finally, by studying the patterns of bad estimates and the laser scan data, we identified the challenges of the identification problem specific to steel modeling.

The Testbed

The testbed used for this study is a training facility used for training ironworkers in erecting a steel structure. The purpose of the scanning was to monitor the process and the progress of the erection. The structure was laser scanned and the data was collected over a period of 4 days from 12 different locations.

The structure was scanned after every component was erected. Hence, the scan data captures the progress of the erection. The scans were performed when there is erection activity on site and workers are present. Scans were registered using fiducial targets.

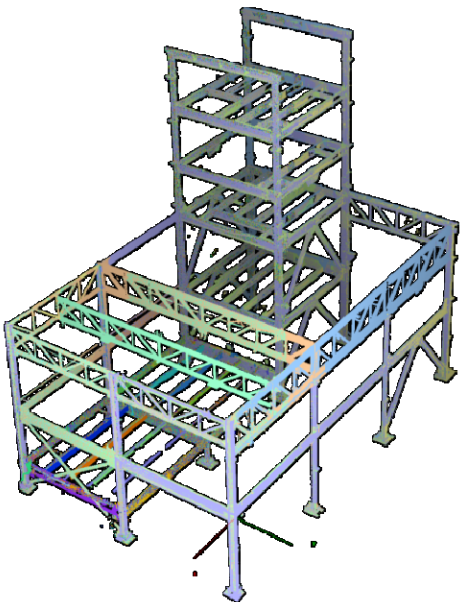

Fig. 1. Registered laser scan data. The coloring refers to the different scans collected at different times.

Fig. 1 presents a snapshot of the registered scan data. Individual scans are colored with different colors. It can be seen that points corresponding to some of the components have consistent colors throughout the components. This means that those components are only scanned from a single location. Some other components have more than one color. This means that those components are partially scanned from multiple locations. Trusses on the second floor are good examples to the second case.

The configuration of the steel structure during the erection composed of a permanent 4 story tower and the lower two stories that were erected during the training. About 36.5 million points were collected over the course of the training. In the figure it can be seen that there are also a partial basement floor (Level 0) and two beams on the top floor of the tower, which we called Level 5 throughout the paper.

A professional company generated the BIM of the structure. The BIM of the structure and the laser scan data constituted the testbed of this study.

Measurement method

Currently, such modeling tasks are carried out manually. Hence, we used a manual method for measuring the dimensions of the steel sections. However, several approaches can be adopted for manually recording the measurements. In order to identify the best suiting method to our testbed, we examined four different manual approaches on a small portion of the laser scan data.

The metrics we used for the evaluation of the manual measurement methods were overall accuracy of identification, easiness, and time. We define easiness as how comfortably the operator performs the measurement task with a given method. We asked the operator to compare the methods verbally to assess the easiness.

The measurement methods used for recording the section dimensions are as following:

1- **Point-to-point measurement:** This method manually selects two points for every dimension (e.g. top flange width) and records the distance between these points.

2- **Distance between edges:** In this method, the distances between the edge lines of a component are recorded. The edges are created by manually selecting the points along the edges and fitting lines to the points belonging to the edges.

3- **Distance between plane-plane intersection lines:** Similar to the second method, the distances between the edges are recorded. However, in this method, first planes are fitted to the surfaces of the components (i.e., flange surfaces and web surfaces). These planes are intersected to find an estimate of the location of the edges.

4- **Cross-section tracing:** In this method, cross-sections are cut through the point cloud. Projecting the points that are within a small tolerance from the cutting planes and connecting the points generates cross-section traces. Dimensions are measured on these traces.

For each of the tested methods, 6 measurements were made along the components for every dimension of the components. For example, AISC steel section tables define I-beams using 6 dimensions (i.e., the depth of the section (D), the width of the top flange (W1), the width of the bottom flange (W2), the thickness of the top flange (t1),the thickness of the bottom flange(t2), the thickness of the web(t3)). Therefore, for an I-beam a total of 36 measurements should be made (6 measurements for each dimension times 6 dimensions).

After recording the measurements, we compared the measurements to the BIM dimensions. Based on the application of the four methods on a small subset of the scan data, we assessed that the first method (point-to-point measurement method) is equally accurate compared to the other three methods while being easier and faster to apply. Though, it should be noted that this assessment cannot be readily generalized without further investigation since the comparison is workflow specific. For our purposes, however, the aim is identifying a method that can produce accurate measurements on the point cloud. Therefore, we selected the first method for the rest of the study.

Best Section Selection

After all the measurements are recorded for all of the components, standard steel sections that correspond to those measurements are identified. Since the measurements were performed independent of the BIM and assuming no prior information, first, the shapes of the sections were identified. It was assessed that all of the components are W-shaped steel sections.

The standard section closest to the measurements was named the best fitting section (or the best estimate). It was observed that the average measurements were not always exactly equal to the dimensions of the standard steel sections from the AISC table. Additionally, the measurements had standard deviations of several millimeters. With such variation, it was not always possible to narrow down the selection to a single standard section. Therefore, we selected the top three standard sections that best fit the measurements.

Ideally, all six measured dimensions of the steel sections should be treated equally when comparing the measurements to the AISC section table. However, flange and web thicknesses were often at the order of the standard deviations recorded. The increments of thicknesses between different standard steel sections were almost always smaller than the standard deviation of the measurements recorded. For example, for a component with a measured flange thickness of 1.73 mm and standard deviation of 0.2 mm, the selection should be made between 33 sections. For the same component with the average section depth of 40 mm and standard deviation of the depth of 0.2 mm, there is only one W-shaped component in the AISC table. Therefore, preference was given to the depth of the section and the flange widths over thickness dimensions.

Table 1. Results of the identification of steel sections from the laser scan data

Results for Columns	
Total number of columns	16
1^{st} best estimate is correct	3 (18.75%)
2^{nd} or 3^{rd} best estimate is correct	8 (50%)
No hit	5 (31.25%)
Results for Beams	
Total number of beams	63
1^{st} best estimate is correct	25 (39.68%)
2^{nd} or 3^{rd} best estimate is correct	24 (38.1%)
No hit	14 (22.22%)

RESULTS

The reference BIM file was used to check the accuracy of the best fitting sections. The three best fitting sections were compared to the steel sections from the as-designed BIM file. The components where the best fitting section was matching with the correct section from the BIM were referred to as the accurate estimates. There were also cases where the 2^{nd} and the 3^{rd} best estimates were matching to the correct sections from the IFC file. The cases where none of the best sections were matching with the correct section from the BIM were classified as 'No hit'.

For the columns, our best estimate was the correct section for only 18.75% of the cases. The second or the third best estimate was the correct section for 50% of the cases. For almost one third of the columns we were not able to determine the correct section.

For the beams, the results were slightly better. We were able to find the correct steel section for 39.68% of the beams. For 22.22% of the beams, we were not able to determine the correct section.

CHALLENGES OF IDENTIFYING STEEL SECTIONS BASED ON THE SCAN DATA

Based on the obtained results and close investiga-

Table 2. Distribution of occlusions by level

	Total number of components	Self -occluded	Unrelated objects	Undefined edges
Columns	16	2 *(12.5%)*	1 *(6.25%)*	1 *(6.25%)*
Level 0 Beams	17	1 *(5.88%)*	13*(76.47%)*	12*(70.59%)*
Level 1 Beams	15	2*(13.33%)*	5*(33.33%)*	8*(53.33%)*
Level 2 Beams	11	7*(63.63%)*	8 *(72.72%)*	7 *(63.63%)*
Level 3 Beams	7	6 *(85.71%)*	3 *(42.86%)*	5 *(71.43%)*
Level 4 Beams	11	7 *(63.63%)*	6 *(54.54%)*	8 *(72.72%)*
Level 5 Beams	2	1 *(50%)*	2 *(100%)*	1 *(50%)*

tion of the scan data, we identified that there are several reasons as to why identification of correct steel sections is a difficult task. Namely, the challenges that can be attributed to this difficulty are associated with occlusions and noise in the laser scan data.

Occlusions

Obstruction of the laser beams by the components in the structure or by other objects (e.g. construction equipment, workers, etc.) causes occlusions in the laser scan data. From the BIM generation perspective occlusions prevents either the identification of the shape of the components or the correct identification of the dimensions of the components. The former is caused when a component is completely occluded. The latter is caused by the fact that due to occlusions, some of the dimensions of the components cannot be measured. A common example is that when a flanged beam is scanned only from one side the web thickness cannot be measured. Therefore in order to measure all dimensions, the components should be scanned from all sides.

In our testbed, for only one of the components the occlusions prevented any dimension measurement. However, partial occlusion rates (ratio of the number of partially occluded components to the total number of components) were high.

Partial occlusion rates for each floor are given in *Table 2*. Rates are calculated as the number of occluded components on a given floor divided by the total number of components on that floor. For example, 85.71% of beams on level 3 had partial self-occlusions. Unrelated objects partially occluded 42.86% of the components on the same floor (Table 2).

Missing data almost always resulted in wrong measurements. We identified a special case, however, where although the standard deviation of the measurements were low the measurements were not close to the actual dimension. In such cases, high precision (low standard deviation of measurements implies high precision) did not mean accuracy of identifying the best section accurately. This special case is caused by consistent occlusion throughout the component. The occluded parts of the section cannot be measured as expected. Additionally, since the occlusion is consistent through the member the spread of the recorded measurements are low although they are far from the actual values. For example, the standard deviation of the measurements of depth of section for Beam 108 was only 0.13mm (less standard deviation implies a precise measurement) and the average measured depth was 3.19mm. But the actual depth of section is 15.88 mm.

Noise

Noise in the data caused the spread of the measurement values to a wide range. Data noise was mostly caused by mixed pixels as the laser beams hit the edges of the components[3]. This spread is measured by the standard deviation of the measurements. Standard deviations were often comparable or even larger than the difference between consecutive sections in the AISC steel sections database. This constitutes a challenge because it is no more possible to narrow down the search to a single component.

In the identification of the best estimate the best estimates were chosen to the sections within the interval of [μ-σ , μ+σ] where 'μ' is the mean value of the measurements and 'σ' is the standard deviation of the measurements recorded.

Table 3 shows example cases of the number of sections falling within one σ range of the mean value of measurements. The sections are counted individually based on each dimension. For example, for

Table 3. Number of sections falling within one standard deviation of the mean measurement [μ-σ , μ+σ] for each dimension.

Section	Based on D	Based on W1	Based on W□	Based on t1	Based on t2	Based on t3
Column 1	5	20	20	115	44	65
Column 2	21	46	30	48	53	33
Column 4	13	64	19	0	52	275
Column 8	37	85	94	105	85	111
Beam 108	0	0	0	0	0	0
Beam 109	0	0	0	77	0	0
Beam 112	0	0	0	52	0	0
Beam 113	0	0	0	0	0	0
Beam 305	21	9	5	0	0	92
Beam 503	21	2	3	47	12	22
Beam 504	19	2	0	36	31	33
Beam 508	16	5	0	0	0	0

Column 1, basing the best fitting section on only the section depth gave 5 choices whereas basing the best fit guess only on the top flange thickness gave 115 choices.

Also, finding the best fitting sections for some sections was a challenge as none of the dimensions were helpful in finding best fits. For example, Beams 108 and 113 have absolutely no best fitting sections from any of the dimensions. The reason was that the mean measurement values did not coincide with any value in the table and within one σ range there were no sections.

Beam 109 has 77 best fitting sections and beam 112 has 52 best fitting sections based on the top flange thickness. The solution sets were considered to be very large. In such cases, the standard sections with the closest values to the mean measured dimensions were chosen as the best fitting sections.

Additionally, the section estimates with respect to individual section dimensions, such as those in **Error! Reference source not found.**, did not have intersecting sets of sections for some of the cases. Therefore, for those cases, using more dimensions did not help with the selection.

Summary and Conclusions

In this paper, we reported on our research on manual identification of the standard steel sections from the as-is point cloud data from laser scanners. For this, dimensions of the sections were measured on the point cloud data and these measurements were compared with the AISC database to find the best fitting sections. These best fitting sections were checked for accuracy using the as-designed BIM file.

The results of the identification showed that the best estimate was only accurate for 18.75% of the columns and 39.68% of the beams. For 31.25% of the columns and 22.22% of the beams, none of the estimates were the correct section.

We assessed the major challenges for recording measurements in point cloud data to be occlusions owing to the shape of the sections and unrelated objects like equipment, mixed pixels at the edges and noisy data. Particularly, large standard deviation of measurements compared to the section dimensions was the major reason behind the low identification rates. The results prove that modeling steel structures with the current equipment technology and processing workflows is a difficult and low accuracy task.

In this paper, we used a manual method for recording the section dimensions and selecting the sections. Additionally, no prior information was used to aid in the selection process.

This study aims at quantifying the accuracy of identification of steel sections from laser scan data. To the best knowledge of the authors', such a quantitative evaluation does not yet exist. The testbed and the results of the manual identification could be used as a baseline for comparison and evaluation of object recognition algorithms on this dataset for future studies. The results of this study points to the low accuracy of laser scanners in imaging small details.

Identification process could be improved using statistical or machine learning methods. Instead of treating all the measurements equally, statistical methods could be used to filter out measurements with lower confidence.

Additionally, prior information about the structure can greatly simplify the process. The search space could be constrained from several angles using such information. For example, the search could be performed among the sections that are known to be used in the structure.

ACKNOWLEDGEMENTS

This material is based upon work supported, in part, by the National Science Foundation under Grant No. 0856558 and by the Pennsylvania Infrastructure Technology Alliance (PITA). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies. We thank Robert Lipman for providing experimental data.

References

1. GSA, General Services Administration, "BIM Guide for 3D Imaging", http://www.gsa.gov/graphics/pbs/GSA_BIM_Guide_Series_03.pdf, Last accessed March 2012.
2. Anil, E.B., Akinci, B., Huber, D., "Representation requirements of as-is building information models generated from laser scanned point cloud data", *ISARC*, 2011.
3. Tang, P., Akinci, B., Huber, D., "Quantification of edge loss of laser scanned data at spatial discontinuities", *Automation in Construction*, Vol. 18(8), pp. 1070-1083, 2009
4. AISC, "Shapes database, V 14.0", http://www.aisc.org/content.aspx?id=2868, Last accessed March 2012.
5. Tang, P., Huber, D., Akinci, B., Lipman, R., Lytle, A., "Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques", Automation in Construction, Vol. 19(7), pp. 829-843, 2010.