Automated Cleaning of Point Clouds for Highway Retaining Wall Condition Assessment

Pedram Oskouie[1], Burcin Becerik-Gerber[2] and Lucio Soibelman[3]

[1] PhD Student, Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA; PH (213) 572-9373; Fax (213) 744-1426; email: oskouie@usc.edu

[2] Assistant Professor, Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA; PH (213) 740-4383; Fax (213) 744-1426; email: becerik@usc.edu

[3] Professor and Chair, Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA; PH (213) 740-0609; Fax (213) 744-1426; email: soibelman@usc.edu

ABSTRACT

Continuous condition monitoring and inspection of under-construction highway retaining walls is essential to ensure that construction performance criteria are met. The use of LIDAR systems by the construction industry has been significantly increased in recent years especially for recording the as-built and as-is conditions of facilities. The high-precision, dense 3-D point clouds generated by 3-D laser scanners can facilitate the process of Asset Condition Assessment (ACA). ACA involves preprocessing the point cloud data, for which point clouds have to be cleaned of any unwanted or occluding objects, and noises. As part of this research, the retaining wall point cloud data from an ongoing highway construction project was processed and analyzed. The project uses 3-D laser scanners for regular monitoring of mechanically stabilized earth walls that retain the soil supporting the highway alignment. The temporary steel and wooden brackets that hold formworks on top of the walls along with other construction materials are defined as unwanted objects. The authors have used a non-deterministic algorithm to remove the brackets and noises from the point clouds. Various settings of the algorithm have been analyzed using different sets of data. This paper presents the accuracy and performance of the tested algorithm and its evaluation when comparing the results with manually cleaned point clouds.

INTRODUCTION

With the advent of remote sensing technologies, 3-D laser scanners have been employed by the construction industry for several different uses such as 3-D as-built model reconstruction (Tang et al. 2010, Xiong et al., 2013), project control and progress monitoring (Turkan et al. 2012), MEP clash detection and as-built reconstruction (Tang et al. 2013, Bosche et al. 2013), 3-D thermal modeling (Wang et al. 2012), and infrastructure surveying and inspection (Tang et al. 2012, Olsen et al. 2012). Laser scanners enable rapid and accurate data collection from under-construction facilities, which makes them a very useful tool for quality control purposes. Hashash et al. 2008, used 3-D laser scanner data to monitor ground movements caused by excavation. The precise and non-contact measurement ability of the laser scanners has facilitated monitoring of infrastructure systems such as bridges, and dams (Tournas and Tsakiri, 2008). Recent research studies have also focused on applying image processing techniques and laser scanner technology for highway asset data collection and road construction (Golparvar et al., 2012, Gong et al. 2012). Olsen et al. 2012, have used laser scanners for real-time change detection of landslides and earth movements for highway retaining walls. While laser scanners can provide precise geometry data, the utilization of point clouds without adequate preprocessing may result in erroneous assessments. Data preprocessing often includes manual or automated filtering and removal of unwanted or occluding data (Tang et al. 2010). Occluding objects such as vegetation are usually filtered from the point cloud by identifying them based on their reflectivity or amplitude values (Schall et al., 2005). Other researchers have looked into semi-automated methods for detection and removal of ground objects such as rock masses from point clouds (Gigli et al. 2011). In general, each scan dataset may have its own particular undesired objects and therefore, available solutions may not be suitable for every point cloud data set. More importantly, the geometrical parameters and features of a scan target have to be considered before the data preprocessing begins. Preprocessing of under-construction facilities is particularly challenging since construction sites are dynamic and usually occupied with laborers, materials, and construction equipment, much of which frequently changes positions.

This paper focuses on point cloud cleaning and noise removal as an essential step of data preprocessing for Mechanically Stabilized Earth (MSE) walls. These walls are constructed with the purpose of retaining soils below the freeway alignment generally where right-of-way constraints preclude the use of a sloped embankment.

The authors have studied the under-construction MSE walls on California Interstate I-405 Sepulveda Pass Widening Project in Los Angeles, CA. The exterior of the wall consists of prefabricated concrete panels that are attached to reinforcing straps that extend into the reinforced soil mass below the freeway lanes.

The walls were scanned by the project surveying group on a regular basis. The scanned data was then processed and post-processed for vertical settlement and lateral movement over time during and following construction of the MSE walls. The final stage of the construction includes the installation of barrier moment slabs on top of the MSE wall. Brackets are installed on top panels of the wall to hold the formwork before concrete is poured. The pouring of concrete for the moment slab takes place in different phases and on different dates, therefore, the as-built point cloud of the walls often contains the brackets and other construction materials occluding the wall's surface.

In order to accurately analyze the wall's point cloud data and to avoid processing errors, the brackets and other occluding objects have to be manually trimmed out of the point cloud. This process is time consuming and labor intensive and often subject to human errors. Errors in trimming out portions of point clouds may result in loosing important data and consequently effecting the data computation during the post-processing stage. In order to improve point cloud data preprocessing, there is a need to automate the process of point cloud cleaning and noise removal to reduce human errors, and to decrease the time and cost associated with the manual process.

RANSAC Algorithm for Cleaning Point Clouds. Random sample consensus (RANSAC) (Fischler and Bolles, 1981) is a non-deterministic method for robustly fitting a model to a data set that is outlier-contaminated. The algorithm is widely used especially in the field of computer vision for robust image feature detection, matching, and visual motion estimation (Brown and Lowe, 2002). The RANSAC algorithm has been also used in the literature for detecting different shapes and objects within the 3-D point clouds (Schnabel et al. 2007). The advantage of the algorithm is its high tolerance to outliers within a data set (Fischler and Bolles, 1981).

The general steps that the algorithm goes through to fit the best model to the data set are: (1) Randomly selecting a subset of data; (2) Fitting a model to the data subset; (3) Classifying the data points as inliers (consensus set) and outliers based on the fitted model; (4) Repeating previous steps for ***N*** iterations; and (5) Selecting the largest set of inliers and re-estimation of the best model. The main variables that the algorithm works with are ***E*** (outlier ratio within the data set), ***p*** (probability of success in finding inliers only), ***N*** (number of iterations, which is calculated based on ***E*** and ***p*** values), and finally ***t*** (distance threshold for determining how close the datum should be to the model to be considered as an inlier). The number of required iterations depends on the value of ***p*** and ***E***. The higher values of ***p*** and ***E*** result in greater numbers of iterations.

This paper presents the results of deploying the RANSAC algorithm for removing noise and cleaning the MSE wall's point clouds. RANSAC was chosen over other existing algorithms because of its capability of fitting a model to a data set and more importantly, its ability to consistently identify unwanted objects which have varying geometric features and therefore cannot be detected and eliminated using point cloud filters. The authors have analyzed the geometric features of the point cloud and have evaluated different settings of the algorithm to achieve the best cleaning performance. The brackets, formworks, fence, and ground were recognized as outliers and the wall data points as inliers.

MSE WALL POINT CLOUD

Figure 1 is an image of the MSE wall along with its 3-D point cloud obtained from the laser scanner. The MSE walls along the I-405 highway vary significantly in terms of height, length, and other geometrical parameters such as wall's orientation and curvature. The prefabricated concrete panels of the walls create near-vertical and generally horizontal joints on the façade. The sets of vertically stacked concrete fascia panels along the wall are referred to as columns. Each column of panels is approximately 1.5m wide. The vertical ridges on the surface of the panels are created for acoustic and traffic noise control as well as an aesthetic treatment. This study uses an MSE wall with the following parameters: Length= *240m*, Height= 5m, Number of Columns= 158. The wall's point cloud contains *25,583,981* points.

METHODOLOGY

Point Cloud Segmentation. In order to evaluate the performance of the RANSAC algorithm, the authors have used different sizes of the wall's point cloud. The point cloud was manually segmented into multiple columns using the vertical joints on the wall. The point cloud cleaning and noise removal process was evaluated for point clouds with 1 to 15 columns.

Figure 1. MSE Wall's Point Cloud and Image

RANSAC Variables. For the purpose of this study, the authors have used a fixed value of 0.99 as the probability of success of data modeling for the algorithm to get the most accurate results possible. For the MSE walls, an outlier, as defined earlier in the paper, is any object that is not located on the surface of the wall. The ratio of outliers is automatically computed by the algorithm. The calculated average ratio of the outliers for the point cloud described above is *0.30*. The most important variable that needs to be carefully studied is the distance threshold value for inliers. The threshold value is generally subjective and dependent on the type of the data under study. As can be seen in Figure 1, the outliers have different Z values than points on the wall's surface considering the project Cartesian axes. The other factor that has to be considered before determining the threshold value is the geometry of the wall. Figure 2 shows a top view of the wall's point cloud. The face of the wall is made up of individual concrete fascia panels and therefore has a surface with variable indentations. The small indentations are fairly uniform and consist of the ridges on the individual concrete panels of the wall while the larger indentations are the vertical joints between the columns of panels along the wall alignment. Looking at the top view of the point cloud (Figure 2), if we take the centerline of the wall as the baseline for the best fitting model to the data (using planar model), we have to determine the right distance threshold value that eliminates the outliers while preserving the joints and data points on the wall surface. Subsequently, the algorithm was evaluated for five different thresholds: $t=0.01m$, $t=0.02m$, $t=0.03m$, $t=0.04m$, $t=0.05m$. Threshold values greater than 5cm were not considered since this depth generally encompasses the ridges in the concrete panels and the joints between columns.

Algorithm Evaluation. The algorithm was tested on different sections of the wall composed of 1 to 15 columns of panels for five different threshold values (total number of tests= $5 \times 15 = 75$). The runtime of the tests were recorded to determine the efficiency of the algorithm for different settings. In order to analyze the accuracy of the algorithm in eliminating the outliers and noises, the point clouds for different numbers of columns were manually cleaned and used as the ground truth. Guidelines were set for manual point cloud cleaning according the current industry practice to verify that the automatically cleaned point cloud is close. The guidelines were as follows:

- Trimming out the brackets using the side section view of the point cloud
- Trimming out the ground points and formwork using the top view of the point cloud
- Trimming out the wooden fence using the front view of the point cloud

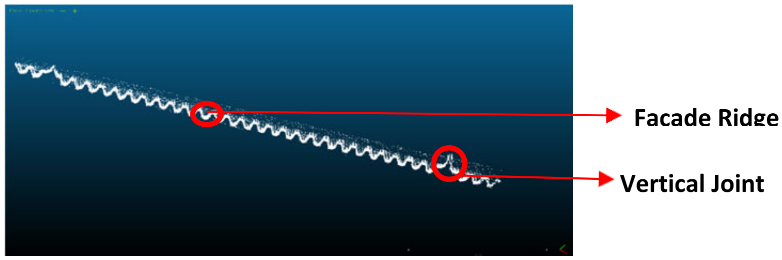

Figure 2. MSE wall's point cloud top view

Table 1 presents the duration of manual point cloud cleaning for point clouds with different numbers of columns.

Table 1. Time duration of manual cleaning process.

No. of Columns in the Point Cloud	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Cleaning Duration (minutes)	1:15	1:25	1:29	1:40	1:53	1:58	2:05	2:11	2:20	2:43	2:50	2:57	3:10	3:18	3:26

In order to compare the RANSAC results with the ground truth (manual cleaning), the two point clouds were compared using the nearest neighbor algorithm (Samet, 2005). The point clouds were associated with an octree data structure with level 9. The octree is a hierarchical data structure for representing spatial relationships of geometrical objects (Meagher, 1980). The octree level 9 was chosen by authors to improve computing speed for large size point clouds as well as to keep the computer memory consumption level down based on a 32-bit memory address. The comparison of two point clouds is based on a point-to-point comparison method, which finds the distance between nearest point neighbors within a voxel search. The manually cleaned point cloud represented the reference model with which, the results from RANSAC are compared. The points on both point clouds are then projected to a 3-D voxel map and the distances between the reference model points and the closest points from the compared model are calculated (Olsen et al. 2012). The low values for standard deviation of total existing distances would confirm the accuracy and act as indicators for algorithm's performance.

RESULTS AND DISCUSSION

Analysis. Figure 3 illustrates the raw point cloud and the automatically cleaned point cloud. The scattered points on the top of the wall in Figure 3a are the remaining parts of the fence. These points are not identified as outliers as they are close to the fitting plane defined by the algorithm. However, this is not a concern for pre-processing of point clouds since these outliers can be easily trimmed out manually in less than 10 seconds. The results indicate that the brackets, formwork and the ground were taken out from the point cloud and only shadows of them were left. Considering the fact that the remaining parts of the fence (areas not in the shadow of a bracket or other obstruction) would cause error during the point cloud comparison process, the authors trimmed them out manually which took less than 10 seconds. In order to further analyze what data points are considered as outliers, the authors studied the histograms of the X, Y and Z coordinates for the raw and cleaned point clouds. The histograms of the X and Y coordinates for both point clouds are very similar, however the histograms for the Z coordinate of the point clouds are quite different. The reason is that most of the outliers are located outside the wall's surface plane and therefore, they have different Z coordinates. Most of the outliers in this data set have Z coordinates ranging from 8 to 9 meters. According to the histograms, the Z points within the approximate interval of (8, 9) meters are considered as outliers and are eliminated by using the RANSAC algorithm.

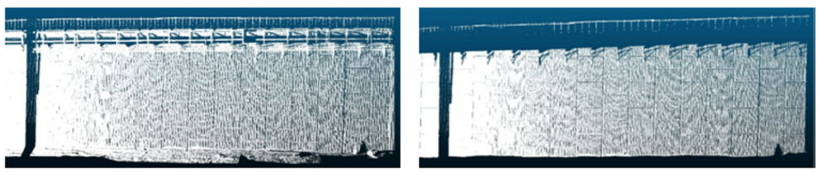

Figure 3. a) Raw point cloud b) RANSAC result

Determination of Optimal Distance Threshold and Number of Columns. Table 2 provides results from the comparison of the ground truth point clouds (manually cleaned point cloud) with ones cleaned by the RANSAC algorithm for different distance thresholds. The low standard deviation (sigma) values indicate that the two point clouds are quite similar. Figure 4 represents a visualization of the sigma changes for different number of columns (point cloud sizes) and threshold values. For all threshold values except t=0.01, the standard deviation follows the same trend.

All threshold graphs have minimum values when given one column and get maximum values for number of columns: 8, 11 and 15. This fluctuation is most probably due to the reason that the wall is not completely straight and is curved at some points. Defining best fitting plane on a curved wall is quite challenging and may result in losing some of the inlier data points. The authors plan to study the effects of a wall's curvature on automated point cleaning process in their future work. The minimum sigma among different distance thresholds occurred with **t=0.05** and it is equal to **0.002395** for one column. The second smallest sigma is for the same threshold for the point cloud with three columns. Additionally, based on Figure 4, the sigma value has remained small up to point clouds with three columns, therefore, for this type of wall, an acceptable accuracy can be achieved with the number of columns ranging from 1 to 3.

Run Time. The total time required for the algorithm to process a point cloud with one column of panels is 5 seconds for t=0.05, whereas the manual cleaning process takes about 75 seconds (Computer's configuration: Core i5 2.4GHz Processor, 8 GB RAM).

CONCLUSION

Post-processing of the 3-D point clouds for monitoring and inspection of highway retaining walls requires high quality and reliable data. Manual cleaning of 3-D point clouds necessitates spending excessive amounts of time and labor and it is often erroneous. The authors evaluated an automated method for point cloud cleaning

Table 2. Cloud comparison results (σ=standard deviation)

No. of Columns	σ(t=0.01)	σ(t=0.02)	σ(t=0.03)	σ(t=0.04)	σ(t=0.05)
1	0.013611	0.008252	0.003128	0.00332	0.002395
2	0.01324	0.007382	0.003146	0.003406	0.002445
3	0.012958	0.007201	0.003128	0.003533	0.002431
4	0.051138	0.008625	0.004944	0.023613	0.012649
5	0.06255	0.008201	0.005214	0.003377	0.003058
6	0.018934	0.008344	0.004897	0.004136	0.003883
7	0.048691	0.009753	0.006192	0.007608	0.009283
8	0.017559	0.041458	0.071614	0.042405	0.019559
9	0.019729	0.010937	0.01059	0.007608	0.0102
10	0.015678	0.009121	0.023722	0.004137	0.002768
11	0.02072	0.073305	0.036373	0.029171	0.005622
12	0.017962	0.015654	0.007845	0.004977	0.003376
13	0.016422	0.015769	0.010341	0.00577	0.00401
14	0.016422	0.015769	0.010341	0.00577	0.00401
15	0.026012	0.023002	0.125316	0.074829	0.019981
Min	0.012958	0.007201	0.003128	0.00332	**0.002395**
Max	0.06255	0.073305	0.125316	0.074829	0.019981

And noise removal. The results showed that the algorithm is promising in terms of removing outliers from the MSE walls' point clouds without sacrificing the inliers. This paper focused on automated point cloud cleaning for one type of MSE walls. The current findings are based on results from one wall.

Future work will include analysis of various types of MSE walls and evaluation of the algorithm's performance for walls with curvatures and changing orientations. In addition, the current methodology requires manual segmentation of the point clouds. The future research will investigate a methodology for automatically segmenting point clouds using the wall features. The goal of this research is to improve the quality of point clouds and more importantly to eliminate human errors by automating the point cloud data processing.

ACKNOWLEDGEMENTS

This material is based upon work supported by the Kiewit Corporation. 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 Kiewit Corporation. The authors would specifically like to acknowledge the support of Todd Shuey and Jason Peters of Kiewit Corporation and thank them for their time and efforts.

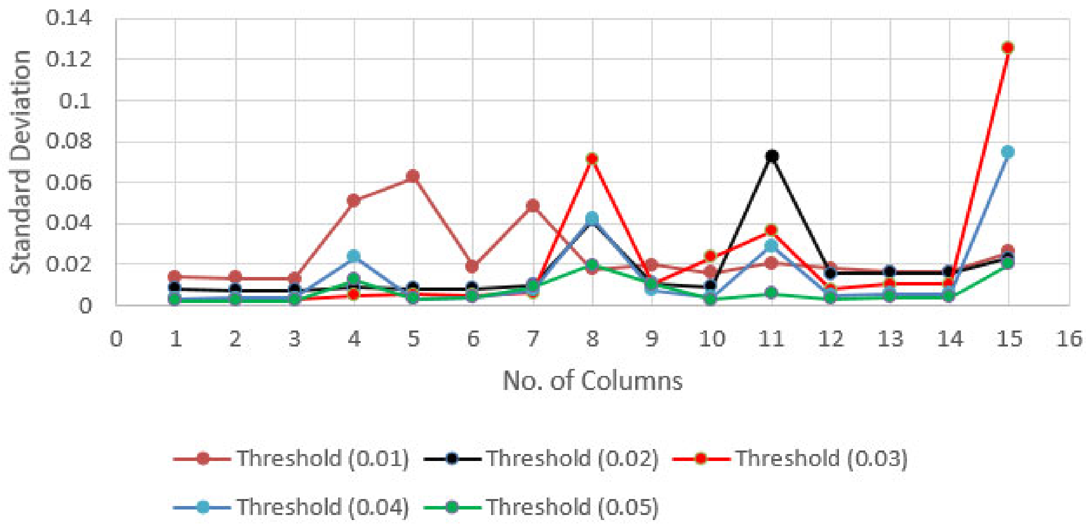

Figure 4. Results for different distance thresholds and number of columns

REFERENCES

Brown, M., and Lowe, D.G. (2002). "Invariant features from interest point groups." *British Machine Vision Conference*, Cardiff, Wales, 656-665.

Bosché, F., Guillemet, A., Turkan, Y., Haas, C., and Haas, R. (2013). "Tracking the Built Status of MEP Works: Assessing the Value of a Scan-vs-BIM System." *J. Comput. Civ. Eng.*

Fischler M. A., and Bolles R. C. (1981). "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography." *Commun.* ACM *24*, 381-395

Giovanni G., and Casagli N. (2011). "Semi-automatic extraction of rock mass structural data from high resolution LIDAR point clouds." *International Journal of Rock Mechanics and Mining Sciences*, Volume 48, Issue 2, 187-198, ISSN 1365-1609.

Golparvar-Fard, M., Balali, V., and de la Garza, J. (2012). "Segmentation and Recognition of Highway Assets Using Image-Based 3D Point Clouds and Semantic Texton Forests." *J. Comput. Civ. Eng.*

Gong, J., Zhou, H., Gordon, C., and Jalayer, M. (2012). "Mobile Terrestrial Laser Scanning for Highway Inventory Data Collection." *Computing in Civil Engineering*, 545-552

Hashash, Y. and Finno, R. (2008). "Development of New Integrated Tools for Predicting, Monitoring, and Controlling Ground Movements due to Excavations." *Pract. Period. Struct. Des. Constr.*, 13(1), 4–10.

Meagher D. (1980). "Octree Encoding: A New Technique for the Representation, Manipulation and Display of Arbitrary 3-D Objects by Computer." *Rensselaer Polytechnic Institute*.

Olsen M. J., Butcher S., and Silvia E. P. (2012) "Real-time Change and Damage Detection of Landslides and other Earth Movements Threatening Public Infrastructure" *Oregon Department of Transportation Research Section.*

Samet H. (2005). "Foundations of Multidimensional and Metric Data Structures." *Morgan Kaufmann Publishers Inc.*, ISBN 0123694469.

Schall O., Belyaev A., and Seidel H. P. (2005). "Robust filtering of noisy scattered point data." *Proceedings of the Second Eurographics / IEEE VGTC conference on Point-Based Graphics (SPBG'05)*, 71-77.

Schnabel, R., Wahl R., and Klein R. (2007). "Efficient RANSAC for Point-Cloud Shape Detection." *Computer Graphics Forum.* Vol. 26. No. 2. Blackwell Publishing Ltd.

Tang P., Huber D., Akinci B., Lipman R., and Lytle A. (2010). "Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques." *Automation in Construction*, Volume 19, Issue 7, 829-843.

Tang P., Akinci B. (2012). "Formalization of workflows for extracting bridge surveying goals from laser-scanned data." *Automation in Construction*, Volume 22, 306-319.

Tang P., Shen Z., and Ganapathy R. (2013). "Automated Spatial Change Analysis of Building Systems Using 3D Imagery Data." *Proceedings of the 30th CIB W78 International Conference*, Beijing, China.

Tournas E., and Tsakiri M. (2008). "Deformation Monitoring Based on Terrestrial Laser Scanner Point Cloud Registration", *13th Symposium on Deformation Measurement and Analysis*, Lisbon, Portugal

Turkan Y., Bosche F., Haas C. T., Haas R. (2012). "Automated progress tracking using 4D schedule and 3D sensing technologies." *Automation in Construction*, Volume 22, 414-421.

Wang, C., Cho, Y., and Gai, M. (2013). "As-Is 3D Thermal Modeling for Existing Building Envelopes Using a Hybrid LIDAR System." *J. Comput. Civ. Eng.*, 27(6), 645–656.

Xiong X., Adan A., Akinci B., Huber D. (2013). "Automatic creation of semantically rich 3D building models from laser scanner data." *Automation in Construction*, Volume 31