Chen, J., Fang, Y., and Cho, Y. (2017). "Performance Evaluation of 3D Descriptors for Object Recognition in Construction Applications." Automation in Construction, Volume 86, February 2018, Pages 44-52 doi.org/10.1016/j.autcon.2017.10.033

Performance Evaluation of 3D Descriptors for Object Recognition in Construction Applications

Jingdao Chen[1], Yihai Fang[2], Yong K. Cho[3*]

[1] Ph. D. Student, Institute for Robotics and Intelligent Machines, Georgia Institute of Technology, 801 Atlantic Dr. N.W., Atlanta, GA 30332, United States, E-Mail: jchen490@gatech.edu

[2] Postdoctoral Associate, Rinker School of Construction Management, University of Florida, Gainesville, FL 32611, United States, E-Mail: yihaifang@ufl.edu

[3]Associate Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, 790 Atlantic Dr. N.W., Atlanta, GA 30332-0355, United States, Email: yong.cho@ce.gatech.edu ([*]corresponding author)

ABSTRACT

3D object recognition from field-acquired point cloud data is important for modelling, manipulation, visualization and other post-processing tasks in the construction domain. However, building semantically-rich models from raw point cloud data is a difficult task due to the high volume of unstructured information as well as confounding factors such as noise and occlusion. Although there exist several computational recognition methods available, their performance robustness for construction applications are not well known. Therefore, this research aims to review and evaluate state-of-the-art descriptors for 3D object recognition from raw point clouds for construction applications such as workspace modelling, asset management and worker tracking. The evaluation was carried out using 3D CAD models with known labels as training data and laser-scanned point clouds from construction sites as testing data. The recognition performance was evaluated with respect to varying level of detail, noise level, degree of occlusion, and computation time. Experimental results show that for all evaluated descriptors, increasing the level of detail and decreasing the noise level results in a moderate increase in recognition accuracy whereas reducing occlusion results in a significant increase in recognition accuracy. In addition, experimental results suggest that the key features that distinguishes an object can be derived around the 10mm level and any further increase in the level of detail do not significantly increase the recognition accuracy.

Keywords: 3D object recognition; point cloud; machine learning; descriptor

1. Introduction

3D point clouds from laser scanning or photogrammetry are widely used to capture the as-built status of a construction jobsite. Point clouds usually consist of millions of points stored in an unstructured format; it is difficult for human agents and 3D modelling software to interpret and analyze the data. Automated object recognition techniques are important to recover contextual information and high-level semantics from raw point cloud data. The result of recognition tasks is relevant for applications such as safety monitoring [1,2], energy analysis [3], defect identification [4], inventory tracking [5], Building Information Modeling [6,7], and workspace modeling for equipment operation [8].

Using 3D CAD models as priors, Bosche and Haas [9] presented a semi-automated approach to recognize building objects (e.g., slabs, columns) from laser scanned point clouds. When the CAD models are extracted from a Building Information Model (BIM), the methods generally referred to “Scan-to-BIM” or “Scan-vs-BIM” [10,11]. By comparing the as-built models recognized from point cloud to the as-design BIM model, dimension discrepancy can be calculated and referred to the corresponding tolerances for construction quality control purpose [4,12]. Combining the 3D object recognition approach with schedule information, Turkan et al. [13] proposed a 4D object oriented construction progress tracking system for both permanent structural objects and secondary and temporary structures [14]. With a focus on energy simulation, Wang et al. [3] proposed an automated approach for extracting building geometries as individual objects and visualize the object as polygons from unorganized point cloud. Kim et al. used curvature information to perform automatic segmentation and 3D modeling of as-built pipelines [15–19]. In addition to building structural objects, recognizing other construction assets such as equipment, material, and foliage are of great value. To facilitate road safety inspection, Pu et al. [5] presented a recognition method that recognizes critical objects (e.g., traffic signs, barriers, trees) from mobile laser point clouds. To assist equipment operation in real-time, Cho and Gai [20] proposed a Projection–Recognition–Projection (PRP) method for rapid 3D modeling of construction equipment workspace using a pre-defined library of target objects.

Despite of the pressing demands of recognizing construction objects from point cloud data, existing approaches and tool sets cannot fully address the challenges such as low recognition rate with incomplete or noisy point cloud data and the need for supervision in the recognition process. One possible approach for object recognition from point clouds is to use descriptor-based recognition methods. A descriptor is a vector of features that allow shape retrieval or correspondence finding algorithms to uniquely identify a keypoint or an object. Using a machine learning framework, a novel object can be classified into a set of learned models based on its descriptors. 2D descriptors have been researched extensively in the computer vision field [21,22], yet the use of 3D descriptors for object recognition from point clouds has not been widely studied, especially in the civil engineering domain. The datasets used for studying 3D descriptors usually involve small objects obtained in a lab setting and there are few existing comparison studies that quantify the effectiveness of different descriptor-based recognition methods in complex, real-world environments. Therefore, this research aims to present a review and evaluation of state-of-the-art descriptors for 3D object recognition for construction applications. In this study, five different local and global 3D descriptors were evaluated on the object recognition task for laser-scanned point clouds obtained from construction sites. The five selected descriptors represent popular approaches to 3D object recognition available in the literature encompassing gradient-based and shape distribution-based approaches. The recognition performance was compared with respect to varying level of detail, noise level, and degree of occlusion. The rest of this paper will present, in order, the literature review, methodology, results, discussion, and conclusions.

2. Literature Review

The use of laser scan data for as-built jobsite modelling and planning has been well studied in the literature [5,23]. Despite the importance of automatically extracting CAD objects from point clouds, there is still a lack of research into descriptor-based recognition that can adequately handle problems in raw data acquired from a construction environment such as outliers, noise, and missing data [6]. This study aims to address this knowledge gap by conducting a performance evaluation of 3D descriptors for object recognition. 3D descriptors can be generally divided into two families, namely local descriptors and global descriptors. Local descriptors are more commonly used for instance recognition tasks whereas global descriptors are more commonly used for classification tasks [23].

Local descriptors characterize the neighborhood of a keypoint in terms of local features such as curvature and gradients. Similar to many 2D descriptors, the technique of histogram binning is used to discretize the feature space. A standard descriptor used in the literature is the spin image [24], where the set of points in the neighborhood of the keypoint of interest is mapped to an image-like grid. The spin image encapsulates the surface features for each keypoint of an object, thus a novel object can be recognized if its spin images can be matched to spin images in an existing model database. This idea is further expanded by the Signature of Histograms of Orientations (SHOT) [25], which builds a histogram of the gradients in a region of support around a keypoint instead of the point distances. A separate idea is to construct a spherical reference frame around the keypoint and accumulate neighboring points in logarithmicallyspaced subdivision. This technique is used in descriptors such as 3D Shape Context [26] and Unique Shape Context [27].

On the other hand, global descriptors summarize the geometry of an entire object in a single feature vector. Global descriptors are able to capture features relevant to the complete object geometry instead of a specific area, but require precise segmentation techniques to extract the object from the raw scan data. One approach to constructing a global descriptor is to calculate the distribution of geometric properties of points sampled throughout the object surface. For example, the Ensemble of Shape Functions (ESF) [28] uses the distributions of distances, areas, and angles to distinguish between different objects. A different approach is to measure the relative pan, tilt, and yaw angles between the point normals and a selected viewpoint direction, used by the Viewpoint Feature Histogram [29]. Alternatively, the Principal Axes Descriptor could be used, where an occupancy grid over the object points can be constructed based on principal component analysis, from which occupancy ratios can be derived and used as features [30].

For this study, a set of five descriptors were selected which represent popular approaches to 3D object recognition available in the literature: (i) spin images (SPIN), (ii) Signature of Histograms of Orientations (SHOT), (iii) Unique Shape Context (USC), (iv) Ensemble of Shape Functions (ESF), and (v) Principal Axes Descriptor (PAD). The five descriptors were selected to compare between a mix of different methods for deriving features from a 3D point cloud (e.g. local vs. global, spherical vs. Cartesian reference frame). A summary of the properties of the different local and global 3D descriptors is given in Table 1.

Table 1: Comparison of properties between different 3D descriptors in the literature

Descriptor	**SPIN** [24]	**SHOT** [25]	**USC** [27]	**ESF** [28]	**PAD** [30]
Region of support	Local	Local	Local	Global	Global
Feature dimension	153	361	1969	640	10
Uses histogram binning	Yes	Yes	Yes	Yes	No
Uses point normals	Yes	Yes	Yes	No	No
Reference frame	Cylindrical	Spherical	Spherical	None	Cartesian

Several works in the literature have attempted to identify the advantages and disadvantages of different 3D descriptors. Campbell and Flynn [31] surveyed different recognition systems for free-form objects in relation to the 3D model building process. Bronstein and Bronstein [32] also examined different feature descriptors for the task of 3D shape retrieval and correspondence finding. Guo et al. [33] presented a survey of local surface features and identified weaknesses such as sensitivity to occlusion, deformation, and point density variation. In terms of experimental work, Salti et al. [34] performed an experimental evaluation for 3D object recognition in terms of robustness to noise, clutter, occlusion, and viewpoint variation. Similarly, Arbeiter et al. [35] evaluated the performance of three local descriptors for different primitive surfaces such as cylinders, edges, and corners. Most of these works concentrate on local descriptors and did not compare their performance to that of global descriptors for object recognition. In addition, there has not yet been any comparison studies using laser-scanned data collected from the field that is subject to occlusion and sensor noise in a real-world setting.

3. Methodology

On overview of the overall experimental design is shown in Figure 1. Detailed explanations of the steps for data acquisition and pre-processing, object recognition framework, and varying the experimental parameters will be presented in the following subsections.

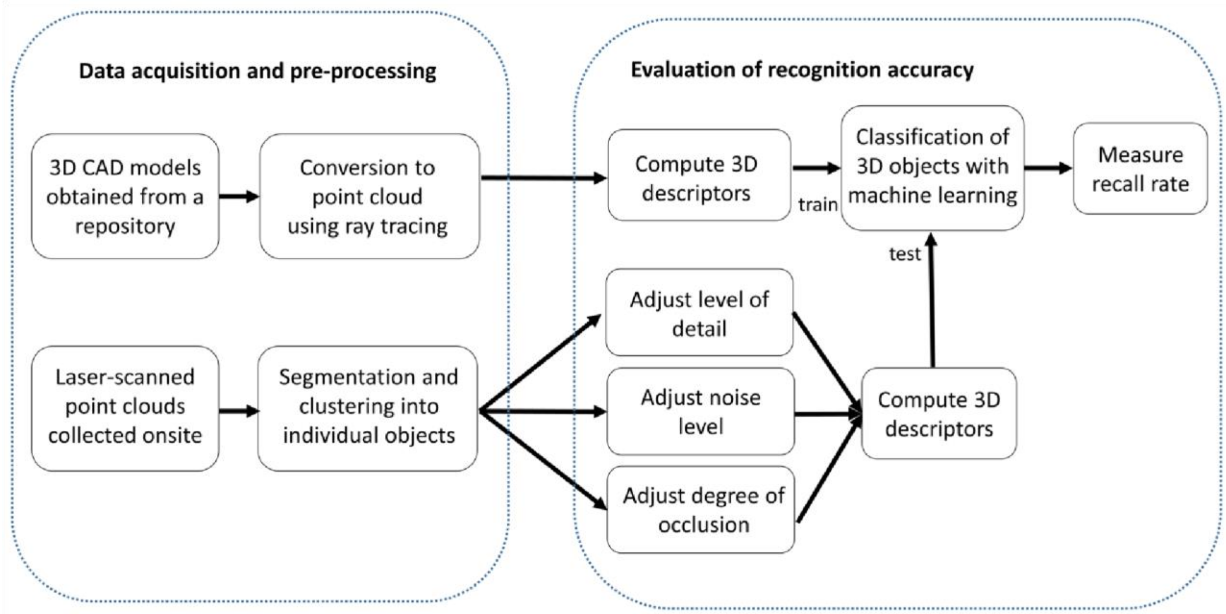


Figure 1: Flowchart of the experimental design to evaluate object recognition performance

3.1 Data acquisition and pre-processing

The 3D object dataset for this study consists of a training dataset, which contains objects with known labels and used to train machine learning classifiers, and a testing dataset, which is used to evaluate the recognition performance of descriptors. The dataset consists of five classes of objects commonly encountered on a construction site, namely (i) trailer, (ii) truck, (iii) worker, (iv) crane, and (v) excavator. The training dataset, shown in Figure 2, is acquired from 3D CAD models downloaded from online repositories. The 3D models in mesh format are converted to a point cloud format using a ray tracing algorithm which works by placing virtual laser scanners around the object. The virtual laser scanner projects a series of rays outwards from the scan origin and the first intersection between each ray and the object surface is recorded as a point sample. This process is repeated to incrementally build a point cloud consisting of points along the object surface. For each object category, virtual laser scans were captured from 8 different view angles for each of 5 CAD models to simulate intra-class variance for that object category.

In total, 200 synthetic point clouds were generated to build the training dataset.

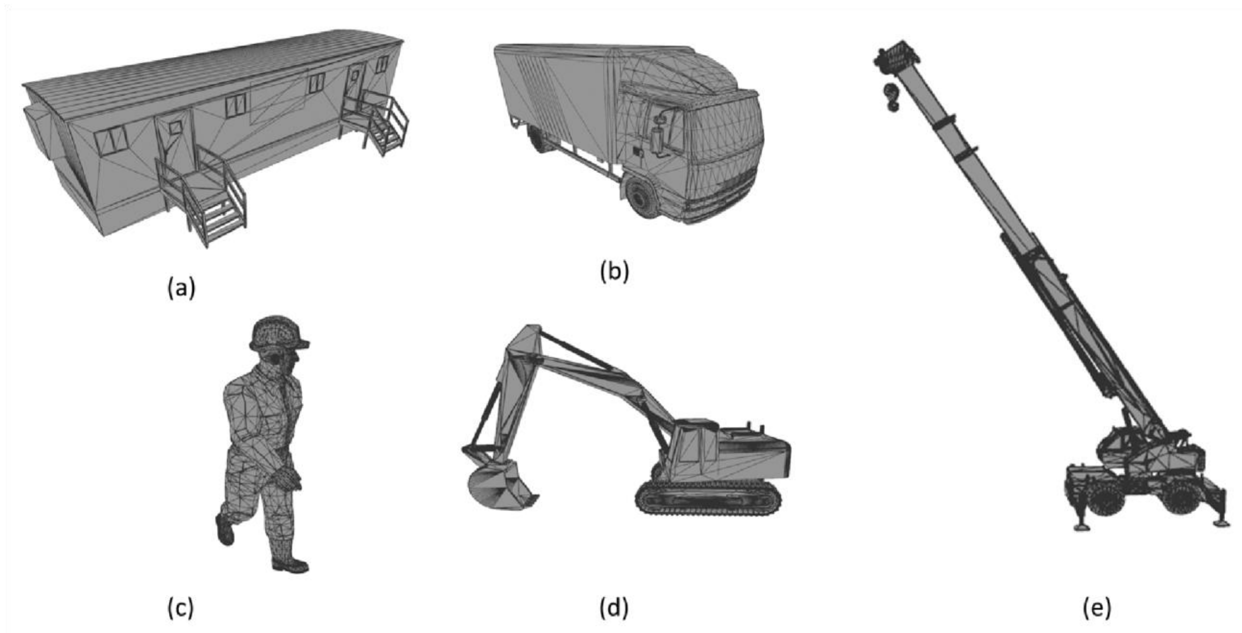


Figure 2: Example training data from 3D CAD models: (a) trailer, (b) truck, (c) worker, (d) excavator, (e) crane

On the other hand, the testing dataset is collected from multiple laser-scanned point clouds from crane yards and construction equipment yards. The raw scan data is pre-processed with a segmentation algorithm and the individual objects corresponding to the classes of interest are extracted as shown in Figure 3. Since the individual objects are laid out in an open area, a simple algorithm of filtering out the ground plane and clustering the remaining points based on Euclidean distance suffices to separate out each object. For cases where the test site is complex and cluttered, a more sophisticated segmentation algorithm utilizing surface roughness and curvatures may be used [36]. Next, the point clouds of individual objects are organized by class according to the predefined categories in the training dataset (Figure 4). The total number of point clouds collected in the testing dataset is 54.

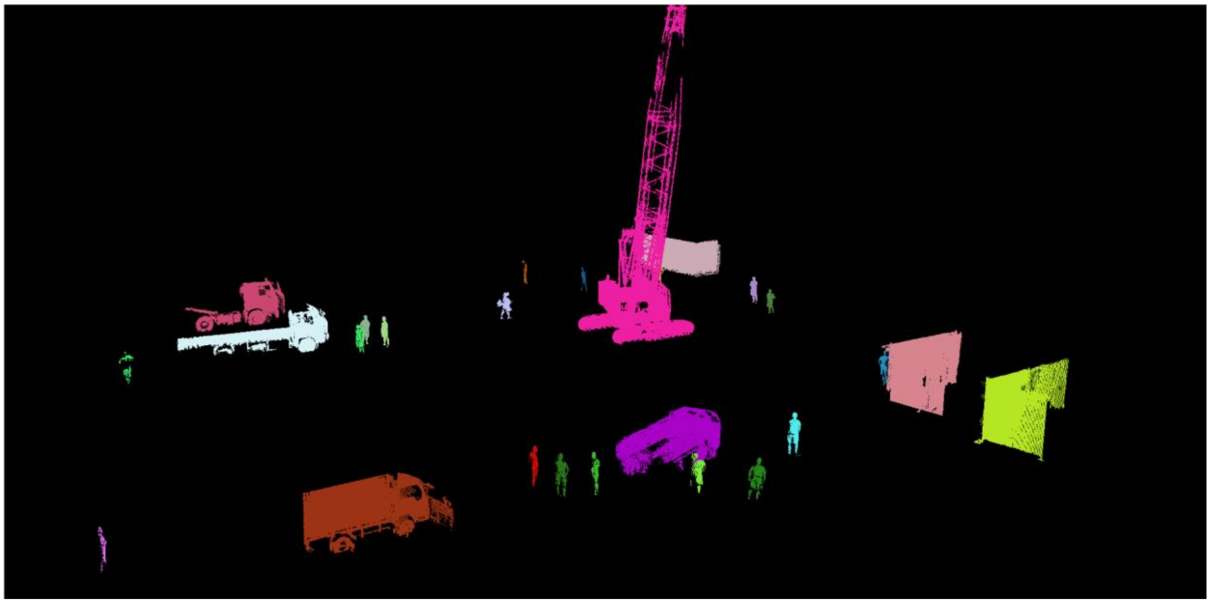
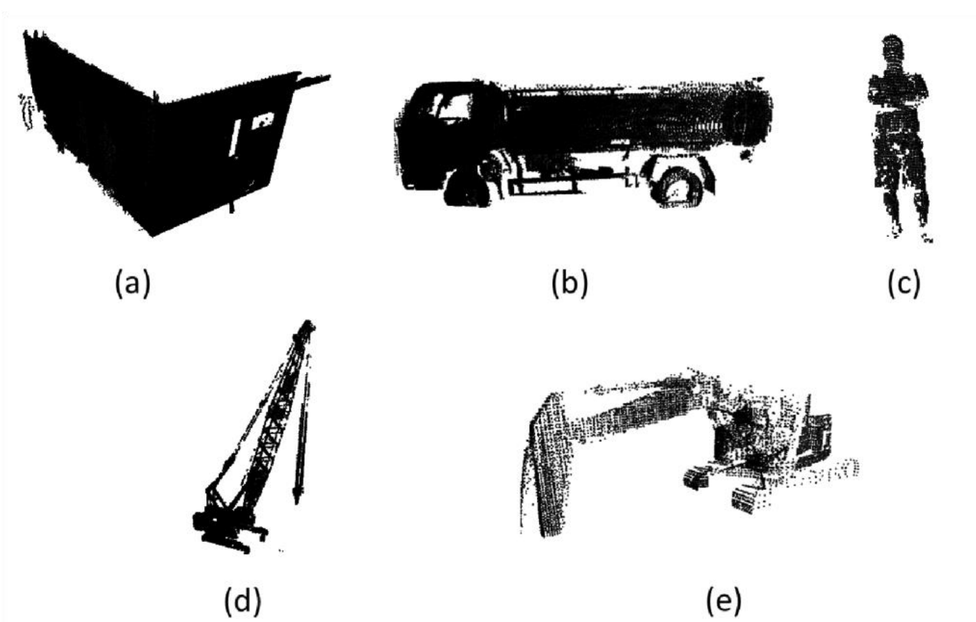


Figure 3: Test point clouds for recognition extracted from a crane yard



Chen, J., Fang, Y., and Cho, Y. (2017). "Performance Evaluation of 3D Descriptors for Object Recognition in Construction Applications." Automation in Construction, Volume 86, February 2018, Pages 44-52 doi.org/10.1016/j.autcon.2017.10.033

Figure 4: Example testing data from laser-scanned point clouds: (a) trailer, (b) truck, (c) worker, (d) crane, (e) excavator

3.2 Object recognition framework

The object recognition framework used in this study is the conventional machine learning framework with a training dataset and a testing dataset as described in the previous section. The recognition framework for local descriptors and global descriptors is different due to the nature of descriptor calculation. For local descriptors, a set of keypoints is randomly sampled from point clouds in the input dataset. A descriptor vector is calculated for each keypoint; thus, a single object will have multiple descriptors. Classification of a test object is performed using a nearest-neighbor method adapted from [37] which was originally used for image classification. The classification method works as follows: (i) first, calculate the Euclidean distance between each descriptor, d_i of the test object and the nearest-neighbor descriptor, $NN(d_i)$, for each category in the training dataset, (ii) next, sum up the distances according to the categories, C of the training data, (iii) finally, classify the test object as the category, $\hat{C}$, that minimizes the distance sum over all descriptors. This relationship can be expressed with Equation 1 as shown below:

$$\hat{C} = \arg\min_{C} \sum_{i=1}^{n} \| d_i - NN_C(d_i) \|^2 \quad \text{(Equation 1)}$$

For global descriptors, only a single descriptor vector is calculated for each object in the training and testing datasets since global descriptors are defined over the entire object. The descriptor vectors together with the class labels of the training dataset are used to train Support

Chen, J., Fang, Y., and Cho, Y. (2017). "Performance Evaluation of 3D Descriptors for Object Recognition in Construction Applications." Automation in Construction, Volume 86, February 2018, Pages 44-52 doi.org/10.1016/j.autcon.2017.10.033

Vector Machine (SVM) classifiers [38]. Multiple SVM classifiers are learned: one corresponding to each class. For each test object, the corresponding descriptor vector is passed as input to the learned SVM classifiers and the class label is determined as the category which maximizes the confidence value of the SVM decision function. This relationship is shown in Equation 2, where w_c and b_c are the weight and bias parameters for the SVM classifier of class C.

$$\hat{C} = \arg\max_{c} \left(w_c d + b_c \right) \qquad \text{(Equation 2)}$$

3.3 Experimental parameters

Several experimental parameters were varied to measure the robustness of the evaluated 3D descriptors in diverse scenarios. The varied parameters were designed to represent the complexity of data that is likely to be encountered in the construction field. The parameters considered in this study are explained in detail in the following subsections.

3.3.1 Level of Detail (LOD)

One of the most important parameters in laser-scanned point cloud data is the level of detail. The level of detail, defined as the resolution of (x,y,z) coordinates in a given region of the point cloud, is the primary factor that affects the scanning time and processing time of a laser-scanned point cloud. Point clouds with a higher level of detail are preferred for visualization and modelling purposes, but usually incur longer data collection time and higher computational costs as well. The level of detail parameter is especially relevant for descriptor-based recognition tasks. A higher level of detail increases the precision of point geometry-derived features available to the recognition system, which potentially increases the accuracy. On the other hand, collecting a point cloud of an object with a higher level of detail necessitates an increase in the number of scan points, and this increases the computation time for descriptor calculation. In this study, the level of detail is varied by using a voxel grid to uniformly downsample the original point cloud. The level of detail parameter is then represented by the size, in mm, of each grid cell.

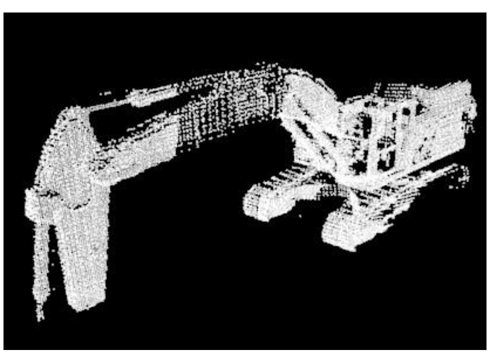
3.3.2 Noise Level

Another important parameter to be considered in point cloud data is the noise level. In the context of point clouds, noise is defined as small deviations in point location or extraneous points that distort the object geometry. Noise can be introduced to point cloud data in the form of (i) sensor noise, where different sensors have different noise characteristics [23], (ii) registration errors, which are caused by inaccuracies when combining multiple point clouds together, and (iii) motion distortion, which occurs when the position or orientation of the scanned object shifts in the middle of the scanning process. The inherent noise level in an acquired point cloud is difficult to measure when the ground truth geometry is unknown, thus in this study, noise is introduced to the dataset in the form of synthetic Gaussian white noise. Each point in the point cloud is perturbed by a small random vector sampled from a Gaussian distribution. The noise level is then represented by the standard deviation, in mm, of the Gaussian distribution.

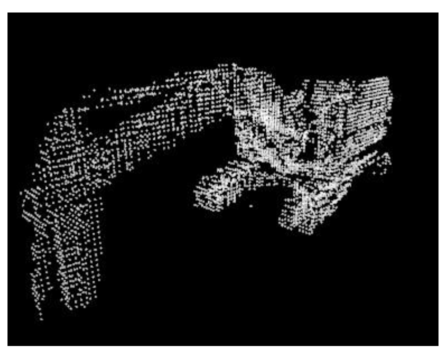
3.3.3 Degree of Occlusion

In a complex and dynamic environment such as a construction site, it is often difficult to acquire a complete point cloud of a target object due to the presence of occlusion. In situations where workers and equipment are moving around the site, multiple objects could potentially interrupt the line of sight between the laser scanner and the target object. This problem is compounded when it is physically difficult to access certain areas of the scanned site with a laser scanner. As a result, most objects in an acquired point cloud are only partially scanned, and usually only from a single direction. Occlusions have an adverse effect on descriptor-based object recognition since the available geometric information of an object is incomplete. In this study, the descriptors are evaluated with respect to test objects divided into three different occlusion scenarios: (i) no occlusion, (ii) partial occlusion, and (iii) heavy occlusion.

Figure 5 below shows example point clouds of an excavator that are modified by adjusting the level of detail, noise level and degree of occlusion parameters. Note that in contrast to level of detail and noise level which are measured numerically, the degree of occlusion is measured quantitatively by assigning a label of either (i) *no occlusion*, where the entire object is captured in the point cloud, (ii) *partial occlusion*, where one part of the object is missing from the point cloud, and (iii) *heavy occlusion*, where multiple parts of the object are missing from the point cloud.

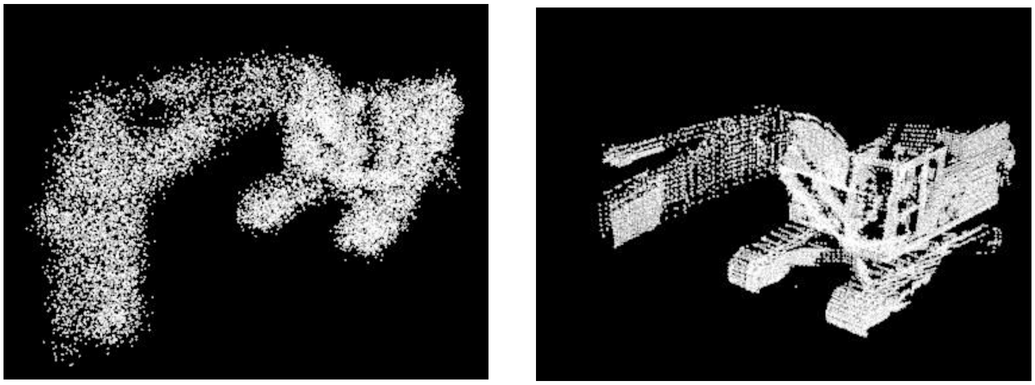


(a)



(b)

Chen, J., Fang, Y., and Cho, Y. (2017). "Performance Evaluation of 3D Descriptors for Object Recognition in Construction Applications." Automation in Construction, Volume 86, February 2018, Pages 44-52 doi.org/10.1016/j.autcon.2017.10.033



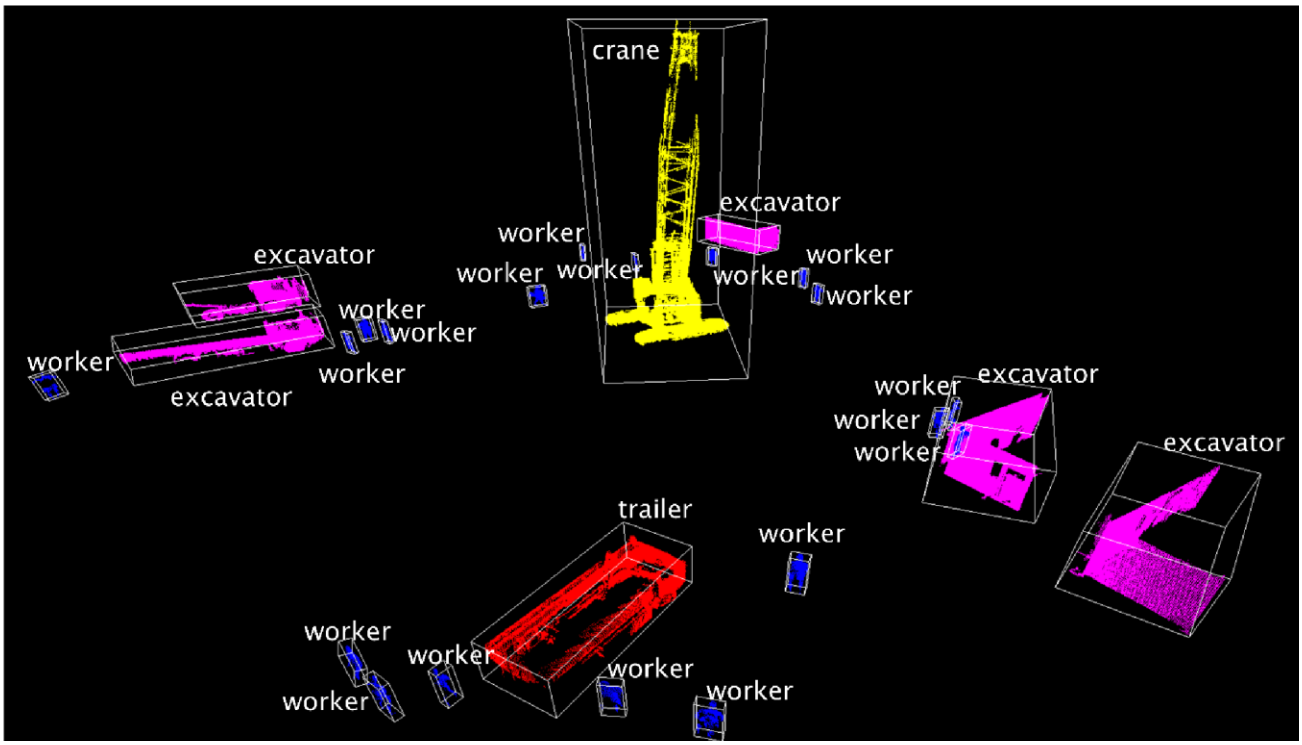
(c) (d)

Figure 5: Modified point clouds of an excavator with: (a) original parameters, (b) decreased level of detail, (c) increased noise level, and (d) increased occlusion

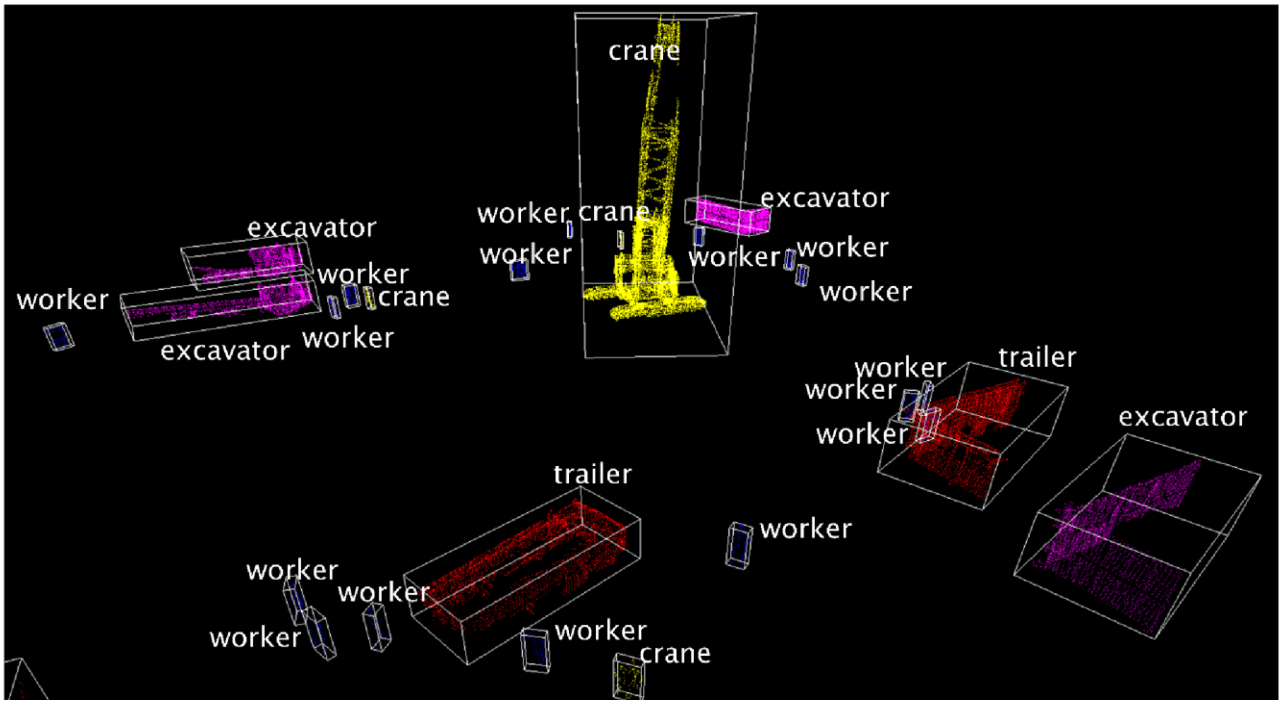
In addition to evaluating the classification performance with varying parameters, a separate experiment was also carried out to measure the relationship between computation time and size of the point cloud for each descriptor. This is important for real-time recognition applications involving large-scale point cloud data streams, in which it is critical to minimize the computational cost of calculating the 3D descriptors.

4. Results and Analyses

The five 3D descriptors evaluated are annotated as follows: (i) spin images (SPIN), (ii) Signature of Histograms of Orientations (SHOT), (iii) Unique Shape Context (USC), (iv) Ensemble of Shape Functions (ESF), and (v) Principal Axes Descriptor (PAD). The first three evaluated descriptors are local descriptors whereas the latter two are global descriptors. The descriptors were evaluated with respect to recognition accuracy measured by the recall rate, which is the ratio between the number of correctly classified objects and the total number of objects, in three experimental scenarios: (i) varying level of detail, (ii) varying noise level, and (iii) varying degree of occlusion. Qualitative results of the recognition results are shown in Figures 6a and 6b, which depicts the test site in two different experimental settings. Class labels for each object were predicted using the methodology described in Section 3.2.



(a)



(b)

Figure 6: Object recognition results for test point clouds with (a) 1mm level of detail and (b) 100mm level of detail

The quantitative experimental results are summarized in Figures 7-8, where the accuracy values are expressed as ratios. The results in Figure 7 show that for most descriptors, the recognition accuracy gradually increases as the level of detail increases (smaller point-to-point distance). Out of the five evaluated descriptors, ESF showed the highest sensitivity to varying level of detail whereas SHOT showed the lowest sensitivity to varying level of detail.

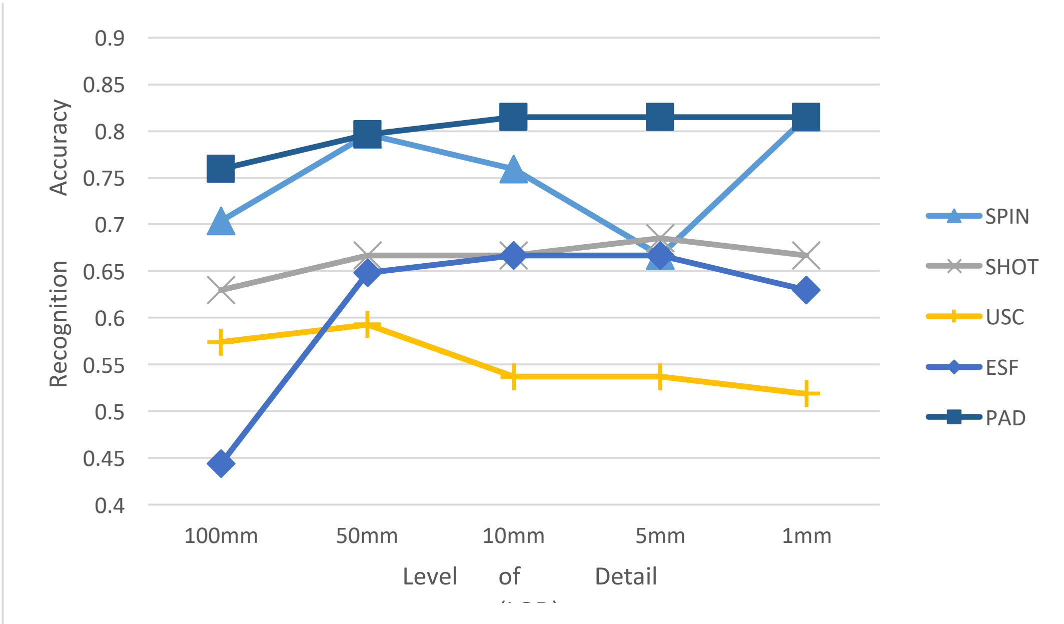


Figure 7: Plot of recognition accuracy with varying level of detail

A similar situation occurs when the noise level is decreased (Figure 8). Most descriptors show a gradual increase in recognition accuracy with ESF having the most improvement and SHOT having the least improvement. In certain cases, such as around the 1mm-5mm level for SPIN and 10mm-50mm level for USC, the recognition accuracy does not improve with lower noise. Overall, PAD performed the best in Figures 7 and 8 since it was originally studied in the context of construction equipment and is able to exploit known geometric structure. Additionally, its low dimensionality offers better regularization against changing experimental parameters.

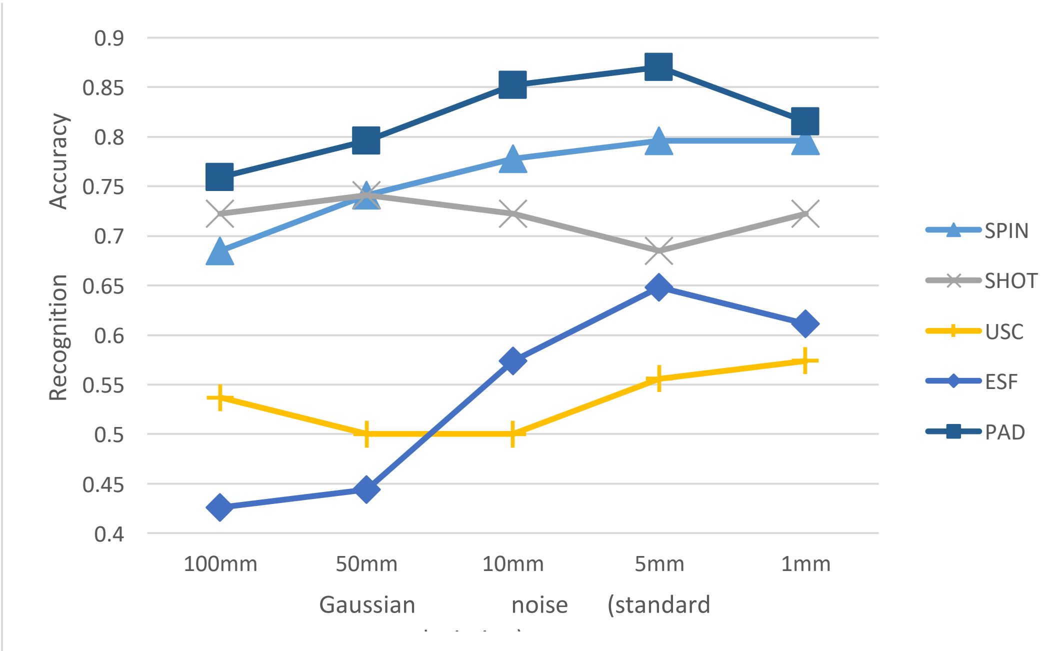


Figure 8: Plot of recognition accuracy with noise level

In terms of degree of occlusion (Table 2), all of the evaluated descriptors show a similar trend where the recognition accuracy significantly increases going from heavy occlusion to partial occlusion and slightly increases going from partial occlusion to no occlusion. This suggests that for both local and global descriptors, a partially occluded object still retains sufficient geometric properties to produce an accurate classification whereas a heavily occluded object is much more difficult to identify due to distorted geometry.

Table 2: Recognition accuracy (recall rate) with varying degree of occlusion

Degree of occlusion	**SPIN**	**SHOT**	**USC**	**ESF**	**PAD**
Heavy Occlusion	0.571	0.500	0.214	0.429	0.500
Partial Occlusion	0.885	0.769	0.654	0.731	0.923
No occlusion	1.000	0.786	0.786	0.714	0.929

The next experimental comparison involves the computation time of calculating a descriptor with respect to the number of points contained in the point cloud. As shown in Figure 9, the scaling of computation time as a function of the number of points is approximately linear except for the cases where the total number of point is small, in which the computation time is dominated by constant factors such as input-output operations and data preprocessing routines. The results showed that USC consumed the most computation time whereas PAD consumed the least computation time.

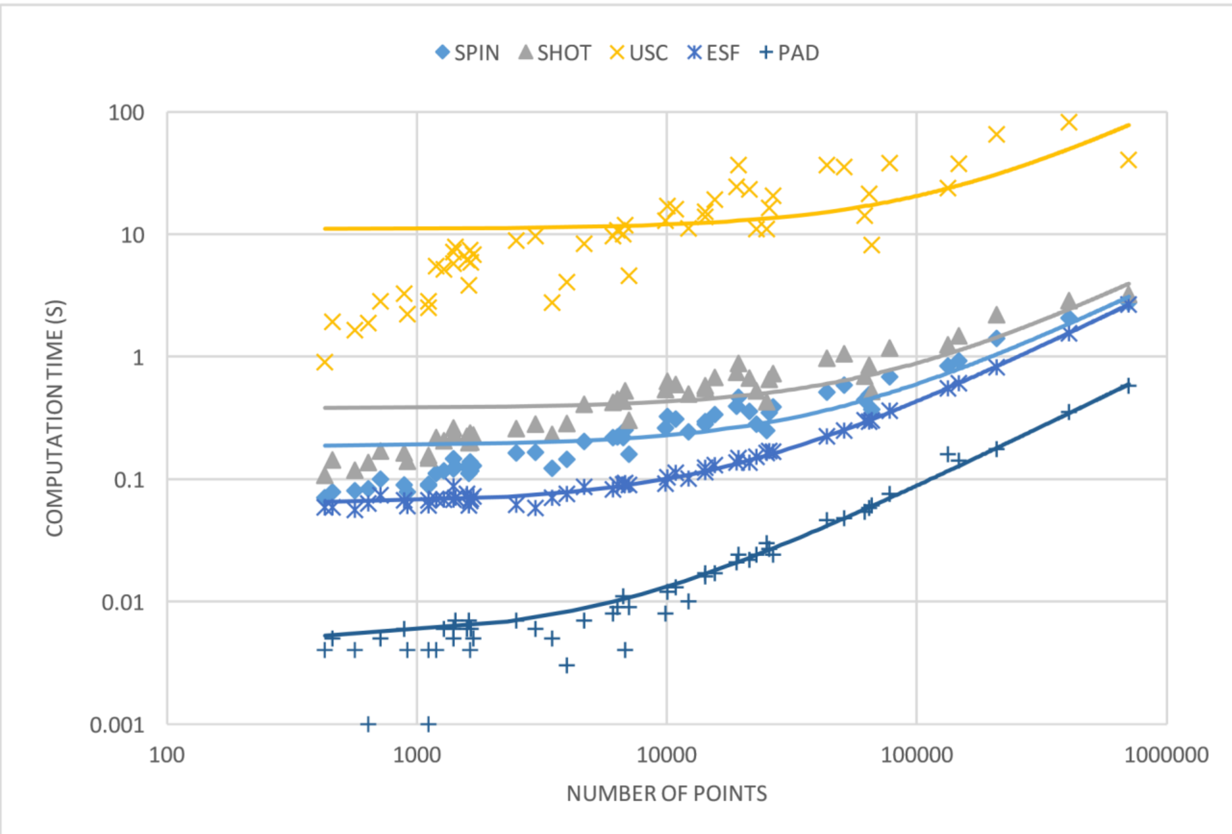


Figure 9: Graph of computation time against number of points for each descriptor

A summary of the comparison test results is given in Table 3. For each experimental parameter measure (level of detail, noise level, degree of occlusion, computational time), the table evaluates the performance relatively as either *good*, *bad*, or *moderate* for each of the 3D descriptors among the five evaluated descriptors. The performance is measured for both undesirable experimental conditions (e.g. high noise level) as well as desirable experimental conditions (e.g. low degree of occlusion).

Table 3: Summary of 3D descriptor performance under different parameter settings

Experimental Parameter Setting	**Object Recognition Performance**				
	SPIN	**SHOT**	**USC**	**ESF**	**PAD**
Low level of detail	Moderate	Moderate	Moderate	Bad	Good
High level of detail	Good	Moderate	Bad	Moderate	Good
Low noise level	Moderate	Moderate	Bad	Moderate	Good
High noise level	Moderate	Moderate	Moderate	Bad	Good
Low degree of occlusion	Good	Moderate	Moderate	Bad	Moderate
High degree of occlusion	Good	Moderate	Bad	Moderate	Moderate
Computational Time	Moderate	Moderate	Bad	Moderate	Good

To verify that the sample sizes used in this study is statistically meaningful, an experiment was conducted to examine the effect of number of training data and number of test data on the classification accuracy. First, the number of samples used for training data was varied from 10% to 100% of the full dataset while the number of test samples was kept constant. Next, the number of samples used for test data was varied from 10% to 100% of the full dataset while the number of training samples was kept constant. The resulting accuracies, shown in Figure 10, indicate that the classification performance shows high variability for low data sizes, but start to stabilize as the data size is increased. Thus, it can be inferred that even if the number of training and test data are further increased, the trend in classification accuracy will be similar.

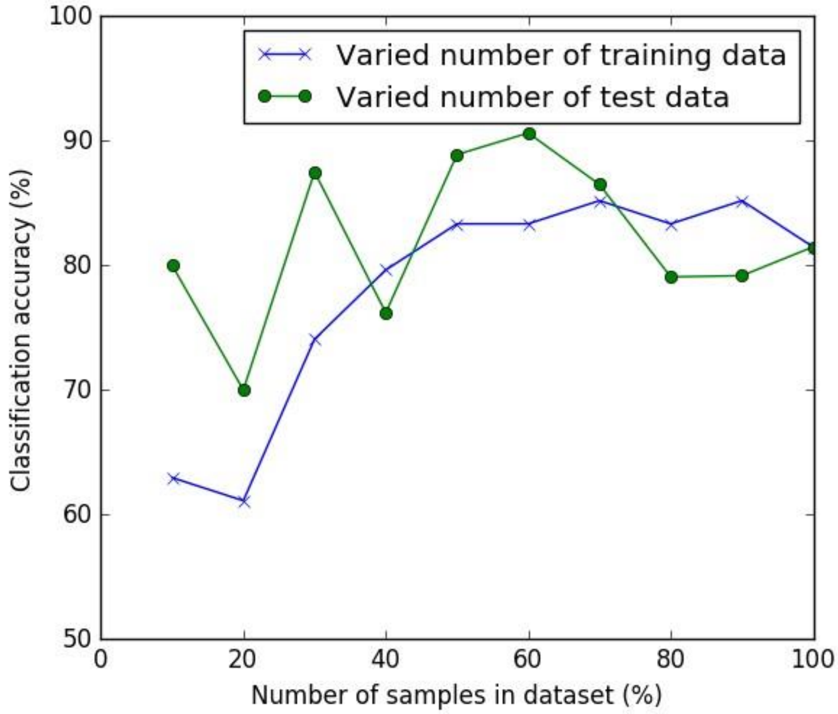


Figure 10: Classification accuracy with varying number of training and test data

5. Discussion

The analyses for this study will mostly be concerned with the trend in recognition accuracy as different parameters in the experimental data are changed and not the absolute accuracy values. This is because that most descriptor calculation routines involve hyperparameters such as normal estimation radius and radius for neighborhood search that have to be optimized for a particular dataset in order to achieve high accuracies. In addition, the amount of training and testing data available in this study is small compared to other machine learning domains such as image classification so the absolute accuracies may not be reflective of the true recognition performance in a general setting.

From the results for the effect of level of detail on recognition accuracy as shown in Figure 7, a general trend that can be observed is that the recognition accuracy tends to plateau around the 10mm level of detail. This suggests that the key features that distinguishes an object can be derived around the 10mm level and any further increase in the level of detail do not significantly increase the recognition accuracy. Thus for practical applications, a high level of detail of 1mm may be initially used to achieve accurate registration but the point cloud may be downsampled to around 10mm for the object recognition stage to save processing time.

The effect of noise level on recognition accuracy showed a general trend of a slight increase in accuracy as the noise level decreases as shown in Figure 8. However, this effect is inconsistent and there may be several reasons why the recognition performance does not continue to improve while the noise level is decreased under certain conditions. First, some descriptors use a histogram binning technique which is designed to diminish the effect of noise. Once a data point is assigned to a bin only the number of points in the bin is taken into account when calculating the descriptor and not the absolute coordinates of the data point. Second, the form of Gaussian noise used in this study is only applied to individual points at a time so the effect evens out once a larger region, whether it is a local patch for local descriptors or the entire point cloud for global descriptors, is taken into account.

Out of all the varied parameters, the degree of occlusion is shown to have the largest effect on recognition accuracy. Under the condition of heavy occlusion, most descriptors only achieved a recognition accuracy of 50% or lower. However, the recognition performance improves greatly if the objects are partially occluded or not occluded. There is not a large difference between the results for partial occlusion and no occlusion, possibly because the 3D descriptors have enough built-in redundancy to classify objects based on just the visible parts in spite of non-visible parts in the point cloud. Overall, this result suggests that for practical applications, the laser scanning process should concentrate on obtaining complete views for objects of interest to increase the likelihood of detection by automatic object recognition methods.

6. Conclusion

This study evaluated the recognition performance of five local and global 3D descriptors on a dataset consisting of objects commonly found in the construction field such as trailers, trucks, workers, cranes, and excavators. The descriptors were evaluated with respect to varying experimental parameters such as level of detail, noise level as well as degree of occlusion. Experimental results show that for both local and global descriptors, increasing the level of detail and decreasing the noise level results in a moderate increase in recognition accuracy. In contrast, reducing occlusion results in a significant increase in recognition accuracy. Thus, occlusion is a major parameter among others which significantly impacts object recognition results. The findings from this study are expected to benefit researchers and field practitioners by serving as a guideline for which object recognition technology to use given the properties of the point cloud data collected. For future work in this research, additional experimental parameters such as number of training samples, number of object classes, and level of intra-class variation will be investigated.

ACKNOWLEDGEMENT

This material is based upon work supported by the National Science Foundation (Award #:

Chen, J., Fang, Y., and Cho, Y. (2017). "Performance Evaluation of 3D Descriptors for Object Recognition in Construction Applications." Automation in Construction, Volume 86, February 2018, Pages 44-52 doi.org/10.1016/j.autcon.2017.10.033

CMMI-1358176) and the US Department of Defense (DOD) (Award No. FA2386-17-1-4655). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF and DOD.

7. Reference

[1] Y. Fang, Y.K. Cho, J. Chen, A framework for real-time pro-active safety assistance for mobile crane lifting operations, Autom. Constr. (2016). doi:10.1016/j.autcon.2016.08.025.

[2] S. Han, S. Lee, A vision-based motion capture and recognition framework for behaviorbased safety management, Autom. Constr. 35 (2013) 131–141. doi:10.1016/j.autcon.2013.05.001.

[3] C. Wang, Y.K. Cho, C. Kim, Automatic BIM component extraction from point clouds of existing buildings for sustainability applications, Autom. Constr. 56 (2015) 1–13. doi:10.1016/j.autcon.2015.04.001.

[4] E.B. Anil, P. Tang, B. Akinci, D. Huber, Deviation analysis method for the assessment of the quality of the as-is Building Information Models generated from point cloud data, Autom. Constr. 35 (2013) 507–516. doi:10.1016/j.autcon.2013.06.003.

[5] S. Pu, M. Rutzinger, G. Vosselman, S. Oude Elberink, Recognizing basic structures from mobile laser scanning data for road inventory studies, ISPRS J. Photogramm. Remote Sens. 66 (2011) S28–S39. doi:10.1016/j.isprsjprs.2011.08.006.

[6] V. Pătrăucean, I. Armeni, M. Nahangi, J. Yeung, I. Brilakis, C. Haas, State of research in automatic as-built modelling, Adv. Eng. Informatics. 29 (2015) 162–171. doi:10.1016/j.aei.2015.01.001.

Chen, J., Fang, Y., and Cho, Y. (2017). "Performance Evaluation of 3D Descriptors for Object Recognition in Construction Applications." Automation in Construction, Volume 86, February 2018, Pages 44-52 doi.org/10.1016/j.autcon.2017.10.033

[7] R. Volk, J. Stengel, F. Schultmann, Building Information Modeling (BIM) for existing buildings - Literature review and future needs, Autom. Constr. 38 (2014) 109–127. doi:10.1016/j.autcon.2013.10.023.

[8] Y. Cho, C. Wang, M. Gai, J.W. Park, Rapid Dynamic Target Surface Modeling for Crane Operation Using Hybrid LADAR System, in: Constr. Res. Congr. 2014, ASCE, 2014: pp. 1053–1062.

[9] F. Bosche, C.T. Haas, Automated retrieval of 3D CAD model objects in construction range images, Autom. Constr. 17 (2008) 499–512. doi:10.1016/j.autcon.2007.09.001.

[10] X. Xiong, A. Adan, B. Akinci, D. Huber, Automatic creation of semantically rich 3D building models from laser scanner data, Autom. Constr. 31 (2013) 325–337. doi:10.1016/j.autcon.2012.10.006.

[11] F. Bosché, M. Ahmed, Y. Turkan, C.T. Haas, R. Haas, The value of integrating Scan-toBIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components, Autom. Constr. 49 (2015) 201–213. doi:10.1016/j.autcon.2014.05.014.

[12] F. Bosche, E. Guenet, Automating surface flatness control using terrestrial laser scanning and building information models, Autom. Constr. 44 (2014) 212–226. doi:10.1016/j.autcon.2014.03.028.

[13] Y. Turkan, F. Bosche, C.T. Haas, R. Haas, Automated progress tracking using 4D schedule and 3D sensing technologies, Autom. Constr. 22 (2012) 414–421. doi:10.1016/j.autcon.2011.10.003.

Chen, J., Fang, Y., and Cho, Y. (2017). "Performance Evaluation of 3D Descriptors for Object Recognition in Construction Applications." Automation in Construction, Volume 86, February 2018, Pages 44-52 doi.org/10.1016/j.autcon.2017.10.033

[14] Y. Turkan, F. Bosché, C.T. Haas, R. Haas, Tracking of secondary and temporary objects in structural concrete work, Constr. Innov. Information, Process. Manag. 14 (2014) 145–167. doi:10.1108/CI-12-2012-0063.

[15] C. Kim, C. Kim, H. Son, Automated construction progress measurement using a 4D building information model and 3D data, Autom. Constr. 31 (2013) 75–82. doi:10.1016/j.autcon.2012.11.041.

[16] J. Lee, H. Son, C. Kim, C. Kim, Skeleton-based 3D reconstruction of as-built pipelines from laser-scan data, Autom. Constr. 35 (2013) 199–207. doi:10.1016/j.autcon.2013.05.009.

[17] H. Son, C. Kim, C. Kim, 3D reconstruction of as-built industrial instrumentation models from laser-scan data and a 3D CAD database based on prior knowledge, Autom. Constr. 49 (2015) 193–200. doi:10.1016/j.autcon.2014.08.007.

[18] H. Son, C. Kim, Automatic segmentation and 3D modeling of pipelines into constituent parts from laser-scan data of the built environment, Autom. Constr. 68 (2016) 203–211. doi:10.1016/j.autcon.2016.05.010.

[19] H. Son, C. Kim, C. Kim, A.M. Asce, Fully Automated As-Built 3D Pipeline Extraction Method from Laser-Scanned Data Based on Curvature Computation, J. Comput. Civ. Eng. 29 (2015) 1–9. doi:10.1061/(ASCE)CP.1943-5487.0000401.

[20] Y.K. Cho, M. Gai, Projection-Recognition-Projection Method for Automatic Object Recognition and Registration for Dynamic Heavy Equipment Operations, J. Comput. Civ. Eng. 28 (2014) A4014002. doi:10.1061/(ASCE)CP.1943-5487.0000332.

[21] H. Bay, A. Ess, Speeded-Up Robust Features (SURF), 110 (2008) 346–359. doi:10.1016/j.cviu.2007.09.014.

Chen, J., Fang, Y., and Cho, Y. (2017). "Performance Evaluation of 3D Descriptors for Object Recognition in Construction Applications." Automation in Construction, Volume 86, February 2018, Pages 44-52 doi.org/10.1016/j.autcon.2017.10.033

[22] D.G. Lowe, Distinctive image features from scale invariant keypoints, Int'l J. Comput. Vis. 60 (2004) 91–11020042. doi:http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94.

[23] P. Tang, D. Huber, B. Akinci, R. Lipman, A. Lytle, Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques, Autom. Constr. 19 (2010) 829–843. doi:10.1016/j.autcon.2010.06.007.

[24] A.E. Johnson, M. Hebert, Using spin images for efficient object recognition in cluttered 3D scenes, IEEE Trans. Pattern Anal. Mach. Intell. 21 (1999) 433–449. doi:10.1109/34.765655.

[25] F. Tombari, S. Salti, L. Di Stefano, Unique signatures of histograms for local surface description, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 6313 LNCS (2010) 356–369. doi:10.1007/978-3-642-15558-1_26.

[26] M. Körtgen, G.-J. Park, M. Novotni, R. Klein, 3D Shape Matching with 3D Shape Contexts, Proc. 7th Cent. Eur. Semin. Comput. Graph. 2003. (2003) 5–17. doi:10.1.1.67.9266.

[27] F. Tombari, S. Salti, L. Di Stefano, Unique shape context for 3d data description, Proc. ACM Work. 3D Object Retr. - 3DOR '10. (2010) 57. doi:10.1145/1877808.1877821.

[28] W. Wohlkinger, M. Vincze, Ensemble of shape functions for 3D object classification, 2011 IEEE Int. Conf. Robot. Biomimetics. (2011) 2987–2992. doi:10.1109/ROBIO.2011.6181760.

[29] R.B. Rusu, G. Bradski, R. Thibaux, J. Hsu, Fast 3D Recognition and Pose Using the Viewpoint Feature Histogram, in: Intell. Robot. Syst., 2010: pp. 2155–2162. doi:10.1109/IROS.2010.5651280.

Chen, J., Fang, Y., and Cho, Y. (2017). "Performance Evaluation of 3D Descriptors for Object Recognition in Construction Applications." Automation in Construction, Volume 86, February 2018, Pages 44-52 doi.org/10.1016/j.autcon.2017.10.033

[30] J. Chen, Y. Fang, Y.K. Cho, C. Kim, Principal Axes Descriptor for Automated Construction-Equipment Classification from Point Clouds, J. Comput. Civ. Eng. (2016) 1–12. doi:10.1061/(ASCE)CP.1943-5487.0000628.

[31] R.J. Campbell, P.J. Flynn, A Survey Of Free-Form Object Representation and Recognition Techniques, Comput. Vis. Image Underst. 81 (2001) 166–210. doi:10.1006/cviu.2000.0889.

[32] A.M. Bronstein, M.M. Bronstein, 3D features, surface descriptors, and object descriptors, (2010) 1–27. doi:10.1.1.637.3222.

[33] Y. Guo, M. Bennamoun, F. Sohel, M. Lu, J. Wan, 3D object recognition in cluttered scenes with local surface features: A survey, IEEE Trans. Pattern Anal. Mach. Intell. 36 (2014) 2270–2287. doi:10.1109/TPAMI.2014.2316828.

[34] S. Salti, F. Tombari, L. Di Stefano, A Performance Evaluation of 3D Keypoint Detectors, 3D Imaging, Model. Process. Vis. Transm. (3DIMPVT), 2011 Int. Conf. (2011) 282–289. doi:10.1109/3DIMPVT.2011.62.

[35] G. Arbeiter, S. Fuchs, R. Bormann, J. Fischer, A. Verl, Evaluation of 3D feature descriptors for classification of surface geometries in point clouds, IEEE Int. Conf. Intell. Robot. Syst. (2012) 1644–1650. doi:10.1109/IROS.2012.6385552.

[36] A. Dimitrov, M. Golparvar-fard, Segmentation of building point cloud models including detailed architectural / structural features and MEP systems, Autom. Constr. 51 (2015) 32–45. doi:10.1016/j.autcon.2014.12.015.

[37] O. Boiman, E. Shechtman, M. Irani, In Defense of Nearest-Neighbor Based Image Classification, IEEE Comput. Vis. Pattern Recognit. (2008). doi:10.1109/CVPR.2008.4587598.

Chen, J., Fang, Y., and Cho, Y. (2017). "Performance Evaluation of 3D Descriptors for Object Recognition in Construction Applications." Automation in Construction, Volume 86, February 2018, Pages 44-52 doi.org/10.1016/j.autcon.2017.10.033

[38] C. Chang, C. Lin, LIBSVM : A Library for Support Vector Machines, in: ACM Trans. Intell. Syst. Technol., 2011: p. 2:27:1--27:27. doi:10.1145/1961189.1961199.