REAL-TIME OBJECT RECOGNITION AND MODELING FOR HEAVY-EQUIPMENT OPERATION

Hyojoo Son, Changwan Kim
Department of Architectural Engineering
Chung-Ang University
221 Hueksuk-dong, Dongjak-gu,
Seoul, Korea 156-756
hjson0908@wm.cau.ac.kr; changwan@cau.ac.kr

Hyoungkwan Kim
School of Civil and Environmetal Engineering
Yonsei University
134 Shinchon-dong, Seodaemun-gu,
Seoul, Korea 120-749
hyoungkwan@yonsei.ac.kr

Kwang-Nam Choi, Jeong-Min Jee
School of Computer Science and Engineering
Chung-Ang University
221 Hueksuk-dong, Dongjak-gu,
Seoul, Korea 156-756
knchoi@cau.ac.kr; jeejjang@vimlab.cau.ac.kr

ABSTRACT

Recognition of free-form objects located in environments that are difficult to characterize or are constantly changing is critical in providing interactive background information for construction worksite modeling. It also allows for accurate, efficient, and autonomous operation of heavy equipment in a broad range of construction tasks. This paper presents a real-time process for 3D modeling of a construction worksite scene and focuses on modeling of the target object via flash LADAR, which could be applied to autonomous heavy-equipment operation. The proposed method consists of three steps: noise reduction, object extraction, and 3D model generation. The whole process is fully automatic and is performed in nearly real time. The method was validated in field experiments with actual construction objects. The results show that the proposed method effectively recognizes construction objects, which could be used to enhance efficiency and productivity in the autonomous operation of heavy equipment.

KEYWORDS

3D Modeling, Flash LADAR, Heavy Equipment Operation, Object Recognition, Point Clouds.

1. INTRODUCTION

On a construction site, the ability to cope with uncertainty is an important issue as regards construction automation. There are many unknown or free-form objects at a construction site, including workers, equipment, and materials, and these can be randomly scattered around the site. The nature of a construction environment is dynamic, with a variety of construction activities going on at once–activities that are constantly changing. In addition, numerous accidents occur on construction sites, caused either directly or indirectly by various types of heavy-equipment operations. Thus an effective 3D modeling process that could be applied to the operation of heavy equipment would be highly desirable–indeed a prerequisite for construction automation [1].

Graphical 3D models that provide geometric information have the potential to enable automation of certain construction tasks. In particular, automated control of heavy equipment may help in a number of construction applications, including obstacle avoidance for improved safety and greater efficiency in picking, placing, and tracking of objects [2, 3]. A 3D model is required to express the complex and dynamic work environment of a construction worksite effectively and in real time [4]. For these types of applications, rapid acquisition of accurate 3D information and a high frequency of updates are crucial. Flash LADAR, which has recently been developed, allows for real-time acquisition of 3D information and can operate day or night. As a result, it enables autonomous operation of heavy equipment and can be used in machine-vision applications [5].

Although laser scanning techniques have brought increasing attention to 3D models generated from 3D information, much more work needs to be done in order to automatically recognize objects in the captured 3D point clouds and build a useful 3D model [6]. Real-time object recognition and modeling on a cluttered construction site is a particularly challenging problem, and, because of the computational complexity of the post-processing phase, few research studies in this area have been undertaken [7]. Although many algorithms

recognize objects in an entire scene, there is no universal recognition procedure for all applications and all kinds of objects. Thus, automated object recognition for an entire scene is still a challenging task, and yet it is crucial for the effective operation of equipment that performs versatile tasks on construction sites [8].

The aim of this research is to develop a method of target object modeling from point clouds obtained with flash LADAR, with the goal of effectively representing the construction environment. For this purpose, a framework for a target object modeling process is proposed, consisting of algorithms that recognize the objects or clusters of objects in the point clouds, together with methods for automatically extracting feature points related to those objects and generating bounded 3D models of the scene. Field experiments have been conducted to test the performance of the proposed 3D modeling method.

2. RELATED STUDIES

2.1. Object Recognition

Automated recognition and modeling of free-form objects in unknown environments has proved to be a significant topic of investigation in the field of automation. Recognition is the process of classifying low-level point data into groups and associating them with pre-existing models. Most modeling applications require automated data processing for recognition of the 3D point clouds in order to extract useful information [6, 9]. Existing methods have various limitations, mainly in their accuracy, efficiency, robustness, and applicability to free-form objects [8]. Nevertheless, they have been used in a large number of applications, including automobile and robot vision, environmental planning, reverse engineering of industrial sites, building reconstruction, and automated searching of hazardous environments.

2.2. Model-driven approach vs. Data-driven approach

Existing approaches to object recognition in 3D point clouds fall into one of two main classes: the model-driven approach and the data-driven approach. The model-driven approach is used to process laser

scanning data; in most cases, it entails extraction—of features such as primitives or atypical shapes—from 3D point clouds and matching of the associated range data to database models [see, for example, 10, 11, 12]. Though the model-driven approach is popular, it cannot be generalized to unknown or free-form objects, and can hardly be considered applicable to those objects [13]. In addition, this approach requires a considerable amount of time to produce a final model. Thus, the model-driven approach is not suitable for recognizing and generating models of free-form objects for use in heavy-equipment operation in a construction environment, since such environments not only are difficult to characterize but also tend to undergo frequent change.

In the data-driven approach, local regions in the 3D point clouds that have smooth surfaces can be extracted by grouping nearby points that share some property, such as the direction of a locally estimated normal vector. This process can be considered a point-cloud segmentation problem, which shares many similarities with the image segmentation problem [9]. Therefore, image processing techniques such as filtering, thresholding, edge detection, and clustering, or a combination of these, can be applied to 3D point clouds. The latest research employing a data-driven approach shows that it can effectively model objects by applying different combinations of data-processing techniques such as filtering, feature extraction, and clustering [6]. In addition, this method is beneficial for providing users with the number of on-site objects as well as their respective volumes, shapes, and spatial orientations. However, experience shows that such an approach has its own limitations: several inputs are required, such as prior knowledge of the site for purposes of pruning the space, and the number of objects for purposes of clustering. Thus, this method is not sufficient to recognize the free-form objects in an unknown construction environment.

In the present study, a new data-driven approach was used to recognize the objects from 3D point clouds automatically and in real time.

3. MODELING PROCESS

The proposed construction-site modeling approach can be divided into (1) environmental object modeling and (2) target object modeling. It can be used to provide interactive background information about a construction site while a piece of heavy equipment is being operated there. In this study, the focus is on modeling of the target object.

Target object modeling applications entail distinguishing one type of point from another (e.g., distinguishing object points from non-object points). The main objective of this study is to extract the feature points of target objects from 3D point clouds. The 3D information acquired via flash LADAR consists of range data with range value. Background noise can be a crucial factor in object recognition in an unknown environment, because the 3D point clouds obtained with flash LADAR are limited by the ambiguity of object edges and the density of spatial noise. Therefore, eliminating unnecessary noise is the first step in our modeling process, and it is the most important phase. There has been extensive research on the use of filtering methods to reduce noise, and many conventional methods are available.

In this study, a new filtering method is proposed, one that uses differences in the range values from one pixel to another to weed out the noise in the range data. After filtering, the feature points related to target objects whose difference value lies below a pre-specified threshold value are extracted. Finally, the remaining noise is filtered out by a median filter, and the result is a collection of feature points related to the object. The median filter preserves real discontinuities while eliminating noise [14].

4. FIELD EXPERIMENTS

4.1. Experimental Setup

Outdoor experiments were conducted on an actual construction site to validate the proposed methods of object recognition and worksite modeling. In our experiments, the 3D data acquisition was accomplished using an SR-3000 mounted on a stationary tripod. The graphical models were displayed on a computer screen using the Matlab^{TM} graphical user interface (GUI).

Construction work was in progress at the construction site while the experiments were being performed. During the construction, pieces of equipment known as a forklift and a heavy-lift were used to lift materials such as piles of wood, stacks of foam, and piles of steel frame and transfer them to other workspaces at the construction site. The site is narrow and complex, so there is a significant chance of obstruction or accidents; moreover, it can be difficult to operate equipment on such a site, since the various construction materials are scattered randomly.

4.2. Experimental Results and Discussion

In the field experiments, target objects (stacks of foam) were chosen for validation of the proposed process of object recognition and modeling on this site; they were bounded by using a convex-hull algorithm. Figures 1(a) and 1(b) show a photographic image and the range image, respectively.
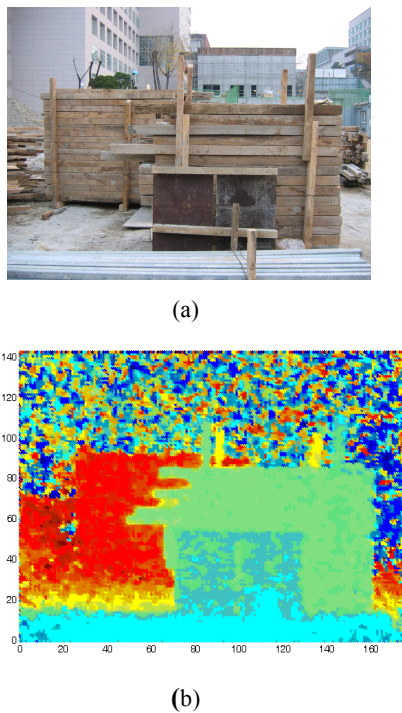
The scene is portrayed in Figure 2(a) as a mesh, which distinctly shows the amount of noise. Different colors represent different ranges of points on the site. Blue is used for the nearest points (those with a range of 0), while red is used for the most distant points (those with a range of ~7.5 m).

Suppose that, above a certain threshold, objects with the same characteristics have similar ranges, while points with values below that threshold represent noise. Then the similarity of range values can be transposed to a difference value by an easy calculation. If the window (or mask) size of the filter is $N \times N$ and the set of range values for the individual pixels in that window/mask is $W = \{w_1, ..., w_{N \times N}\}$, then the difference value DV is given by

$$DV = \sum_{\substack{i,\,j=1 \\ i \neq j}}^{N} \frac{\left| w_i - w_j \right|}{2} \approx \mathrm{Max}(W) - \mathrm{Min}(W) \qquad \textbf{(1)}$$

(a)

(b)

Figure 1. (a) Photographic Image, (b) Range Image
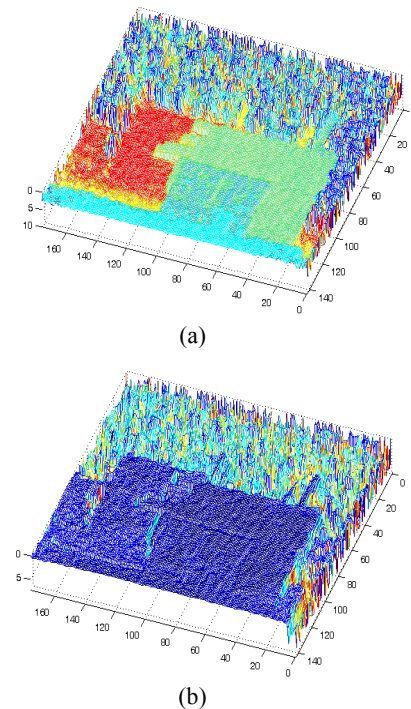

(a)

(b)

Figure 2. (a) Original Mesh, (b) After Difference Filtering with 3×3 Mask

235

The central pixel of the window is replaced with the difference value as the window slides over the image. Fig. 2(b) shows the result of applying the difference filter to the mesh in Fig. 2(a). In the filtered mesh, objects had values close to 0, and noise typically had high values. The mask size was set to 3×3. A smaller mask size was found to be more sensitive to object boundaries and small-sized objects, but it may be weak in terms of noise reduction.
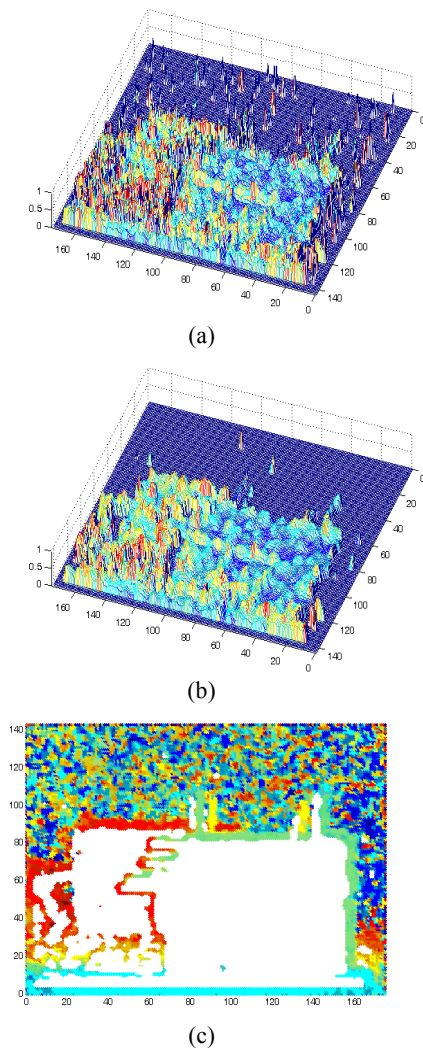

(a)

(b)

(c)

Figure 3. (a) Result of Extracting Points with Values below the Threshold Value of 0.7, (b) Result of Applying the Median Filter, (c) Object Extracted via the Proposed Process

Points with ranges below the threshold value were extracted after difference filtering over the image, as shown in Figure 3(a). The resulting image range was 0.0066~6.9124 m, according to the output of the depth scope, and objects were found to be distinguishable when the threshold value was set at 0.7 (10% of total depth). The final objects were extracted after use of a median filter to reduce the remaining noise (Figure 3(b)). As shown in Figure 3(c), this process effectively delineated the boundaries of objects found on the scene.

5. CONCLUSIONS

Achieving free-form object recognition rapidly and effectively from randomly scattered 3D point clouds remains an important topic of investigation. In this paper, a novel, real-time approach for object recognition and modeling of construction worksite scenes based on a data-driven approach was presented. The proposed target-object modeling method applies fundamental principles to object recognition and 3D representation of the environment. An actual construction worksite scene was used to show the results produced by the proposed algorithm, and the method was successfully implemented and tested using Matlab^{TM}. The experimental results presented in this paper demonstrate that filtering of noisy data is crucial to object recognition and modeling. This preliminary study also shows that the model achieves a good balance between speed and capacity, since it can rapidly process millions of 3D points. This method will be applied to a number of different environments in the future in order to test the efficiency of noise reduction in recognizing and modeling objects in an entire scene.

The spatial model generated from this method demonstrates the potential of such a model for assisting in the movement and operation of heavy equipment, thereby enhancing efficiency and productivity of automated equipment. In addition, it can help the workforce and engineers to make better and faster decisions and represents a preliminary step toward full construction automation.

6. ACKNOWLEDGEMENT

REFERENCES

[1] Kim, C., Haas, C. T., and Liapi, K. A.(2005) Rapid, On-site Spatial Information Acquisition and Its Use for Infrastructure Operation and Maintenance, Automation in Construction, Vol. 14, No. 5, 666–684.

[2] Haas, C. T., Skibniewski, N., and Bundy, E.(1995) History of Robotics in Civil Engineering, Microcomputers in Civil Engineering, Vol. 10, No. 5, 371–381.

[3] Kim, C. W., and Kwon, S. W.(2003) Human-assisted Rapid 3D Workspace Modeling Using Sparse Point Clouds for Heavy Equipment Operation on Construction Sites, Architectural Research, Vol. 5, No. 1, 69–76.

[4] Kim, C., Haas, C. T., Liapi, K. A., and Caldas, C. H.(2006) Human-Assisted Obstacle Avoidance System Using 3D Workspace Modeling for Construction Equipment Operation, Journal of Computing in Civil Engineering, ASCE, Vol, 20, No. 3, 177−186.

[5] Teizer, J., Kim, C., Bosche, F., Caldas, C. H., and Haas, C. T.(2005), Real-Time 3D Modeling for Accelerated and Safer Construction Using Emerging Technology, Proceedings of 1st International Conference on Construction Engineering and Management, Seoul, Korea, 539−543.

[6] Gong, J., and Caldas, C.(2007) Processing of High Frequency Local Area Laser Scans for Construction Site Resource Management, Proc. of the 2007 ASCE International Workshop on Computing in Civil Engineering, Pittsburgh, PA, 665−672.

[7] Lee, S., Jang, D., Kim, E., Hong, S., and Han J.(2005) A Real-Time 3D Workspace Modeling with Stereo Camera, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005), August 2–6, 2140–2147.

[8] Mian, A. S., Bennamoun, M., and Owens, R. A.(2004) A Novel Algorithm for Automatic 3D Model-based Free-form Object Recognition, IEEE Conference on Systems, Man and Cybernetics (IEEE SMC), 6348−6353.

[9] Vosselman, G., Gorte, B. G. H., Sithole, G., and Rabbani, T.(2004) Recognising Structure in Laser Scanner Point Clouds, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 46, Part 8/W2, Freiburg, Germany, October 4–6, 33–38.

[10] Bosche, F., Haas, C. T., and Caldas, C.(2005) 3D CAD Drawing as A Priori Knowledge for Machine Vision in Construction, Proceedings of the 1st Annual Inter-University Symposium on Infrastructure Management, Waterloo, Ontario, Canada, August 6.

[11] Cho, Y., Haas, C. T., Liapi, K., and Sreenivasan, S. V.(2002) A Framework for Rapid Local Area Modeling for Construction Automation, Journal of Automation in Construction, Vol. 11, No. 6, 629–641.

[12] Kwon, S., Bosche, F., Kim, C., Haas, C. T., and Liapi, K. A.(2004) Fitting Range Data to Primitives for Rapid Local 3D Modeling using Sparse Range Point Clouds, Journal of Automation in Construction, Vol. 13., No. 1, 67–81.

[13] Barnea, S., Filin, S., and Alchanatis, V.(2007) A Supervised Approach for Object Extraction from Terrestrial Laser Point Clouds Demonstrated on Trees, PIA07 Photogrammetric Image Analysis, Munich, Germany, September 19–21, 135–140.

[14] Stone, W., Cheok, G., and Lipman, R.(2000) Automated Earthmoving Status Determination, Proceedings of the 4th International Conference and Exposition/Demonstration on Robotics for Challenging Situations and Environments, Albuquerque, NM, USA, 111–119.