Rapid 3D Object Modeling Using 3D Data from Flash LADAR for Automated Construction Equipment Operations

Hyojoo Son[1], Changwan Kim[2] and Yoora Park[3]

[1] Research Assistant, Dept. of Architectural Engineering, Chung-Aug University, 221 Hueksuk-dong, Dongjak-gu, Seoul, Korea 156-756, PH (82) 2-825-5726; FAX (82) 2-825-5726; e-mail: hjson0908@wm.cau.ac.kr
[2] Assistant Professor, Dept. of Architectural Engineering, Chung-Aug University, 221 Hueksuk-dong, Dongjak-gu, Seoul, Korea 156-756, PH (82) 2-820-5726; FAX (82) 2-812-4150; e-mail: changwan@cau.ac.kr
[3] Research Assistant, Dept. of Architectural Engineering, Chung-Aug University, 221 Hueksuk-dong, Dongjak-gu, Seoul, Korea 156-756, PH (82) 2-825-5726; FAX (82) 2-825-5726; e-mail: saeara10@nate.com

Abstract

Automated recognition and modeling of 3D objects located in a construction work environment that are difficult to characterize or are constantly changing is critical for autonomous heavy equipment operation. Such automation allows for accurate, efficient, and autonomous operation of heavy equipment in a broad range of construction tasks by providing interactive background information. This paper presents 3D object recognition and modeling system from range data obtained from flash LADAR, with the goal of rapid and effective representation of the construction workspace. The proposed system consists of four steps: data acquisition, pre-processing, object segmentation on range images, and 3D model generation. During the object segmentation process, the split-and-merge algorithm, which separates a set of objects in a range image into individual objects, is applied to range images for the segmentation of objects. The whole process is automatic and is performed in nearly real time with an acceptable level of accuracy. The system was validated in outdoor experiments, and the results show that the proposed 3D object recognition and modeling system achieves a good balance between speed and accuracy, and hence could be used to enhance efficiency and productivity in the autonomous operation of heavy equipment.

Keywords: 3D object modeling; 3D object recognition; construction automation, construction heavy-equipment, flash LADAR, range image segmentation

1. Introduction

In recent years, with fundamental advances in sensor technology, it is becoming ever more feasible to automate construction equipment in ways that would help improve efficiency and safety of equipment operations on construction sites. While automation of heavy equipment has the potential to make an important contribution to productivity improvement, what is needed is an efficient way to represent a workspace in 3D and incorporate that representation into control of equipment operations. There are requirements that 3D modeling methods for use in construction automation have to satisfy; data acquisition speed, level of intricacy, versatility, and automated data processing.

Recently, 3D modeling methods have been investigated to represent construction workspace for several applications such as as-built drawings, visual feedback to equipment operators, and construction materials tracking. Most research on 3D modeling in the construction industry have employed large and expensive 3D laser scanner to produce dense point clouds. While 3D laser scanning can produce very detailed models of the scanned scenes, which are useful for obtaining as-built drawings of existing structures, extensive processing of the received point clouds is needed in order to construct the 3D model, thereby processing the entire 3D scanning process too laborious and time-consuming for the intended applications. The burdens imposed by a 3D laser scanner in terms of processing time generally preclude the use such technology for real-time decision-making.

An alternative method is based on the use of flash LADAR which encompasses a new generation of scannerless LADAR devices. A flash LADAR device produces an image of the observed scene in which each pixel value represents the intensity and range of the corresponding image area (Uijt de Haag et al. 2008). When compared to a 3D laser scanner, a scannerless flash LADAR device is smaller in size, lighter in weight, and lower in cost and has advantage of acquiring data in real-time (Stone and Juberts 2002; Habbit et al. 2003). With the capability of capturing range data at high speed, flash LADAR produces a 3D image an entire scene in a single data acquisition step; moreover, it can capture moving objects, and hence can provide both static and dynamic information (Habbit et al. 2003). As a result, flash LADAR is beneficial for real-time applications such as obstacle detection, equipment navigation and object recognition (Price et al. 2007). Although flash LADAR has brought a new means of achieving real-time 3D model generation from range data, few research studies on this technology have been undertaken in the construction industry (Teizer et al. 2007; Gong and Caldas 2008; Kim et al. 2008). Much more work needs to be done in order to achieve automated recognition and modeling of objects on a construction site from range data and reconstruction of useful 3D model in near-real-time.

The aim of this research is to develop a system for 3D object recognition and modeling using range data from flash LADAR, with the goal of rapidly and effectively representing construction workspaces. To achieve this purpose, a system is proposed, consisting of algorithms that recognize the objects in the scene, together with methods for automatically extracting feature points related to those objects and generating bounded 3D models of each object. And outdoor experiments have been performed to test the performance of the proposed method.

2. Framework for 3D Object Modeling for Use in Heavy Equipment Operation

In this section, an overview of the framework for the proposed automatic 3D object recognition and modeling system using range image is presented. The process used in the 3D object recognition and modeling system proposed herein is outlined in the flowchart in Figure 1.

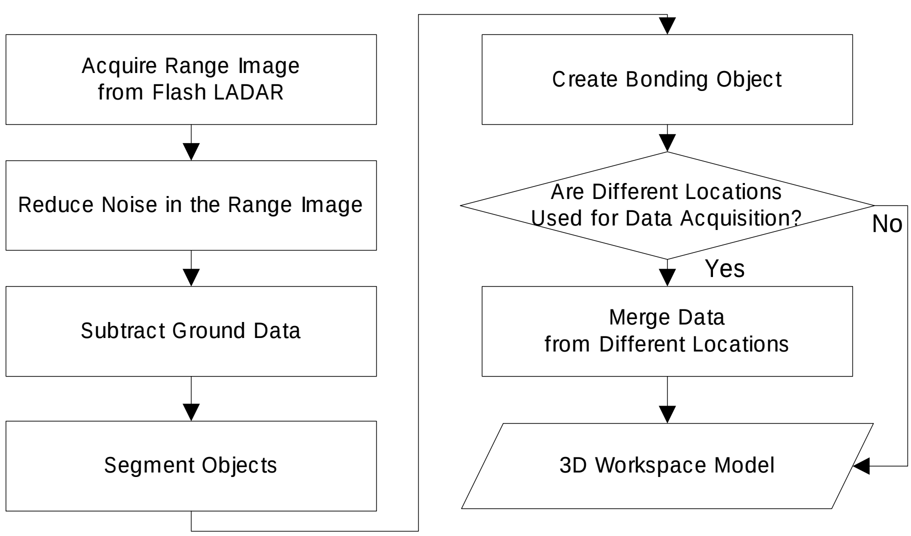


Figure 1 Proposed Process for 3D Workspace Modeling

The first step of the 3D object recognition and modeling process is to acquire the 3D data that adequately covers the 3D scene. As mentioned earlier, workspace models are required to express dynamic work environment of a construction site effectively and in near-real-time for construction automation applications (Kim et al. 2006). For this reason, not only high frequency of updated of 3D data acquisition but also an acceptable level of accuracy are needed for reliable and successful 3D object modeling. Thus, in this study, the 3D data of the objects were acquired by using a SwissRanger SR-3000 flash LADAR, since it provides the most adapted trade-off between the data acquisition speed and the accuracy of the 3D data in case of real-time applications (Bosche et al. 2005).

Once the range image acquired, the range image acquired from high speed range scanner such as flash LADAR contains noise of considerable level (Frome et al. 2004). If the noise is not reduced, it may affect a negative effect on object recognition; therefore, pre-processing is needed. For this, data filtering method is performed to reduce the noise influence. In addition, objects on the ground are one that its boundary between object region and ground region is hard to recognize. To detect and remove ground data, ground subtraction technique is used.

Pre-processing step is completed; the feature points of objects are extracted without distinction between different objects. The range image should be separated into individual objects. This is the segmentation process which is required to recognize the objects in a scene. In this study, a split-and-merge algorithm is applied to separate the set of extracted feature points into individual objects

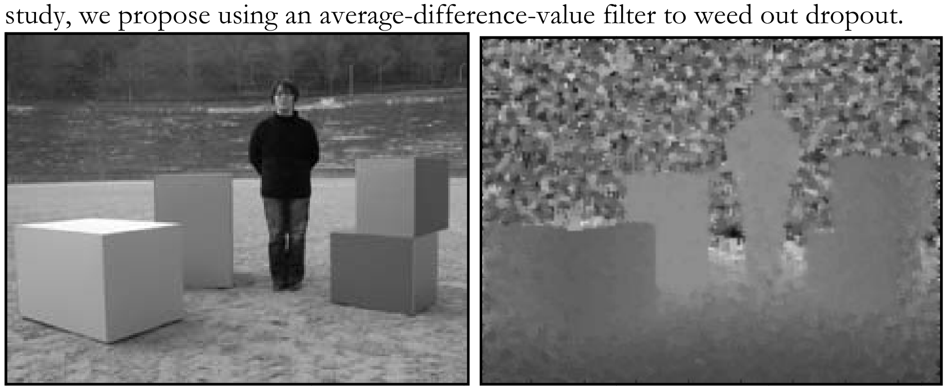
Then bounding models of objects representing various construction site scenes are created through the use of a general class of geometric primitives. In cases where the operator's view from any one reference point is limited and the data have to be acquired from two or more locations, the data sets obtained from the different locations are merged into a single data set having one common coordinate system by applying an ICP (iterative closest point) algorithm. The 3D models generated from data acquired on-site via the process described above are useful as a tool for providing interactive feedback to equipment operators. Additionally, the models can be shown as 3D simulations that offer equipment operators an opportunity to experience the results of certain aspects of the tasks at hand before actual operation.

3. Experiments and Results

This section describes the detailed methods of 3D object recognition and modeling system with the discussion of outdoor experimental results. Outdoor experiments were conducted to establish validation for the proposed object recognition and modeling system.

3.1. Pre-processing

After the data acquisition using flash LADAR, 25,344 data points were acquired per range image. Range image acquired from flash LADAR in an outdoor environment tend to contain large regions of dropout, because of the measurement limits of flash LADAR and outdoor environmental conditions (Frome et al. 2004). Figure 2(b) shows a range image in which fluctuations in the level of gray indicate false range values. Thus, in this study, we propose using an average-difference-value filter to weed out dropout.



(a) (b)

Figure 2 (a) Photographic Image, (b) Range Image

The average difference value ADV employed in this research is the average of the differences between the value of a given pixel and those of its eight neighbors in the 3×3 window centered at that pixel. If the value of ADV is larger than some predefined threshold, the central pixel is assumed to be corrupted by dropout and is eliminated. Otherwise, the central pixel is left unchanged. Throughout the process, points with a range above a threshold value of 0.6 were weeded out. This threshold was selected after performing a set of experiments and finding that it successfully detected the noisy regions of range images acquired from flash LADAR.

After the results of the average-difference filtering were applied, there was still speckle noise left in the image, especially in the object region, which caused measurement errors to creep in. In this study, a 3×3 median filter was used to remove speckle noise and render the surfaces of objects more uniformly. A median filter is useful in eliminating speckle noise in a range image while preserving edge information (Doss 2004). As shown in Figure 3(a), noise was reduced to an acceptable level after using the proposed filtering method, while object regions were preserved.

However, in this case, there is the large part of the points corresponding to the ground floor which makes difficult to segment the objects. Thus, the next step undertaken in the object modeling is removal of data corresponding to the ground floor of the scene. In cases where the sensor is located on a rigid body, and the height and vertical angles are known, ground plane detection is pretty straightforward. After the data are subjected to an appropriate coordinate rotation, the ground data can be extracted using a threshold value (Bostelman et al. 2006). The result of such an approach used for removal of ground data in this research is shown in Figure 3(b). After the pre-processing step, about the 65% (16,431 points) of the raw data points were filtered and removed.

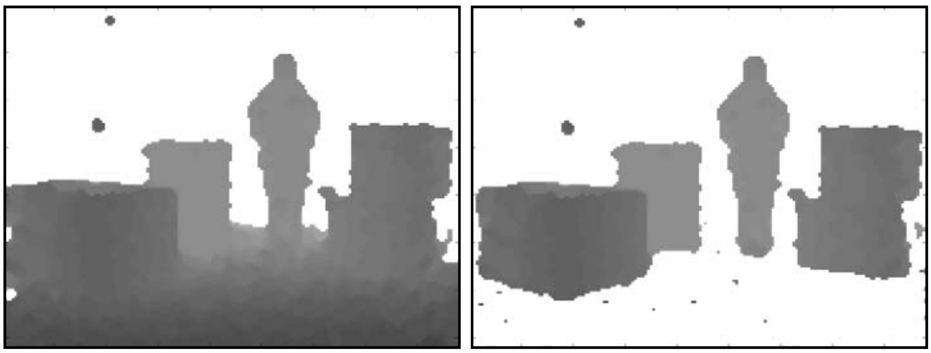


Figure 3 (a) Result of Data Filtering Process, (b) Filtering Out of Ground Floor Related Data Points

3.2. Object Segmentation

In the pre-processing stage, range values for the unwanted regions are effectively extracted. Once that is accomplished, a process of decomposition of the resulting image into separated objects is undertaken. Various methods for object segmentation, which are classified as boundary-based approaches and region-based approaches, have been developed (Xiang and Wang 2004). Boundary-based approaches identify the edges based on discontinuities (pixel differences) and link up edges to produce closed boundaries for individual objects (Lin and Talbot 2001). However, their applicability is limited by the fact that it is difficult to find complete boundary information for an object, especially in a noisy image (Ikeuchi 1987). Region-based approaches take a noisy image and distinguish coherent regions that satisfy a predefined homogeneity condition, and then use those coherent regions to identify the objects of interest (Kelkar and Gupta 2008). In this study, split-and-merge segmentation, which is one of the main methods used in region-based approaches, was adopted (Lin and Talbot 2001). This method has the advantage of simplicity of use as well as computational efficiency by a combination of splitting and merging of regions in the image (Salih and Ramli 2001; Sun and Du 2004).

The object segmentation process based on split-and-merge algorithm comprises following two steps: splitting and merging. At the first split operation, the splitting process starts with the complete range image Figure 3(b) as a single region R . If R is inhomogeneous, it is split into four subregions—in particular, four rectangular blocks of equal size. This process is repeated recursively until all subregions are homogeneous. Since the splitting process might have split up homogeneous regions, a merging process is then used to test the homogeneity of adjacent regions and merge them into a single region if their union is homogeneous. Through the object segmentation process, four objects were successfully separated as shown in Figure 4.

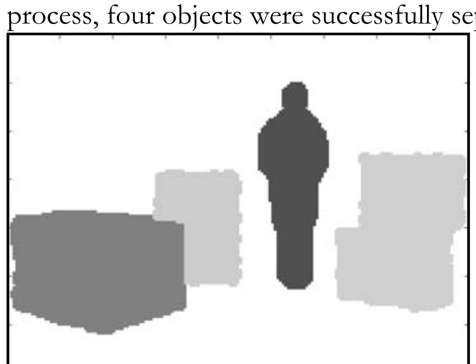


Figure 4 Result of the Object Segmentation Process

3.3. Model Generation

In the object segmentation process, the feature points of objects were successfully segmented and identified into four objects. Based on the segmented feature points of each object, each object points were connected to generate 3D model by using a convex hull algorithm as shown in Figure 5. The convex hull algorithm used in this research is an incremental algorithm developed by Barber et al. (1996); it is well suited for rapid 3D object modeling, because its use in the generation of bounding models to represent the spatial shapes of 3D objects is compact, fast, and relatively efficient (Kim et al. 2006).

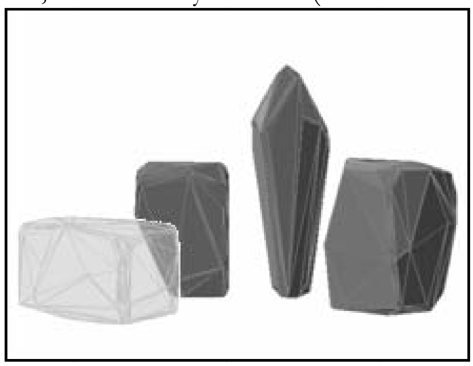


Figure 5 3D Model Generated from One Point of View (the object at left is shown as a translucent model)

In the case at hand, there were two objects in the left portion of the scene, one of which was partially occluded by the other, as shown in Figure 5. Also, in some cases, it may be impossible to acquire an image of the complete scene due to limitations of the sensor's field of view or for geometric reasons (Neugebauer 1997; Mure-Dubois and Hugli 2008). Thus, the supplemental 3D data was acquired from different location and segmented into separated objects. And data sets acquired from different locations were registered together with the respect to the same coordinate system. In this study, an ICP algorithm developed by Besel and McKay (Besl and McKay 1992) was used for that purpose. The ICP algorithm utilized here, which was designed for use with free-form surfaces, works by automatically matching the closest point in one set of 3D data to another set of 3D data (Allen et al. 2003). Through this process, complete 3D models were generated by making up for range image (for the second object in the left portion of the range) that were not acquired in the first round of data acquisition, as shown in Figure 6.

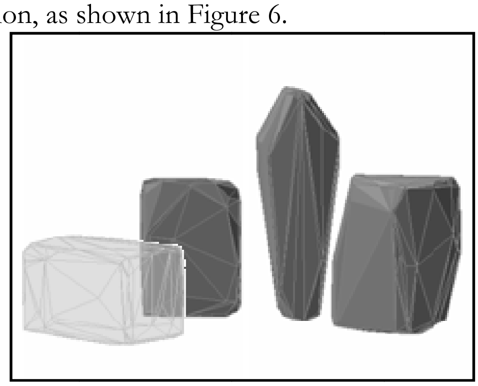


Figure 6 Merged 3D Models (the object at left is shown as a translucent model)

5. Conclusions

This study presented rapid and efficient 3D object recognition and modeling system for use in automated operation of construction equipment. Flash LADAR is shown to be a viable means of acquiring real-time spatial information in the form of a range image of a construction-site scene. The proposed data processing scheme, which includes pre-processing, object segmentation, and model generation, utilizes various algorithms to recognize the objects in the scene from the acquired data and then automatically generate a 3D model. The effectiveness of the proposed object recognition and modeling system was validated. Outdoor experimental results demonstrated that it takes only a few seconds to generate 3D models with the proposed method, hence that it can be used for automated object recognition and modeling of construction objects in near-real-time. Adoption of such a method could facilitate the safe and efficient operation of heavy equipment, since it not only provides a 3D graphical representation of the scene but also enables spatial analysis of the current set of conditions on a construction site.

Although the proposed modeling method has been applied in this research to the modeling of static objects only, we see no reason why it could not be extended to the modeling of dynamic objects. Also, the proposed method would seem to lend itself to use in control systems for heavy construction equipment, where it could be used, for example, as part of an obstacle avoidance system.

6. Acknowledgements

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (KRF-2008-313-D01164).

References

[1] Allen, P. K., Troccoli, A., Smith, B., Murray, S., Stamos, I., and Leordeanu, M. (2003). “New methods for digital modeling of historic sites.” IEEE Comp. Graph. Appl., 23(6), 32–41.

[2] Barber, B., Cobkin, D., and Huhdanpaa, H. (1996). “Quickhull algorithm for convex hull.” ACM Trans. Math. Software, 22(4), 469–483.

[3] Besl P. and McKay N. (1992). “A method for registration of 3-D shapes.” IEEE Trans. Pat. Anal. and Mach. Intel., 14(2), 239–256.

[4] Bosche, F., Haas, C. T., and Caldas, C. H. (2005). “3D CAD drawing as a priori knowledge for machine vision in construction.” Proc., 1st Ann. Inter-University Symp. on Infrastructure Management (AISIM), Waterloo, Canada.

[5] Bostelman, R., Russo, P., Albus, J., Hong, T., and Madhavan, R. (2006). “Applications of a 3D range camera towards healthcare mobility aids.” Proc., 2006 IEEE Int. Conf. on Networking, Sensing, and Control (ICNSC ’06), Ft. Lauderdale, Fl., 416–421.

[6] Frome A., Huber, D., Kolluri, R., Bulow, T., and Malik, J. (2004). “Recognizing objects in range data using regional point descriptors.” Proc., 8th European Conf. Computer Vision (ECCV), Prague, Czech Republic, 224–237.

[7] Doss, N. (2004). “3D modeling of biological structures.” BS thesis, Univ. Western Australia, Nedlands, WA 6907, Australia.

[8] Gong, J., and Caldas, C. H. (2008). “Data processing for real-time construction site spatial modeling.” J. Autom. Constr., 17(5), 526–535.

[9] Habbit, R. D., Nellums, R. O., Niese, A. D., and Rodriguez, J. L. (2003). “Utilization of flash ladar for cooperative and uncooperative rendezvous and capture.” Proc., SPIE, 5088, Orlando, Fl., 146–157.

[10] Ikeuchi, K. (1987). “Generating an interpretation tree from a CAD model for 3D-object recognition in bin-picking tasks.” Int. J. Comput. Vision, 1(2), 145–165.

[11] Kelkar, D., and Gupta, S. (2008). “Improved quadtree method for split merge image segmentation.” Proc., 2008 1st Int. Conf. on Emerging Trends in Engineering and Technology, Nagpur, India, 44–47.

[12] 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.” J. Comput. Civ. Eng., 20(3), 177–186.

[13] Kim, C., Son, H., Kim, H., and Han, S. (2008), “Applicability of flash laser distance and ranging to three-dimensional spatial information acquisition and modeling on a construction site.” Can. J. Civ. Eng., 35(11), 1331–1341.

[14] Lin, Z., Jin, J., and Talbot, H. (2001). “Unseeded region growing for 3D image segmentation.” Proc., Selected papers from Pan-Sydney Area Worksh. on Visual Information Processing (VIP2000), Sydney, Australia, 31–37.

[15] Mure-Dubois, J., and Hugli, H. (2008). “Fusion of time of flight camera point clouds.” Proc. IEEE Worksh. on Multi-camera and Multi-model Sensor Fusion Algorithms and Applications (M2SFA2), Marseille, France.

[16] Neugebauer, P. J. (1997). “Geometrical cloning of 3D objects via simultaneous registration of multiple range images.” Proc., 1997 Int. Conf. Shape Modeling and Applications (SMA ’97), 130–139.

[17] Price, M., Kenney, J., Eastman, R. D., and Hong, T. (2007). "Training and optimization of operating parameters for flash ladar cameras." Proc., 2007 IEEE Int. Conf. on Robotics and Automation, Roma, Italy, 3408–3413.
[18] Salih, Q. A., and Ramli, A. R. (2001). "Region based segmentation technique and algorithms for 3D image." Proc., Int. Symp. on Signal Processing and Its Applications (ISSPA), Kuala Lumpur, Malaysia, 747–748.
[19] Stone, W. C., and Juberts, M. (2002). "Towards the ultimate construction site sensor." Proc., 19th Int. Symp. on Automation and Robotics in Construction (ISARC), National Institute of Standards and Technology, Gaithersburg, Maryland, 393–400.
[20] Sun, D., and Du, C. (2004). "Segmentation of complex food images by stick growing and merging algorithm." J. Food Eng., 61(1), 17–26.
[21] Teizer, J., Caldas, C. H., and Haas, C. T. (2007). "Real-time three-dimensional occupancy grid modeling for the detection and tracking of construction resources." J. Construct. Eng. Manag., 133(11), 880–888.
[22] Uijt de Haag, M., Venable, D., and Soloviev, A. (2008). "Implementation of a flash-ladar aided inertial navigator." Proc., IEEE/ION Position, Location, and Navigation Symp., Monterey, Ca., 832–837.
[23] Xiang, R., and Wang, R. (2004). "Range image segmentation based on split-merge clustering." Proc., 17th Int. Conf. on Pattern Recognition, Cambridge, UK, 614–617.