REAL-TIME SPATIAL DETECTION AND TRACKING OF RESOURCES IN A CONSTRUCTION ENVIRONMENT

Jochen Teizer[1], Frederic Bosche[2], Carlos H. Caldas[3], and Carl T. Haas[4]

ABSTRACT

Construction accidents with heavy equipment and bad decision making can be based on poor knowledge of the site environment and in both cases may lead to work interruptions and costly delays. Supporting the construction environment with real-time generated three-dimensional (3D) models can help preventing accidents as well as support management by modeling infrastructure assets in 3D. Such models can be integrated in the path planning of construction equipment operations for obstacle avoidance or in a 4D model that simulates construction processes. Detecting and guiding resources, such as personnel, machines and materials in and to the right place on time requires methods and technologies supplying information in real-time.

This paper presents research in real-time 3D laser scanning and modeling using high range frame update rate scanning technology. Existing and emerging sensors and techniques in three-dimensional modeling are explained. The presented research successfully developed computational models and algorithms for the real-time detection, tracking, and three-dimensional modeling of static and dynamic construction resources, such as workforce, machines, equipment, and materials based on a 3D video range camera. In particular, the proposed algorithm for rapidly modeling three-dimensional scenes is explained. Laboratory and outdoor field experiments that were conducted to validate the algorithm's performance and results are discussed.

KEY WORDS

Occupancy Grid Algorithm, Range Sensing, Real-Time 3D Modeling, Resource Detection and Tracking, Safety, Voxel

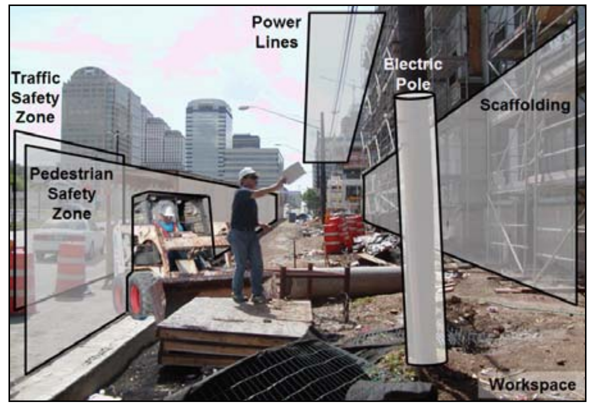
INTRODUCTION

Real-time three-dimensional modeling of construction environments is of fundamental as well as technological interest to the construction community. Fundamentally, three-

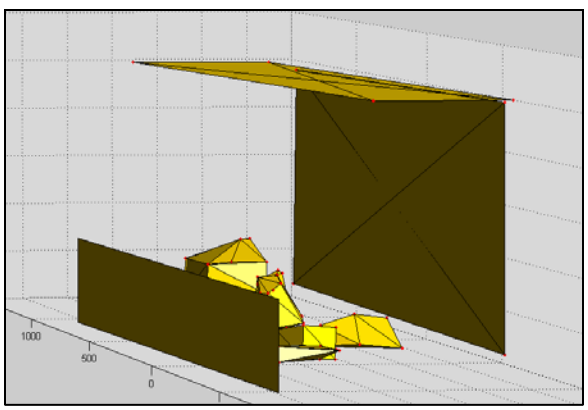
1 Ph.D. Candidate, Department of Civil, Architectural, and Environmental Engineering, The University of Texas, Austin, TX 78712, Phone +1 512/554-7857, FAX 512/471-3191, teizer@mail.utexas.edu
2 Ph.D. Candidate, Department of Civil Engineering, University of Waterloo, Ontario, Canada N2L 3G1, Phone +1 519/888-567 ext. 3872, FAX 519/888-4300, fbosche@engmail.uwaterloo.ca
3 Assistant Professor, Department of Civil, Architectural, and Environmental Engineering, The University of Texas, Austin, TX 78712, Phone +1 512/471-6014, FAX 512/471-3191, caldas@mail.utexas.edu
4 Professor, Department of Civil Engineering, University of Waterloo, Ontario, Canada N2L 3G1, Phone +1 519/888-4567 ext. 5492, FAX 519/888-4300, chass@civmail.uwaterloo.ca

dimensional modeling enables the comparison between the planned and actual spatial status of complex systems in order to support better management decisions. Technologically, the development of ever smaller and cheaper electronic components for safer and faster operation of heavy equipment or for materials tracking requires the investigation of multi-disciplinary fields and principles from areas such as manufacturing, logistic, remote sensing, and transportation. Ultimately, such an approach can assist in solving some of the current problems in construction and make construction processes more effective and efficient by reducing accidents, cost, schedule, and waste (Goodrum and Haas, 2002).

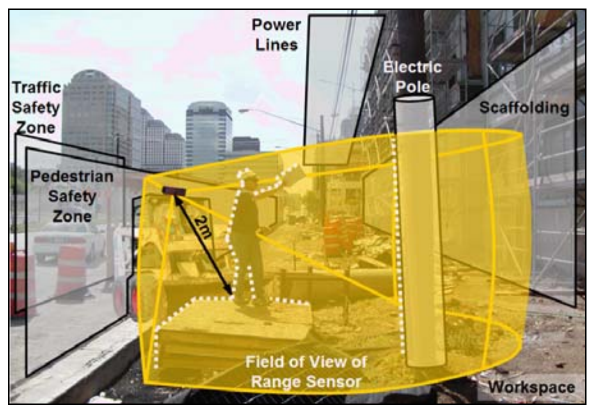
(a) Original scene of potential hazards



(b) Sparse point cloud modeling approach



(c) Sparse point cloud model of static objects



(d) 3D range camera modeling approach

Figure 1: Real-time 3D Modeling Approach for Obstacle Avoidance System

The approach of this research is presented in Figure 1. To execute construction sites safer, faster, and at higher productivity levels, an obstacle avoidance system was designed to increase the limited perception of heavy equipment operators through real-time three-dimensional modeling. The obstacle avoidance system is based upon five steps: (1) Acquire range information of construction scenes using existing or emerging technologies (Figure 1a), (2) create 3D models of objects with permanent static hazardous nature (Figure 1b and 1c), (3) detect and track 3D models of objects with a temporary hazardous nature (Figure 1d),

(4) integrate 3D models in a so called World Model, that includes all scene relevant information of object location, dimension, velocity, and direction, and (5) run and operate an obstacle avoidance system based on range sensing in a simulated environment and on realistic construction job site equipment. Since research steps 1 and 2 were developed in previous research efforts, this paper presents algorithm developments, experiments and results of research step (3).

BACKGROUND REVIEW

Construction site layouts usually host multiple sources of objects with hazardous potential, which in general are of permanent static nature (street with ongoing traffic, pedestrians in walkways, high voltage in power lines), of temporal static nature (erected scaffolding walls and structures, and at various locations placed and stored materials), or of moving nature (workforce, equipment and machines). All of which might have the potential of critical impact on productivity, cost, or on safety.

Reach-in and interference of machines, workers, or construction related materials into vehicular and pedestrian traffic space can cause collateral damage as well as bodily injuries or fatalities. As a result, the typical construction environment needs boundaries to divide the "civil space" from the "construction workspace". The civil space is the area in where construction work is generally not permitted unless certain safety protection guidelines are followed. The civil space is usually separated from construction workspace by installing safety fences, protective barriers such as traffic cones, or covered pedestrian sidewalks. The original scene image in Figure 1 shows that objects in the construction workspace itself are not protected, thus offer a hazardous potential to resources such as workforce, materials, and equipment. Variously placed materials, moving workforce and objects such as power lines make it difficult for heavy equipment, such as a skid steer loader or crane, to navigate on a job sites safely and effectively.

An obstacle avoidance system which is based on the generation of rapid three-dimensional models must be able to detect, track, and characterize each object within the workspace. The idea of rapidly building three-dimensional barriers to areas where construction equipment faces elevated danger (Figure 1b), e.g. street, walkway, power lines, scaffolding walls, or electric poles, allows granting i.e. machine access to work only in safe zones.

In this research approach objects can be modeled in 3D by using two different approaches. Permanently located objects that don't frequently change their position and shape are modeled after the Sparse Point Cloud approach (Song, 2004 and Kim, 2004). All other objects which very frequently change their location or geometric shape may need real-time updates to accurately determine their position in the construction workspace. Integrating both approaches allows building an obstacle avoidance system.

RANGE DATA ACQUISITION OF STATIC SCENES

LAser Detection and Ranging (LADAR) sensing technology such as commercially available laser scanners capture millions of range points of static scenes (Leica, 2006). The range data acquisition process of laser scanners involves a manual sensor installation and allows the automated acquisition of range values to image pixels of entire scenes within several minutes. The large number of collecting range points allows applications in a wide field of construction work, e.g. comparison of as-built to planned information, defect detection, etc. (Teizer et al., 2005). The range data processing and supply of 3D information to a CAD model, however, can take up to several hours or days. As a result, object dimension and location cannot be analyzed instantly due to the lack of real-time data acquisition and processing. New sensing and modeling methods are required for application where objects have a dynamic nature.

In the goal of developing a real-time obstacle avoidance system, one step towards the generation of a more rapid three-dimensional modeling approach was successfully demonstrated using a Sparse Point Cloud modeling approach. The basic modeling principle is described in Figure 1b. The Sparse Point Cloud modeling approach focuses on single range point data acquisition which leads to faster range data processing and modeling.

The manual selection of range points using low cost commercial Laser Range Finder technology eases the search in finding the significant geometric object boundaries, such as corner points of a box to determine its bounding volume. Sparse Point Cloud modeling is a rapid semi-automated approach that shortens the overall 3D model generation process to a few minutes. However, this approach does not allow creating 3D models of any kind of object that has a frequent change in location including moving objects such as workforce or equipment (Teizer et al., 2005).

RANGE DATA ACQUISITION OF STATIC AND MOVING OBJECTS

This research used emerging technology like 3D video range cameras to be mounted on heavy equipment. It intended to create an obstacle avoidance system based on three-dimensional modeling of the static and moving construction environment in real-time.

This paper discusses experiments and results in detecting and tracking moving targets in the field of view of a 3D video range camera and demonstrates the general feasibility of applying this research approach to other areas, such as comparison of as-built to as-planned range data or in the integration to 3D or 4D CAD models (Bosche, 2006).

All of these applications require overcoming the discussed difficulties of the existing three-dimensional modeling approaches. The research objective was to find and address the three-dimensional geometric characteristics of static or moving objects within a larger field of view of a range camera.

The developed approach uses 3D video range camera, a.k.a. Flash LADAR that is based on a contactless distance measurements principle. In this research a SwissRanger 2 range sensor emits a continuous near-infrared light wave (880nm) in a scene at a modulation frequency of 20Mhz. Amplitude samples of the reflected wave determine the distance to each of the 160x124 pixels (resolution) after the phase-shift principle. Range image frames are collected at frame update rates of up to 30Hz in a field of view of horizontal 42° and vertical 45°. The range accuracy is less than 5cm at a non-ambiguous distance of 7.5m. One of the current limitations of the 3D video range sensing technology is the prototype stage of the hardware. More importantly, range image processing algorithms need to be developed to accurately detect and track objects in the field of view of a range camera (CSEM, 2004).

THREE-DIMENSIONAL OCCUPANCY GRID ALGORITHM

Initial developments in processing the raw range data were based on existing image processing techniques, such as Canny edge detection and clustering techniques such as k-means. Results were successful, however, the duration to process a single range frames in a MatLab® environment was up to 4 minutes. Moreover, k-means clustering and other clustering techniques required a priori knowledge to supply and detect the correct number of objects contained in a scene. These downsides of knowledge based approaches limited the use of these processing techniques and thus asked for a more robust range data processing algorithm.

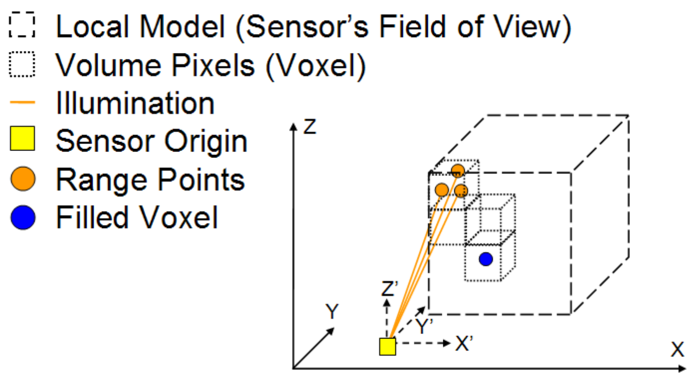


Figure 2: Three-dimenstional occupancy grid

Three-dimensional occupancy grids offered a different approach. The working principle of an occupancy grid can be seen in Figure 2. In a first step, the three-dimensional occupancy grid allocates the originally collected range matrix of range points into so called voxels (volume pixels). A voxel is defined by the grid size of the Local Model, also known as the field of view of the sensor. The developed algorithm used a variety of empirically found threshold values to filter noise measurements (local extremes called "salt and pepper") and to cluster the remaining range values.

With this 3D occupancy grid approach an object can be detected and tracked by collecting the following information:

- Location (Position of the center of gravity of each cluster in 3D)
- Dimension (Length of object in all axes)
- Velocity (Speed of object)
- Direction (Orientation of velocity vector)

The knowledge of where objects are, knowing their accurate dimension, velocities, and direction can help to plan the path of heavy equipment operation. Furthermore, it allows detecting and tracking objects such as materials for tracking purposes.

An occupancy grid algorithm with the grid size of 0.1m in all axes (X, Y, and Z) was used to analyze the raw range data collected in indoor and outdoor experiments. The grid size was based on empiric values and allowed to capture small sized objects as well as reduced the overall number of range points for processing. The basic working mechanism of the algorithm is described next: Two range points in each voxel were needed to keep a voxel filled. If more than six surrounding neighbors of this voxel were filled as well, the voxel was kept, otherwise deleted with no further computational burden. Segmentation into single clusters was based on grouping at least 10 voxels together and if the distance to the next group was less than 2 voxels, or respectively 20cm, the cluster was combined with its neighbor cluster using hierarchical agglomerative clustering (Elfes, 1989). The algorithm calculated the center of gravity of each cluster which was used to track moving objects. The cluster kept its identity if its volume from one frame to the next did not change more than 25%. If the difference in the location of the center of gravity of each cluster changed between one or two voxels the cluster was identified moving, otherwise identified as a static object or assumed to be a new object appearing in the scene.

EXPERIMENTS AND RESULTS

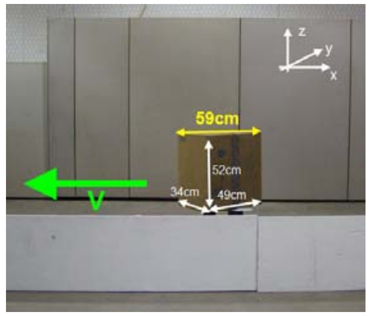
The functionality of the developed occupancy grid algorithm was verified in experiments. Range data of static as well as moving objects placed in the field of view of the range camera was collected. The occupancy grid algorithm was used to create 3D models in real-time. The following example output of one experiment is demonstrated in Figure 3.

A target object, e.g. a box, was propelled at various speeds and different angles guarded on a rail through the field of view of the 3D video camera (see Figure 3a). The target object varied from different sized boxes, round aluminum pipes, and a human representing a construction worker. In the front of the image a fascia board covered the cart which carried the target object. To reference the range data obtained from the 3D video camera, the dimensional values of the target object were measured with a commercially available laser range finder (a Total Station). Additional collected values characterized the experimental environment, e.g. temperature inside laboratory and humidity level, and were stored in an experimental log book.

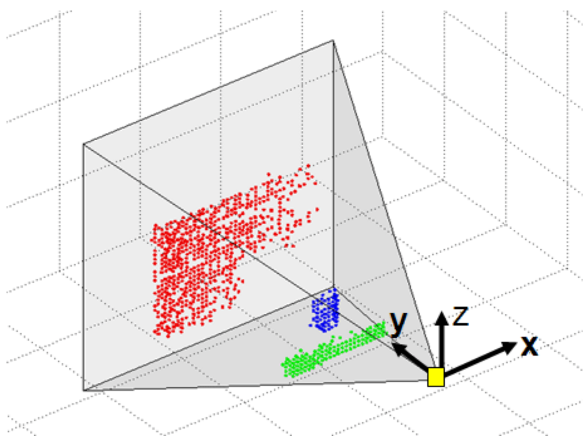
The target object was propelled through the field of view of the range sensor and range images at frame update rates of 15.2 Hz were recorded. The described occupancy grid algorithm converted the raw range data into an occupancy grid model. The result of the algorithm is displayed in Figures 3b to 3e.

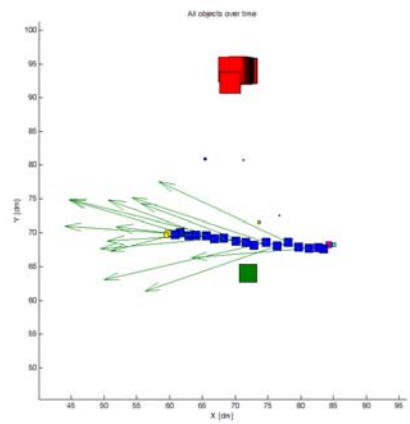
In Figure 3b the three-dimensional view of the occupancy grid can be seen. The three objects placed within the field of view of the range camera were successfully detected. The fascia board, the box, and the background wall are well separated in different clusters. To each cluster and each frame a center of gravity was generated. The center of gravity is calculated using the average of voxel locations in each cluster. The number of voxels in each cluster determines the size of the square of the each center of gravity. In Figure 3c the plane view of 52 consecutive range frames taken in this experiment are plotted. The trajectory path and the velocity vectors of the box can be seen. The static fascia board received a single location for its center of gravity. The square, demonstrating the center of gravity of the fascia board is bigger than those ones generated from the occupancy grid algorithm for the box. Since the number of points in the “fascia board” cluster is larger then in the “box” cluster this particular displaying and tracking element of the occupancy grid algorithm was successful.

The size of the "background wall" cluster is the biggest of all three detected objects. Its location in the plot in Figure 3c is varying slightly, since some parts of the background wall are covered by the moving box from one frame to the other.



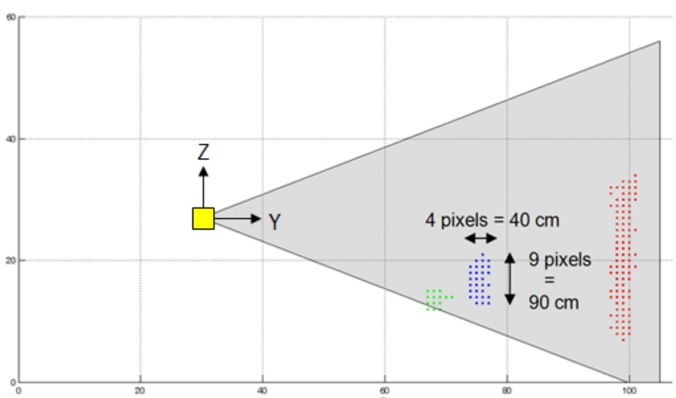
(a) Front view of original scene with propelled box

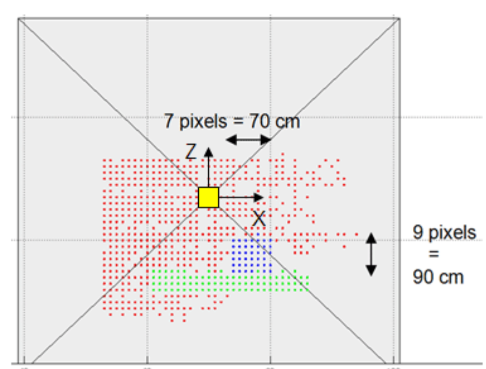




(b) 3D view of single range frame

(c) Plane view of 52 consecutive frames





(d) Elevation view of single range frame

(e) Front view of single range frame

Figure 3: Occupancy grid model of experiment

The grey pyramid in Figure 3b demonstrates the field of view of the range camera. In Figure 3d and 3e the elevation view and front view of a single processed range frame is presented. All voxels generated in the occupancy grid fall into the grey area. The developed algorithm automatically counts the dimensional values (volume) of each clusters as well as the number of voxels each cluster contains. Counting the voxels respectively to each cluster in the following frame allows measuring the position, dimension, direction, and speed of moving objects in a scene.

Modeling Accuracy	Axis	AVG Deviation
Position [m]	X	0.03 (+8.9%)
	Y	0.22 (+4.6%)
	Z	0.01 (+1.0%)
Dimension [m]	X	0.09 (+15.9%)
	Y	0.15 (+40.7%)
	Z	0.11 (+9.9%)
Direction [°]	X	0.3
	Y	1.7
	Z	4.8
Speed [m/s]	X	0.08 (+3.7%)
	Y	0.07 (+5.4%)
	Z	0.01 (+1.1%)

Figure 5: Experimental results at occupancy grid size 0.1m

In Figure 5 the results of the accuracy of the developed occupancy grid at a grid size of 0.1m are presented. Positive deviation values mean that the value measured in the occupancy grid is larger than the real value. The position error of single points is maximal 8.9% off compared to the original position. The dimension of objects varies significantly in y-direction (depth). Since single range cameras capture the face of objects (2½D instead of entire 3D), this large error was expected. Objects are maximum 4.8 degrees of the original path. The speed of the measured object compared to the original velocity is maximal 5.4% higher than the real speed. Calibration and historical measured error values may allow predicting the correct depth value. In summary, the observed average deviation of reality to model is mostly in the positive direction of the axes. This may conclude to have systematic errors which can be eliminated through calibrating the range camera. Indeed, no calibration technique for 3D video range camera is known. As a result, the development of a calibration tool for the range sensor as well as improvements to the experimental environment need to be addressed in future research to avoid the systematic and random error sources described in this paper.

CONCLUSIONS

Three-dimensional modeling of construction environments becomes increasingly necessary for good management. It becomes feasible through modeling approaches based on sparse and dense point cloud algorithms which use range data collection methods based on laser scanning and 3D video range cameras.

This paper demonstrated that real-time detection and tracking of objects in the field of view of a 3D video range camera is possible. This research created computational algorithms based on sparse and dense point cloud approaches and validated in experiments its functionality. A prototype three-dimensional video range camera was used to capture dense range point clouds of object surface geometry. Range data from resources such as humans, equipment, materials, or structures at frame rates above 15 Hz was collected. This research accomplished two tasks: Firstly, it validated distance and position to an accuracy level where the approach can be applied to the construction field, e.g. in an obstacle avoidance systems for safety; secondly, the development and working principle of real-time range data processing techniques and real-time data analyses was verified in extensive experimentation of modeling static and moving resources. This occupancy grid based modeling approach successfully demonstrated complete within the prototype sensor's field of view, accurate, stable, and fast visualization and modeling of complex structures and sites.

ACKNOWLEDGMENTS

This work is supported in part by the National Science Foundation under grant CMS #0409326 and the National Institute of Standards and Technology under solicitation number SB1341-04-Q-0898. NSF's and NIST's support is gratefully acknowledged. Any opinions, findings, conclusions or recommendations presented in this paper are those of authors and do not necessarily reflect the views of the National Science Foundation or of the National Institute of Standards and Technology.

REFERENCES

Bosche, F. and Haas, C.T. (2006). "The Challenges of the Integration of 3D Sensed and 3D CAD Information for Construction Automation", Accepted for publication, *1st CSCE International Construction Specialty Conference*, Calgary, Canada.

CSEM AG (2004). SwissRanger 2 at http://www.swissranger.ch. Accessed October 20, 2004.

Leica Geosystems AG (2006*).* http://www.leica-geosystems.com/media/new/product_solution/HDS4500_25m_and_53m.pdf. Accessed February 15, 2006.

Elfes, A. (1989). "Using Occupancy Grids for Mobile Robot Perception and Navigation." Computer, 22(6), 46-57.

Goodrum, P.M. and Haas, C.T. (2002). "Partial Factor Productivity and Equipment Technology Change at Activity Level in U. S. Construction Industry." J. Constr. Engrg. and Mgmt., ASCE, 128 (6) 463-472.

Kim, C., Haas, C.T., Liapi, K.A., McLaughlin, J., Teizer, J., and Bosche, F. (2004). "Rapid Human-Assisted, Obstacle Avoidance System Using Sparse Range Point Clouds." *9th Aerospace Division International Conference*, ASCE, 115-122.

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 Point Range Clouds." *Automation in Construction*, 13 (1) 67-81.

Teizer, J., Kim, C., Haas, C.T., Liapi, K.A., and Caldas, C.H. (2005). "A Framework for Real-time 3D Modeling of Infrastructure." *Transportation Research Record*, 1913, 177-186.