Real-Time Three-Dimensional Object Detection and Tracking in Transportation

J. Teizer[1], C.T. Haas[2], C.H. Caldas[3], and F. Bosche[4]

[1] Jochen Teizer, ASCE Student Member, Ph.D. candidate, Dept. of Civil, Arch. and Env. Eng., Univ. of Texas at Austin, 1 University Station C1752, Austin, TX 78712; PH (512) 554-7857; FAX (512) 471-3191; email: teizer@mail.utexas.edu
[2] Carl T. Haas, ASCE Member, Professor, Dept. of Civil Eng., Univ. of Waterloo, 200 University Ave West, Waterloo, Ontario, Canada N2L 3G1; PH (519) 888-4567 x5492. FAX (519) 888-4300; email: chaas@civmail.uwaterloo.ca
[3] Carlos H. Caldas, ASCE Member, Assistant Professor, Dept. of Civil, Arch. and Env. Eng., Univ. of Texas at Austin, 1 University Station C1752, Austin, TX 78712; PH (512) 471-6014; FAX (512) 471-3191; email: caldas@mail.utexas.edu
[4] Frederic Bosche, ASCE Student Member, Ph.D. candidate, , Dept. of Civil Eng., Univ. of Waterloo, 200 Univ. Ave West, Waterloo, Ontario, Canada N2L 3G1; PH (519) 888-4567 x3872. FAX (519) 888-4300; email: fbosche@civmail.uwaterloo.ca

Abstract

Flash LADAR (Laser Detection and Ranging) devices are one of the recent technology developments which allow rapid spatial data acquisition of scenes. Algorithms that can process and interpret the output of such enabling technologies into three-dimensional models have the potential to significantly improve work processes. A particular important application in transportation is modeling the location and path of objects to make construction processes safer and more secure. Once people and equipment are detected and mapped into a three-dimensional computer model, their path can be analyzed and access to hazardous areas can be restricted. This paper presents experiments and results of a real-time three-dimensional modeling technique to detect static and moving objects within the field of view of a high-frame update rate laser range scanning device. Applications related to transportation and construction are specified.

Introduction

The ability to locate, describe in 3D, control, and track objects within a field of view has become an important factor in intelligent transportation systems, construction, maintenance, and asset management. Assets usually are somehow scanned and then modeled in 3D at varying frequencies, depending on the application. Asset management may require no more than bi-annual updates, whereas vehicle navigation or construction activities may require real-time updates (Kim and Haas, 2002). An overall framework of data acquisition and 3D model building is presented in Figure 1. It illustrates that in practice, 3D modeling requires combinations of top down design, bottom-up data acquisition, and comparison of both sources of information in many cases for individual assets. Typically, design processes provide well defined information including perfectly parallel, perpendicular, flat, etc. forms (strong forms) like pipes, beams, columns and floors, whereas weak and non-parametric forms are produced from existing infrastructure conditions. Hirschberg (1996) defines "weak forms" as forms related to strong forms, but previous design information was improperly documented. Over time, for example, a rectangle may become an irregular four sided polygon to fit a distorted wooden beam or a cylinder which may grow a joint to represent a bent pipe. Non-parametric forms include wire nets that may represent contour data, polylines that can represent cracks and occupancy arrays or octrees that can represent amorphous volumes or deformed objects. These forms can generally be derived from range point data which contains the distance information in an array of pixels of the original scene image.

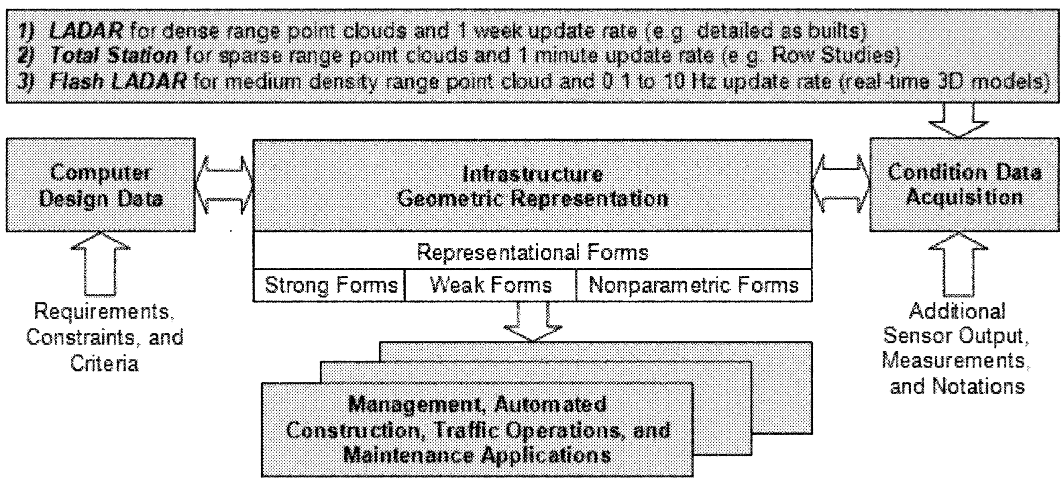


Figure 1. Real-time 3D Modeling Framework (Teizer et al. 2005)

Existing and emerging laser ranging technologies like LADAR (Laser Detection and Ranging), Total Stations, and Flash LADAR can help in the need for fast and accurate geometric modeling. New sensor technology innovations now allow us to address problems of the highest priorities in the transportation and construction area.

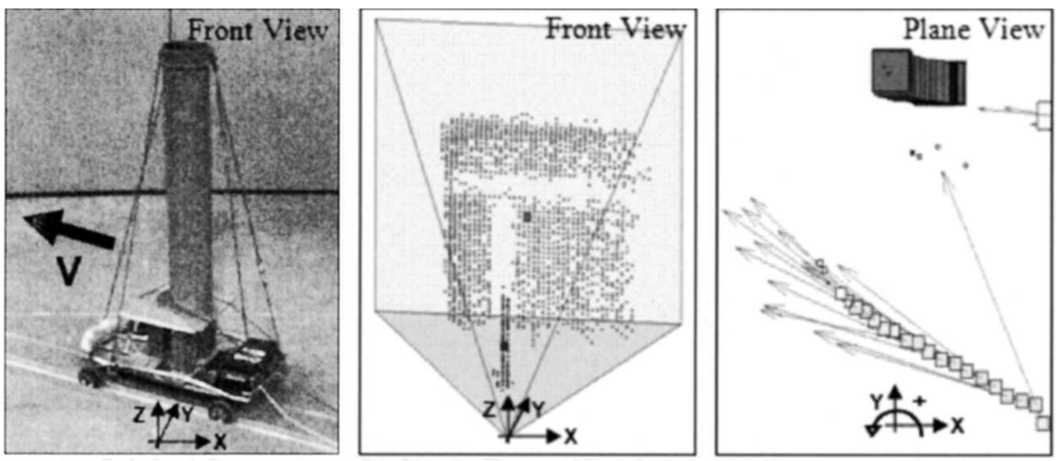
Background

Range scanning methods in 3D modeling can be categorized in three different categories. All are able to capture range information of as-built environments which is later on processed to 3D models. Namely they are dense point cloud approach, sparse point cloud approach, and Flash LADAR approach (Teizer et al. 2005). A dense point cloud approach uses LADAR technology that is able to scan high details of static environments. Due to a complex mirror based scanning system, the acquisition phase of millions of points requires anywhere from 20 seconds up to several hours and does not include the processing phase to convert range data into a 3D model. Moreover, laser range scanning equipment purchase cost is high and easily can reach several hundred thousand US dollars (Stone et al. 2004). A sparse point cloud approach uses selected points to model static environments. Major edges, corners, and lines may already describe an object in a significant manner. Thus, the advantage of manually selecting fewer points with lower purchase cost of mass manufactured commercially available laser range finders and faster range data acquisition in the magnitude of seconds, reduces the overall modeling approach to minutes (Kwon et al. 2004). To overcome the limitations of not modeling in real-time at frame rates above 1 Hz and to advance the applicability to dynamic environments where objects are moving, technology is needed that can address static and moving situations in a rapid manner. Emerging Flash LADAR technology offers to scan static or dynamic environments at high frame update rates as well as provides a dense point cloud required for accurate real-time three-dimensional modeling.

Flash LADAR Approach, Experiments, and Results

Conventional laser range scanning methods are based on interference and triangulation techniques as well as of time-of-flight measurements. We used emerging technology based on the time-of-flight principle that uses a CMOS sensor and pulsed active illumination to generate a 3D array of pixels (image). To each pixel on-board the sensor chip, the system measures the phase shift between a reference signal sent directly to the detector's smart pixels and light reflected by the scene. Since the measured wave is cyclic, the distance measurement is not unambiguous. Under the current prototype development of the sensor, each pixel stores information values of range and intensity to objects contained in a scene at the unambiguous range of 7.5m. Changing the modulation frequency of 20 MHz increases the acquisition range, but reduces the illuminated power send out into the scene in form of an intensity-modulated wave front, ultimately limiting the accuracy of range measurements. Technological improvements are under way and will use higher power laser diodes instead of conventional near-infrared LEDs, or focus on applying two different modulation frequencies. As a result, the current inexpensive technological development stage limits the potential use to important feasibility studies rather than commercial end product driven solutions in real-time 3D modeling.

This research focuses on the application of detecting and tracking objects in the field of view of a 3D range camera. Rapid range measurements in transportation construction are needed, to navigate vehicles safely in paths were potential hazards and obstacles are avoided. We have conducted 27 indoor experiments to validate the accuracy of range measurements as well as to test our algorithm development for processing range information. The sensor's field of view (horizontal 42° and vertical 46°) has a resolution of 124x160 pixels, thus achieving an angular accuracy of 4.6cm at the furthest distance (7.5m). Once range data is acquired and stored in an array, our processing algorithm based on occupancy grid system is applied to the range information. Grouping range points into voxels, filtering noise measurements, and applying a hierarchical clustering technique reduced data significantly to define features in each frame. Comparing features in each cluster such as volume, speed, direction, and location in consecutive frames, allowed separating static from moving objects. Since the location to all centroids in the clusters where known, velocity vectors were applied. Figure 2 demonstrates the process of building an occupancy grid from an original scene. In the example, we used a cart with a mounted pipe to simulate a moving vehicle (Figure 2a). Guided on rails at angles of 0/90/150 degrees to the X-axis, range frames were collected from static and moving boxes and pipes at frame update rates of 15.2 Hz. The processing algorithm converted each range frame in real-time (greater than 1 Hz) into voxel based models. In Figure 2b the pipe is represented by one cluster and also the background wall; the ground and cart are out of the sensor's field of view. Further analyzes classified features such as centroid locations, direction of movement, and velocities (Figure 2c). On average the errors in accuracy were: Object position +4.5% (positively off in all axis directions X, Y, and Z), object dimension +30% (half of which is due to the fact that line-of-sight prevents from modeling the depth Z, thus only 2½D instead 3D), direction -1.2° (further away than in reality), and speed +6.5% (faster than in reality). Errors most likely stem from measurement errors, but can also partially come from systematic errors. The results are promising and show that real-time 3D detection and tracking of objects is feasible.



(a) Original Scene (b) Single Frame 3D Model (c) Centroids of all Frames

Figure 2. Real-time 3D Shape and Path Modeling of a Simulated Vehicle

Future research steps will be extensive testing of the algorithm in outdoor environments. Initial tests indicate that the algorithm works successfully, however, due to the prototype stage of the sensor, it physical working principle does not allow full sunlight on the reflecting scene. Since sunlight is 400 times brighter than existing LEDs, power emitted by the sensor will not be reflected. For this reason, we successfully tested our algorithm in cloudy weather. We add at this point that the next generation Flash LADAR (to be released in early 2006) will significantly increase the range of non-ambiguity by the use of two or more frequencies and supply active background light suppression that will permit measurements in full sunlight.

Applications in Transportation

Many application areas for 3D cameras exist. One predominant area of current interest in range cameras for transportation is their use in safety. In 2003 the US experienced 42,643 deaths from motor vehicle crashes. The mileage death rate dropped from 4.88 deaths per million miles traveled to 1.56, and deaths from 100,000 population dropped from 26.8 to 15.4 (O'Neill, 2005). Nevertheless, the overwhelming bulk of serious crashes stems from drivers breaking the traffic laws or not paying attention. Brian O'Neill argues that scientific evidence has shown that more education does not change the drivers' behavior. Instead progress can be achieved through stricter law enforcement and political action on state level to create incentives for the use of existing or emerging technologies which can improve safety. The Federal Highway Administration (FHA) and National Highway Safety Administration (NHSA) have eased the requirements to remove hazardous road design features and upgraded safety on roadsides under renovation, by establishing new standards for roadside clear zones, breakaway lights and poles (O'Neill, 2005).

Continental Automotive Systems (2001) states, that 40% of all road casualties can be prevented once the entire response time of a vehicle breaking system can be reduced by 0.5 seconds. The human judgment of analyzing a road scene and applying force to the break system of a vehicle is at around 1 second. We believe that assisting the human eye with a real-time 3D sensor based vision system can help make aware of upcoming hazardous situations, since the response time can be reduced to the minimum time a sensor needs to acquire, process, and analyze a frame shot. Once a danger appears in the vehicle's path, a Flash LADAR device could reduce the entire response time to better than 0.1 seconds (greater than 10Hz). This largely depends on developing range sensing technologies to a commercial level where longer distances in outdoor environments and faster processing algorithms of dense point cloud data are fully developed. The car industry, for example, is working on a system for the robust measurement of the distance between two vehicles.

More applications exist in the vehicle guidance, such as the assistance of the drivers' perception in rear or side parking situations by automated scanning of confined spaces, slower inflation of airbags when persons are close to the dashboard, steering wheel, or airbags that should not inflate once baby seats are installed. Similar to belt use, three-dimensional laser scanning technology and processing can help to reduce injuries and fatalities in pedestrian/car related accidents. Running red lights becomes one of the most dangerous situations at intersections when pedestrians or bicycles are involved. Range scanning of entire or significant street areas might detect and track the movement of pedestrians or bicyclists who willingly, by fault, or by ignorance enter crosswalks or change lanes. Alarms can be set off (in cars or at pedestrian traffic lights) to prevent impacts or help in making a decision to react for a best-case scenario. Other application areas are in the robotics industry where automated robots need to navigate autonomously and, in the security industry, inexpensive measurement systems are being planned for the surveillance of complete rooms. Observation systems may even assist the prioritization and scheduling of elevator transportation (Opto & Laser Europe, 2004)

In construction related to transportation, the industry reported in 2005 that 23% of all 1,124 fatalities were related to transportation and 22% had contact to objects and equipment (BLS, 2005). These figures indicate that close to one half of all accidents in construction are related to transportation where vehicles, such as cars, trucks, or heavy equipment, apply deadly forces to workers. By modeling workspace in real-time and in 3D, the operator's perception can be improved, safety work zones can be defined and restricted access can be granted. Equipment or robots that track or handle materials can be equipped with Flash LADAR devices to safely navigate construction sites. Flash LADAR sensing can be applied in simulating work processes by comparing 3D as-built models vs. as-planned data. Real-time 3D modeling further can help to monitor health progress in life-cycle management for transportation infrastructure, such as inspection processes for defect detection.

Conclusion

We have presented in our research the general feasibility of using real-time three-dimensional modeling as a detection and tracking tool for static and moving objects. Efficiently and rapidly built 3D models after the occupancy grid principle allowed to acquire, convert, and store range information data from a high frame update rate laser range camera. Since this research approach includes the dimension of time, tracking and path planning for static and moving objects from a dynamic platform or vehicle becomes feasible. This real-time 3D modeling technique offers a wide field of potential applications in transportation applications related to safety, such as vehicle navigation and pedestrian protection. Other application areas in the field of transportation exist and may be pursued in the future by modeling infrastructure, tracking materials, detection of structural defects, or life-cycle monitoring of assets.

References

Bureau of Labor Statistics (2005). http://www.bls.gov/iif/oshcfoi1.htm (Dec. 8, 2005).

Continental Automotive Systems (2005). "Nach dem Bremsweg wird der Reaktionsweg kürzer." http://www.conti-online.com (June 10, 2005)

Hirschberg, U., and Streilein, A. (1996). "CAAD meets Digital Photogrammetry: Modeling 'Weak Forms' for Computer Measurement." *Automation in Construction*, 5, 171-184.

Kim, Y.S., and Haas, C.T. (2002). "A Model for Automation of Infrastructure Maintenance Using Representational Forms." *Automation in Construction,* 10(1), 57-68.

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, 67-81.

National Highway Safety Administration (2005). http://www.nhtsa.gov (Dec 1, 2005).

Stone, W.C., Juberts, M., Dagalakis, N. Stone, J., and Fronczek, C. (2004). *Performance Analysis of Next-Generation LADAR for Manufacturing*, Construction, and Mobility, NISTIR-7117, Gaithersburg, MD.

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

O'Neill, B. (2005). "Improving U.S. Highway Safety – Have We Taken the Right Road?" TR News, 239, 24-27.

Opto & Laser Europe (2004). http://optics.org/articles/ole/9/12/2/1 (June 9, 2005).