Ground Truth Data Using 3D Imaging for Urban Search and Rescue Robots

Nicholas A. Scott, Alan M. Lytle
National Institute of Standards and Technology
100 Bureau Drive Stop 8611
Gaithersburg, MD, 20899-8611, USA
{nicholas.scott, alan.lytle}@nist.gov

ABSTRACT

The National Institute of Standards and Technology (NIST) is leading an effort to develop performance standards for urban search and rescue robots (US&R). An important component of developing performance standards for these robots is capturing ground truth data that represents the geometry of the robot operating environment. This paper describes two ground truth data collection efforts conducted in 2006 and 2008 at the Texas Engineering Extension Service Disaster City training facility in College Station, Texas. Several indoor and outdoor training scenarios were captured with 3D imaging systems and the data is now publicly available through NIST to support research and development of robotic technologies for the US&R domain.

Keywords

robotics, ground truth, 3D imaging, urban search and rescue

1. INTRODUCTION

The National Institute of Standards and Technology (NIST) is leading an effort to develop performance standards for urban search and rescue (US&R) robots [5]. As part of this effort, NIST organizes events that allow emergency responders, robot manufacturers and robotics researchers to work shoulder-to-shoulder within world-class responder training facilities [4]. Commercial-off-the-shelf products and laboratory prototypes are operated by responders in realistic operating scenarios while being observed and supported by the technical experts. These events allow responders to better understand state-of-the-art robot capabilities and limitations, and provide manufacturers and researchers unfiltered access to subject matter experts. An additional activity at these responder events is the exercise of various test methods and performance metrics under development to support the overall goal of creating a suite of performance standards for US&R robots. An important component of developing performance standards for these robots is capturing ground truth data of the training scenarios and the test methods.

PerMIS'09 September 21-23, 2009, Gaithersburg, MD, USA
ACM 978-1-60558-747-9/09/09.

This paper describes two ground truth data collection efforts conducted in 2006 and 2008 at the Texas Engineering Extension Service (TEEX) Disaster City training facility in College Station, Texas. Several indoor and outdoor training scenarios and two test methods were captured with 3D imaging systems[1] and the data is now publicly available through NIST to support research and development of robotic technologies for the US&R domain.

This paper begins by describing the motivation behind the data collection efforts. Section 3 provides information regarding the captured scenarios and Section 4 presents metrics for the data collected. Finally, Section 5 discusses future NIST efforts using this data.

2. MOTIVATION

When a disaster occurs, previously benign terrain may become difficult or impossible to traverse. Buildings collapse, roads and bridges are destroyed, and previously smooth, obstacle free terrain may contain large obstacles and discontinuities. In order to perform search and rescue operations, responders must asses the terrain in order to employ assets that posses the correct mobility techniques to get to desired locations. For responders to effectively use robotic technologies on US&R missions, they must understand how different robotic platforms perform in diverse terrain. Developing tests and performance metrics to enable this understanding not only supports the responder's use of robots in the field, but also provides important information to support further research and development.

An essential element in defining these performance metrics is independently capturing an accurate ground truth representation of the robot's operating environment. This ground truth data can support a wide range of research including mobility performance metrics, terrain characterization, mapping algorithm evaluation, and virtual environment construction.

For US&R robotics, both qualitative and quantitative measures of the environments in which platforms are tested and deployed to support mobility performance metrics and terrain characterization are of great interest. For examples of qualitative measures of an environment, consider trail rating systems for ski slopes or the Beaufort Wind Force Scale for estimating wind speed from sea state. Quantitative US&R terrain characterization metrics would enable predictable and consistent ways of representing difficult ter-

[1] A 3D imaging system is a non-contact measurement instrument used to produce a 3D representation (for example, a point cloud) of an object or a site [1].

rain (e.g., rubble) and provide fair comparison of platforms. An example of quantitative metrics in the US&R context could be a specific measure of the traversibility of the terrain surface derived using techniques such as height, slope, and roughness estimation from plane fitting, fractal dimension analysis or wavelet energy statistics. Traversibility is a well-studied discipline, particularly in the context of unmanned ground vehicle path planning. The challenge is to standardize a universally accepted measure for US&R robot evaluation.

In addition to understanding the terrain, ground truth data supports mapping algorithm evaluation and virtual environment construction. More and more developers are including map creation in their operator toolkits, and map generation is a highly desired capability amongst responders. The ground truth data collected can serve as a baseline for evaluating the performance of 2D and 3D mapping systems deployed on mobile sensors. Ground truth data used to evaluate mapping can also be used in simulation environments. NIST, along with partner organizations, is investigating how to represent the point clouds and/or derivative terrain models within simulation environments such as USARSim [2]. Importing point, polygonal, or surface models of realistic training scenarios into simulation systems can make the training scenarios themselves accessible to a wider set of developers. Responders, researchers, developers, and other interested personnel will be able to navigate the scenarios, to some degree of fidelity, without having to physically travel to the location. Intelligent behaviors for semi-autonomous robots can also be virtually tested within the models.

3. DATA COLLECTION

To support the need for ground truth data, NIST researchers gathered high-resolution 3D image data for five training scenarios and two test methods at Disaster City, the TEEX National Emergency Response and Rescue Training Center in College Station, Texas. Disaster City is a 52-acre site that provides full-scale collapsible structures, rubble piles, and wrecked transportation structures for search and rescue training and is considered by many to be the most comprehensive emergency response training facility presently available. The ground truth data was collected during two separate NIST organized events at Disaster City.

3.1 Collection One

The first data collection effort was held on April 4-6, 2006 and focused on outdoor US&R environments. Data for three different scenarios was collected:

- Concrete Rubble Pile
- Wood Rubble Pile
- Passenger Trains

The concrete rubble pile scenario is depicted in Fig. 1. This scenario simulates a fully collapsed concrete structure with interior voids. The rubble pile primarily consists of concrete and rebar. Large concrete slabs, barriers, and pipes, generally several meters in length, provide the support for many of the voids. Small concrete rocks, typically 30 cm to 50 cm in diameter, fill-in the the space around the larger

Figure 1: Concrete rubble pile training scenario which simulates a fully collapsed concrete structure.

concrete pieces. Rebar and other metal structures are scattered throughout the rubble pile. Robots are deployed from the perimeter either directly into subterranean voids or over top of the rubble pile to search for victims and map the area.

The wood rubble pile scenario is shown in Fig. 2. This

Figure 2: Wood rubble pile training scenario which simulates a fully collapsed concrete structure.

scenario simulates a fully collapsed wood structure with interior voids. The rubble pile consists of several meter length wood planks and wood pallets. Interior voids are created by using several meter sections of concrete piping. Robots are deployed from the perimeter of this pile by climbing, throwing, or launching into the central area to look for victims and map the area.

The passenger trains scenario is depicted in Fig. 3. This scenario mimics the collision and partial derailment of passenger rail cars and industrial hazardous material tanker cars carrying an unknown substance. Robots are deployed from the perimeter of the wreck and circumnavigate the trains, tracks, and rubble to map the perimeter of the scene and determine the location of each car. The underside of the elevated car and the interior of the car on its side is explored for victims and to look for placards describing what hazardous material is onboard.

3.2 Collection Two

The second data collection effort was held on November 17-21, 2008 and focused on indoor US&R environments.

Figure 3: Passenger trains training scenario which simulates the collision and partial derailment of passenger rail cars.

Data for two training scenarios and two test methods was collected:

- Single Family Dwelling
- House of Pancakes
- Theater Maze
- Tube Maze

The single family dwelling scenario is depicted in Fig. 4. This scenario simulates a partially collapsed single family

Figure 4: Single family dwelling training scenario simulating a partially collapsed home. The top picture shows the area inside one of the rooms of the building and the bottom picture shows the building from the outside.

home due to an earthquake. Entrances to the building and doorways between rooms are compromised with many either fully or partially blocked. The floors are scattered with debris from the concrete structure and furniture. The ceiling is mostly collapsed in one room and there is a large breach in the floor in another room. There is also a basement accessible from the outside through a long set of stairs to a welled exit. Robots are deployed into the building to identify victims, hazards, and all entrances and exits to inform responders of the situation.

The house of pancakes scenario is depicted in Fig. 5. This

Figure 5: House of pancakes training scenario simulating a partially collapsed concrete building. The top picture shows the main area inside with the collapsed sloped roof and the bottom picture shows the building from the outside.

scenario mimics a partially collapsed concrete building of unknown use. The roof structure is collapsed on one side of the building causing the roof to angle downward such that it is almost in contact with the ground. The interior of the building contains various wood and concrete structures as well as office desks and tables. Robots are deployed into the building to search for victims and map the environment.

The theater maze test method is depicted in Fig. 6. This

Figure 6: Theater maze test method which tests the ability of a robot to navigate a complex environment without getting lost.

environment tests the ability of the robot and its operator to fully explore a complex environment for victims and identify standard hazardous material placards without getting lost. The maze consists of rolling wooden floor planks on a slope and 2.44 m tall wooden walls.

The tube maze test method is depicted in Fig. 7. This

Figure 7: Tube maze test method which tests the mapping ability of US&R robots in an environment littered with occlusions.

environment tests the mapping ability of robots in an environment full of occlusions. The flooring is made of angled sheets of wood. Raising up from the flooring are PVC pipes of varying height that provide occlusions for this environment.

4. DATA SETS

The 3D image data was collected using commercial laser scanners. A laser scanner is a 3D imaging device that uses a laser to measure the distance to an object. The laser beam is scanned both horizontally and vertically over time to image the operator-designated field of view. The distance, azimuth, and elevation information collected from each measurement in the scan is used to create high-resolution point clouds containing hundreds of thousands of points for a single scan.

Two different laser scanners were used in the two data collection efforts. The data for collection effort one was collected using a pulse-based time-of-flight laser scanner. The manufacturer specifies a range uncertainty of 7 mm and a point uncertainty of 12 mm at 100 m range for this instrument. Collection effort two data was collected using a phase-based time-of-flight laser scanner. For this instrument, the manufacturer specifies a range uncertainty $\leq$ 6 mm for ranges up to 50 m. A point uncertainty was not specified for this instrument.

4.1 Sample Data and Metrics

Figures 8 and 9 show screen captures of scenes generated in point cloud software for the tube maze test method. Each point is colored based on the intensity of the laser return. Within the software, camera viewpoints can be changed to examine the 3D data from multiple viewing angles and measurements such as point-to-point distance can be readily determined.

Figure 8 shows an elevated view of the point cloud data for a single scan of the tube maze. Since laser scanners are line-of-sight instruments, a single scan is unable to capture the entire environment when there are occlusions in the scene. Occlusions cause "shadows" of missing data where the laser scanner cannot sense. By design, the tube maze contains many occlusions and areas of missing data (shown in black) are prevalent in the individual scans.

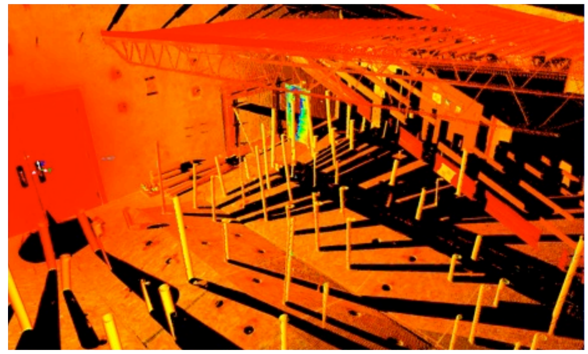

Figure 8: An elevated view of the point cloud data for a single scan of the tube maze test method. Occlusions cause "shadows" of missing data.

To fill in the missing data, scans are taken from multiple locations around the scene. Individual scans are then merged through a process called registration to create complete point clouds of the scenes. While some of the scans required manual registration, most of the scans were registered using stationary targets placed in the scene to provide common points of reference to register the scans. The data was registered and segmented using commercial software tools. Figure 9 shows the complete registered point cloud data set for the tube maze test method from a similar viewpoint.

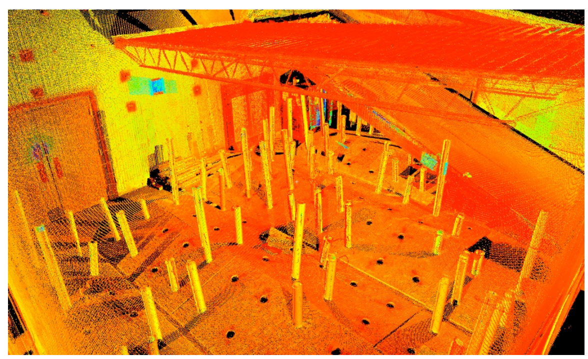

Figure 9: An elevated view of the fully registered point cloud data for the tube maze test method.

The number of points collected for each scenario is given in Table 1 and the number of scans is given in Table 2.

4.2 Data Availability

All of the 3D image data outlined in this paper is available free of charge to the public. Please contact the authors to obtain any data of interest. In the near future, the data will be made available through the NIST website.

5. FUTURE WORK

Stepfield pallets are a fabricated and repeatable terrain for evaluating robot mobility [3]. As a first step towards developing terrain traversibility metrics, NIST researchers will use the ground truth data presented in this paper to investigate the design of a multi-unit stepfield approximation of a representative segment of one of the Disaster City rubble

Table 1: The number of points for each scenario and test method data set from collection efforts one and two.

Collection	Scenario/Test Method	# Points (Millions)
1	Concrete Rubble Pile	5.77
	Wood Rubble Pile	7.75
	Passenger Trains	4.87
2	Single Family Dwelling	848.35
	House of Pancakes	2550.60
	Theater Maze	12.62
	Tube Maze	296.90

Table 2: The number of scans for each scenario and test method data set from collection efforts one and two.

Collection	Scenario/Test Method	# Scans
1	Concrete Rubble Pile	45
	Wood Rubble Pile	23
	Passenger Trains	41
2	Single Family Dwelling	26
	House of Pancakes	29
	Theater Maze	2
	Tube Maze	10

piles. If this is achieved, the existing mobility metrics captured for the stepfields can be applied to predict how well a given mobility platform will perform in the rubble pile scenario and more generally, any terrain that can be modeled in this fashion.

The ground truth data collected in this work will also be used to explore methods for evaluating the quality of a maps generated by the US&R robots which traversed the same scenario environments. While there is currently some work being investigated for evaluating 2D maps, there is little work being done for 3D maps.

Finally, the data will be used to support the modeling of the Disaster City scenarios for use in virtual training and testing environments such as USARSim.

6. ACKNOWLEDGMENTS

The authors would like to thank Adam Jacoff and Christopher Scrapper whose support made this work possible. Also, a hearty thanks to Gerry Cheok, Kamel Saidi, and Nicolas Philippon for their tireless work in helping plan and execute these data collection efforts.

7. REFERENCES

[1] ASTM. *ASTM E2544 - 09b Standard Terminology for Three-Dimensional (3D) Imaging Systems*. ASTM, 2009.

[2] S. Balakirsky, F. M. Proctor, C. J. Scrapper, and T. R. Kramer. An integrated control and simulation environment for mobile robot software development. In *Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conferences*, August 2008.

[3] A. Jacoff, A. Downs, A. Virts, and E. Messina. Stepfield pallets: Repeatable terrain for evaluating robot mobility. In *Proceedings of the 2008 Performance Metrics for Intelligent Systems (PerMIS) Workshop*, 2008.

[4] A. S. Jacoff and E. R. Messina. Dhs/nist response robot evaluation exercises. In *Proceedings of the 2006 IEEE International Workshop on Safety, Security and Rescue Robotics (SSRR)*, January 2007.

[5] A. S. Jacoff and E. R. Messina. Urban search and rescue robot performance standards: Progress update. In *Proceedings of the 2007 SPIE Defense and Security Conference*, June 2007.