DEVELOPMENT OF A PROBABILISTIC SENSOR MODEL FOR A 3D IMAGING SYSTEM

Alan M. Lytle
National Institute of Standards and Technology
100 Bureau Drive
Gaithersburg, MD 20899-8611
alan.lytle@nist.gov

ABSTRACT

The use of high frame-rate 3D imaging systems for construction equipment control is part of a broader research effort at the National Institute of Standards and Technology investigating innovative technologies for automated construction, performance and evaluation of 3D imaging systems, and construction object recognition and tracking. A probabilistic measurement model is desired for these imaging systems to better understand and describe the instrument performance, and to create synthetic data sets for algorithm development, training and testing. The initial development of a probabilistic sensor model and 3D image simulator based on an open-source graphics package is presented in this paper.

KEYWORDS

3D Imaging System, Flash LADAR, 3D Range Camera, Sensor Model.

1. INTRODUCTION

The NIST Construction Metrology and Automation Group is conducting ongoing research to provide standards, methodologies, and performance metrics that will assist the development of advanced systems to automate construction tasks. Research efforts in this project have produced an automated construction testbed that includes a 6 degree-of-freedom robotic crane (the NIST RoboCrane), laser-based real-time tracking, ultra-wide band (UWB) tracking, radio-frequency identification (RFID), and both high-resolution and high-frame rate 3D imaging sensors (i.e., construction-grade laser scanning and range cameras)[1].

An initial demonstration project involved autonomous assembly of a multi-component test steel structure using the robotic crane with pose tracking provided by a laser-based site measurement system and assembly scripts generated from a commercial 4D CAD package. More recent work includes the addition of a prototype site visualization system – JobSight – that allows monitoring of site sensor data, the current position and planned action of robot agents, and other object data such as CAD models, point clouds, etc. Figure 1 (a) and (b) depicts RoboCrane and JobSight, respectively. Additional information can be found in [1].

In the first demonstration project, there were no exteroceptive sensors[2] mounted on the crane. The next effort involves the use of a high-frame rate (approximately 30 Hz) low-resolution (160 pixels x 124 pixels) 3D imaging system (range camera) as an obstacle avoidance and docking sensor

---

[1] Certain trade names and company products are mentioned in the text or identified in an illustration in order to adequately specify the experimental procedure and equipment used. In no case does such an identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily the best available for the purpose.

[2] A newer model of this range sensor is available. Examination of the next-generation sensor is an element of future work

Target recognition objects are structural steel elements.

In probabilistic robotics, a percept action is characterized as arriving at a conclusion about the world given a sensor measurement, a measurement model, and prior state information. This problem is usually not directly solvable, and instead Bayesian inference is applied to produce the best estimate of the solution given a certain sensor model. The initial development of a range camera sensor model to use in 3D construction object recognition and tracking is the focus of this research project.
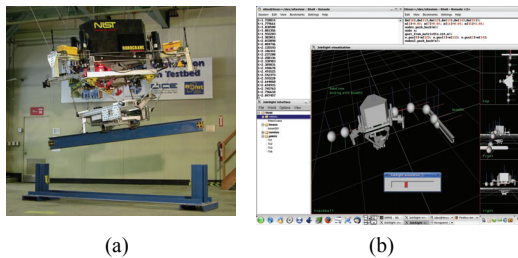

(a)   (b)

Figure 1 (a) RoboCrane Carrying Element of Test Structure and (b) Screen shot of the Job Sight visualization software

2. RELATED WORK

Zheng et al. [2] present a flash LADAR (laser detection and ranging) object-recognition system which fit 10 pixel x 20 pixel range images to a template generated from a laser physics based simulation. The algorithm used a projection-based filter to trim the candidate pool and then used both pixel matching (shape) and silhouette matching (boundary) for recognition. The system was trained using synthetic image data for seven military vehicle targets developed from a LADAR simulation package.

The use of commercial models and commercial graphics animation software for developing automatic target recognition algorithms is discussed in [3]. The use of the open source Persistence of Vision raytracer (POV-Ray) [4] and the unofficial MegaPOV extension [5] to develop 3D data for generating and testing 3D video communication systems is reported in [6], though implementation details are not provided. An example of a high-fidelity, physics-based LADAR

simulation environment used for hardware design is provided in [7].

3. MEASUREMENT MODEL

3.1. Terminology

The following terms are provided as a basis for communication within this paper.

Range uncertainty: The uncertainty associated with a distance measurement reported by an instrument to a given object of interest (or target). Uncertainty is a parameter that characterizes the dispersion of the measurement values that could reasonably be attributed to the measurand. Range uncertainty for active 3D imaging systems varies as a function of target range, target reflectivity, target angle of incidence, and sensor azimuth.

Beam spot size: The size of the light or laser beam as it hits a plane perpendicular to its travel path. Because of beam divergence, the beam spot size is dependent on, and should specify, the distance beween the target and the beam source. For example, the beam spot size is 50 mm at 100 m.

Mixed pixels: Errant 3D data resulting from the way most instruments process multiple returns. These multiple returns occur when a laser beam hits the edge of an object and the beam is split. In this case, part of the beam is reflected by the object while the other part continues and may be reflected by another object beyond. The measured reflected signal therefore contains multiple range returns, and typically the reported range measurement for that particular ray vector is an average of those multiple returns.

3.2. Sensor Description

The range camera used for these experiments is a commercially available, phase-based AM-homodyne device that uses a bank of 870 nm IR LEDs (infrared light emitting diodes) for the illumination source. The source is modulated at 20 MHz, yielding an unambiguous range of 7.5 m. The focal plane array (FPA) is 160 pixels x 124 pixels with a corresponding nominal field of view (FOV) of 43° (h) x 46° (v). Data returned from the device include range, intensity, and amplitude per pixel at a frame rate of approximately 30 Hz. Settings for the integration time, amplitude

threshold, and distance offset are user-controlled. The sensor is shown in Figure 2.

Figure 2 The Range Camera

A summary of manufacturer's specifications is provided in Table I.

Table I Manufacturer's Specifications

ATTRIBUTE	VALUE
Frame Rate	30 Hz
Range	7.5 m
Resolution	< 1 cm
Field of View	43˚ x 46˚
Pixels	160 x 124

3.3. Sensor Evaluation

The instrument range uncertainty was determined as described in [8]. Measurements of an 18 % reflectivity, 40.5 cm (16 in) x 25.4 cm (10 in) planar target were taken at ranges of 1 m through 7 m at intervals of 1 m, at a zero angle of incidence. All tests were conducted indoors under ambient lighting (fluorescent) conditions. Ranges to the targets were measured using a laser distance meter with an uncertainty of ±1.5 mm (1σ) mm over a range of 0.2 m to 200 m. At each range N = 1000 images were obtained and range data were sampled from a 7x3 pixel region at the center of the targets. The sampling was repeated two additional times at each range for a total of three sets per range increment. The results of the range error measurement are shown in Figure 3.

All samples are with the target angle of incidence = 0° and target reflectance = 18 %. There are three sets for each range increment. Plotted points are offset slightly for visibility. Error bars indicate uncertainty (± 1 σ).

Although there is an oscillatory behavior to the mean range error, for simplicity in this first iteration of the sensor model the error is fit as a linear function given by:
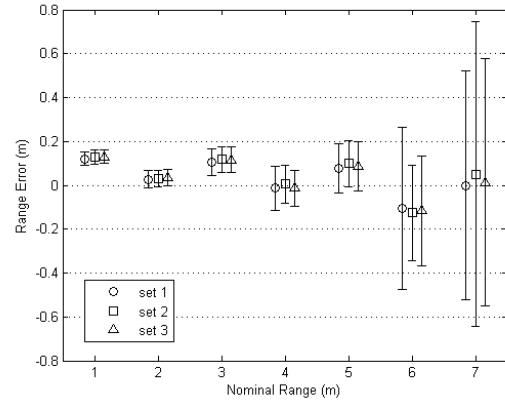

Figure 3 Range Uncertainty test results for N=1000 image samples and 7x3 pixels sampled per range image.

$$\mu_{Error} = -0.027R + 0.16 \qquad (1)$$

where

μ_{Error} = mean range error

R = range

The standard deviation of the range error σ_{Error} is given by:

$$\sigma_{Error} = 0.0025R^4 - 0.031R^3 + 0.13R^2 + 0.22R + 0.14 \quad (2)$$

3.4. Sensor Model

Thrun et al. [9] presented a probabilistic measurement model of a beam finder that is a mixture of four types of errors. This mixture model includes:

(1). Normal distribution of a correct range measurement with noise

(2). A uniform distribution at the max range of the sensor for a failure to intersect any objects

(3). An exponential distribution for shorter than expected range measurements due to unexpected objects

(4). A uniform distribution over all ranges for random returns.

Figure 4 is a representation of the four-component measurement model. Of note, only components (1) and (2) of the mixture model were incorporated

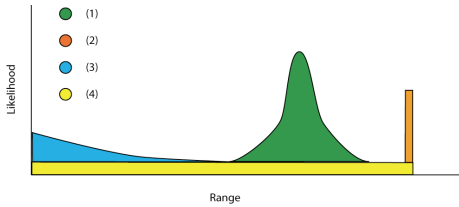into the synthetic data generator[3]. Components (3) and (4) require additional experimental data.

Figure 4 The Sensor Model (adapted from Thrun et al.)

4. SYNTHETIC 3D IMAGE GENERATION

4.1. Motivation and Design Criteria

A system for generating synthetic 3D image data from CAD models is needed to systematically evaluate and test various algorithms for model-based construction object recognition and tracking research. The ability to vary the noise levels in the synthetic data and generate representative noisy data sets would not only significantly aid in algorithm development, but would also provide data-driven recommendations to manufacturers for future system requirements.

The primary performance criterion for the system was the ability to generate 3D image data from CAD models that was representative of the type of data that would be provided from the tested instrument. Specifically the FOV (field of view), range, and pixel field had to be consistent with the range camera parameters. The mean range error and variance (both as a function of range) also had to be incorporated into the system model. Finally, the ability to simulate the mixed pixel effect was highly desired, as this is a significant contributor to the system noise.

4.2. Implementation

The system consisted of four basic elements: a scene generator, a raycasting function, a Monte Carlo simulation for generating modeled noise, and a mixed pixel simulation. In all cases, specific efforts were made to use existing software to meet base-level capabilities where available.

MegaPOV (an unofficial modified version of POV-Ray) was chosen as the raytracing application. POV-Ray is an open source raytracer with a large community following. MegaPOV is an extension of POV-Ray with numerous post-processing features, one of which is the creation and storage of a depth map from the ray trace data. Unfortunately, the depth data from MegaPOV only provides 8-bit resolution (e.g., ≈ 3 cm for 7.5 m). VLPov [10], which is an annotation patch developed at UCLA for MegaPOV, was used to provide increased depth resolution. The VLPov package also provided MATLAB scripts for reading and processing the depth map data. The disadvantage of using the VLPov patch was that it necessitated recompiling MegaPOV and prevented the use of the application's graphical user interface.

The CAD scene was generated using a commercially available software package that could export to the POV-Ray scene description format. Of note, this is a common capability with most current CAD packages. The camera parameters in the scene description file were then manually edited with the range camera parameters. The desired resolution was set with an initialization file. In this case, a resolution of 804 (width) x 624 (vertical) was used. This was a fivefold increase in the resolution required for single raycasting (160 x 124) and allowed the use of a 9 x 9 kernel to simulate the beam spot size.

The binary depth map resulting from the MegaPOV rendering was read and initially processed with the VLPov Matlab files. The depth matrix was then modified to represent the max range effect by substituting 7.5 m for all pixels without object intersections, and then reduced to 160 x 124 using the kernel (in this case using an equal weighting from each of the 81 rays). This averaging of the rays produced a rudimentary mixed pixel effect.

The data were then adjusted with additive Gaussian noise using the mean error (1) and standard deviation (2) functions derived from experimental data. The resulting point cloud was then imported to the original CAD package for viewing and comparison with the original model.

---

[3] Since the range camera is a phase-based system with only one modulation frequency, there is the additional factor of errant range returns for objects greater than the unambiguous range (7.5 m) of the system. This effect was not modeled

4.3. Sample Data

Sample data sets were generated using a CAD model of an existing 3D imaging artifact that is used in testing the performance of different 3D imaging systems. The star artifact is shown in Figure 5.

The CAD model was rendered using the synthetic 3D image generator, and point clouds were produced with the following parameters:

(1) Single ray for each range element – no noise

(2) Single ray for each range element – with noise

(3) Multiple rays (81) for each range element – no noise

(4) Multiple rays (81) for each range element – with noise

Figures 6 through 8 depict the point clouds generated from each of those parameter settings.
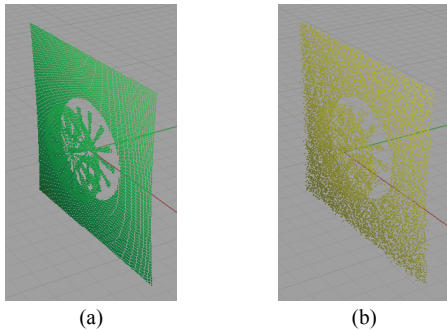

Figure 5 The Star Artifact

(a)   (b)

Figure 6 Star Artifact at 2 m from Range Camera, Single Ray for Each Range Element, (a) No Noise, (b) Noise
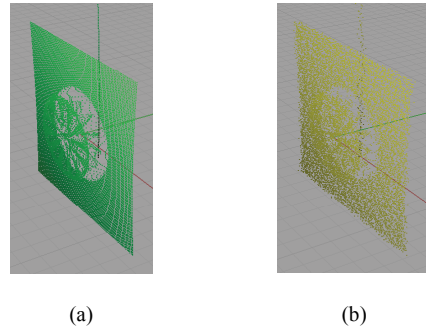

(a)   (b)

Figure 7 Star artifact at 2 m from Range Camera, Multiple Rays for Each Range Element, (a) No Noise, (b) Noise, (note mixed pixels between star and wall)
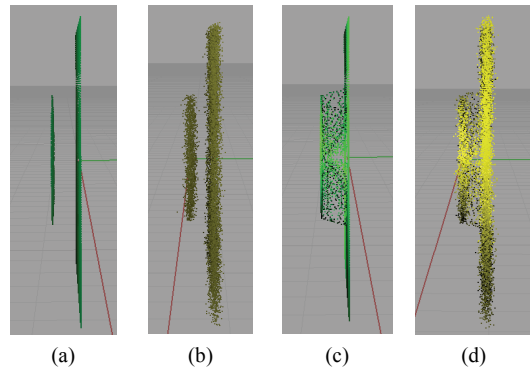

(a)   (b)   (c)   (d)

Figure 8 Star Artifact (profile) at 2 m from Range Camera, Single Ray for Each Range Element, (a) No Noise and (b) Noise Multiple Rays for Each Range Element, (c) No Noise and (d) Noise

5. CONCLUSION

This report presents initial results in the development of a probabilistic sensor model for a high frame rate, noisy, 3D imaging system. A synthetic 3D image generator based upon open source raytracing software is also presented. From a visual inspection, the point clouds appear representative of the types of data sets one would achieve from the actual instrument, though thorough validation testing is required. The validation testing will be conducted as part of future work.

6. FUTURE WORK

Recommendations for future efforts include the following:

- Validate synthetic imagery against real data.
- Investigate different kernels and raycast oversample sizes for generating mixed pixel effects.
- Repeat range uncertainty experiment and model development for third-generation sensor (second-generation reported in this paper).
- Develop data sets of various steel beam configurations for object recognition research.

7. REFERENCES

[1]   A. M. Lytle and K. S. Saidi, "NIST research in autonomous construction," Autonomous Robots, vol. 22, pp. 211-221, 2007.

[2]   Q. Zheng, S. Z. Der, and H. I. Mahmoud, "Model-based target recognition in pulsed ladar imagery," Image Processing, IEEE Transactions on, vol. 10, pp. 565-572, 2001.

[3]   G. Powell, R. Martin, D. Marshall, and K. Markham, "Simulation of FLIR and LADAR data using graphics animation software," Computer Graphics and Applications, 2000, pp. 126-134.

[4]   "POV-Ray, http://www.povray.org/."

[5]   "MegaPOV, http://megapov.inetart.net/."

[6]   R.Olsson and Y. Xu, "An interactive ray-tracing based simulation environment for generating integral imaging video sequences," in Three-Dimensional TV, Video, and Display IV, Boston, MA, USA, 2005, pp. 60160F-8.

[7]   R. J. Grasso, G. F. Dippel, and L. E. Russo, "A model and simulation to predict 3D imaging LADAR sensor systems performance in real-world type environments," in Atmospheric Optical Modeling, Measurement, and Simulation II, San Diego, CA, USA, 2006, pp. 63030H-12.

[8]   A. M. Lytle, I. Katz, and K. S. Saidi, "Performance Evaluation of a High-Frame Rate 3D Range Sensor for Construction Applications," in ISARC Ferrara, Italy, 2005.

[9]   S. Thrun, W. Burgard, and D. Fox, Probabilistic Robotics MIT Press, 2006.

[10] VLPov, http://vision.ucla.edu/~vedaldi/code/ vlpov/ vlpov.html."