Copyright

by

Jochen Teizer

2006

المنارة للاستشارات

www.manaraa.com

The Dissertation Committee for Jochen Teizer certifies that this is the approved version of the following dissertation:

REAL-TIME SPATIAL MODELING TO DETECT AND TRACK RESOURCES ON CONSTRUCTION SITES

Committee:

Carl T. Haas, Co-Supervisor

Carlos H. Caldas, Co-Supervisor

James T. O'Connor

Katherine A. Liapi

J. K. Aggarwal

المنارة للاستشارات

www.manaraa.com

REAL-TIME SPATIAL MODELING TO DETECT AND TRACK RESOURCES ON CONSTRUCTION SITES

by

Jochen Teizer, Diplom-Ingenieur

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2006

المنارة للاستشارات

www.manaraa.com

UMI Number: 3267849

UMI®

UMI Microform 3267849

Copyright 2007 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

المنارة للاستشارات

www.manaraa.com

Dedication

An meine Mama und meinen Papa,

Marlitt und Franz Teizer

المنارة للاستشارات

www.manaraa.com

Acknowledgements

This dissertation completes another segment of my life that was truly very enjoyable and on the other hand, not very often to recognize, very hard to accomplish.

During my time as a Ph.D. student I have learned that excellence in research comes from intelligent leadership. Leadership means to pursue innovative goals, to teach, to support, to motivate, and to learn from others. Rapid progress is successful when work is reflected quickly, critically, and in a fair manner. People have spent time with me in recent years and they are true leaders. They are listed in a chronological way:

Critical with me from the beginning but always strongly supporting my decisions have been my mom and dad, Marlitt und Franz Teizer. Next significant impact is the unity among us brothers, Winfried, Klaus, Peter, and me as an unbeatable stronghold. In the meanwhile their wives and kids gave me a lot of confidence too.

Through the support of Dr. Gerhard W. Schmidt, Prof. Fritz Gehbauer, and Prof. Kunibert Lennerts, all at the University of Karlsruhe (TH) in Germany, I was awarded my German degree as a Diplom-Ingenieur. They also offered me the first insights into the academic life. As a visiting scholar with Prof. David N. Ford at Texas A&M University, I was allowed to make my first experiences with the educational system in the United States of America. This experience contributed to my decision to pursue my Ph.D. At the University of Texas at Austin my Ph.D. supervisors, Prof. Carl T. Haas and Prof. Carlos

المنارة للاستشارات
www.manaraa.com

H. Caldas, strongly supported me in decision making, education, research, and research travel. Their dedication to me has been greatly appreciated.

I have to be very thankful to a lot of other people, since they had significant contributions or inspirations in my life. They are: Prof. Katherine A. Liapi, Frédéric Bosché, Jongchul Song, Dr. Richard Gebken II, Giovanni Migliaccio, Alan Lytle, Dr. Kamel Saidi, Dr. Wai Kiong Oswald Chong, Dr. Rayson Chou, Dr. Changwan Kim, Dr. Issam Srour, David Grau Torrent, Dr. Lilin Liang, Dr. Soon-Wook Kwon, Dr. Haitham Logman, Prof. Richard Tucker, Ruwini Weerasooriya, Shan-Pin Yeh, Raquel Esquatel, Jakub Felkl, Cecile Rey, Leif and Ingrid Ristroph, Hilda Ramos, Esther Maier, Mikko Börkircher, Jochen Gierich, Jens Harder, Ruth Jirauscheck, Bernhard Wunsch, and Jochen Merkenschlager. To all of you above thank you very much. I also like to thank the National Science Foundation under grant CMS #0409326, the National Institute of Standards and Technology under solicitation number SB1341-04-Q-0898, the Construction Industry Institute, and the "Shirley and Richard Tucker Endowed Scholarship" for their indirect and direct financial support.

Since so many challenges and expectations remain in this world to be solved, each of us has the responsibility to work hard to solve some of these goals. A special thank you goes to my brothers Winfried, Klaus and Peter who were always available to reflect and support my career path. Thanks to y'all!

August 11, 2006

المنارة للاستشارات
www.manaraa.com

REAL-TIME SPATIAL MODELING TO DETECT AND TRACK RESOURCES ON CONSTRUCTION SITES

Publication No._____

Jochen Teizer, Ph.D.

The University of Texas at Austin, 2006

Supervisors: Carl T. Haas and Carlos H. Caldas

For more than 10 years the U.S. construction industry has experienced over 1,000 fatalities annually. Many fatalities may have been prevented had the individuals and equipment involved been more aware of and alert to the physical state of the environment around them. Awareness may be improved by automatic 3D (three-dimensional) sensing and modeling of the job site environment in real-time. Existing 3D modeling approaches based on range scanning techniques are capable of modeling static objects only, and thus cannot model in real-time dynamic objects in an environment comprised of moving humans, equipment, and materials. Emerging prototype 3D video range cameras offer another alternative by facilitating affordable, wide field of view, automated static and dynamic object detection and tracking at frame rates better than 1Hz (real-time).

This dissertation presents an imperical work and methodology to rapidly create a spatial model of construction sites and in particular to detect, model, and track the

المنارة للاستشارات

www.manaraa.com

position, dimension, direction, and velocity of static and moving project resources in real-time, based on range data obtained from a three-dimensional video range camera in a static or moving position. Existing construction site 3D modeling approaches based on optical range sensing technologies (laser scanners, rangefinders, etc.) and 3D modeling approaches (dense, sparse, etc.) that offered potential solutions for this research are reviewed. The choice of an emerging sensing tool and preliminary experiments with this prototype sensing technology are discussed. These findings led to the development of a range data processing algorithm based on three-dimensional occupancy grids which is demonstrated in detail. Testing and validation of the proposed algorithms have been conducted to quantify the performance of sensor and algorithm through extensive experimentation involving static and moving objects. Experiments in indoor laboratory and outdoor construction environments have been conducted with construction resources such as humans, equipment, materials, or structures to verify the accuracy of the occupancy grid modeling approach. Results show that modeling objects and measuring their position, dimension, direction, and speed had an accuracy level compatible to the requirements of active safety features for construction. Results demonstrate that video rate 3D data acquisition and analysis of construction environments can support effective detection, tracking, and convex hull modeling of objects. Exploiting rapidly generated three-dimensional models for improved visualization, communications, and process control has inherent value, broad application, and potential impact, e.g. as-built vs. as-planned comparison, condition assessment, maintenance, operations, and construction activities control. In combination with effective management practices, this sensing approach has the potential to assist equipment operators to avoid incidents that result in reduce human injury, death, or collateral damage on construction sites.

المنارة للاستشارات
www.manaraa.com

Table of Contents

Chapter 1: Introduction 1
1.1 Background 1
1.2 Motivation and Challenges 3
1.2.1 Need for Safety in Construction 4
1.2.2 Need for Automated Assistance 5
1.2.3 Need for Advanced Data Processing Methods 6
1.3 Research Hypothesis 8
1.4 Research Objectives and Scope 8
1.5 Research Methodology 10
1.6 Structure of Dissertation 11

Chapter 2: Literature Review 12
2.1 Potential Range Data Acquisition Technologies and Methods 12
2.2 Optical and Infrared Range Sensing 14
2.2.1 The Physics behind Light Waves and Photons 15
2.3 Vision Based Ranging Technologies and Methods 21
2.3.1 Triangulation, Photogrammetry, and Stereo Vision 23
2.3.2 Interferometry 24
2.3.3 Time-of-Flight (TOF) 24
2.4 State-of-the-Art of Optical Range Sensing Using TOF 29
2.4.1 Contactless Distance Measurement Approaches 30
2.4.2 LADAR, LIDAR, and 3D Laser Scanner Approach 31
2.4.3 Time Constraint and Clean Data 35
2.4.4 Sparse Point Cloud Approach 37
2.5 Review of Range Data Processing Methods 40
2.5.1 Model-Based and Data-Driven Approaches 41
2.5.2 Occupancy Grids 41
2.5.3 Data Clustering 42
2.5.4 Image and Model Registration 45

المنارة للاستشارات
www.manaraa.com

2.5.5 Tracking Objects 46
2.5.6 Classification of Objects 47
2.5.6 Modeling Approaches of Large Data for Large 3D Terrains 48
2.5.7 Localization for Three-Dimensional Maps 49
2.5.8 Obstacle Detection and Motion Detection Methods 50
2.5.9 Using Intensity and Range Information to Create 3D models 51
2.5.10 Image Processing Techniques Based on Intensity Data 52
2.6 Selection of Range Sensing Technology & Data Processing Method 55
2.6.1 3D Range Camera for Range Sensing 56
2.6.2 Real-time 3D Occupancy Grid for Range Data Processing 59

Chapter 3: Preliminary Study of 3D Range Sensing Technology 60
3.1 Three-Dimensional Range Cameras 60
3.2 Working Principle of 3D Range Cameras 61
3.3 Preliminary Experiments and Findings 66
3.4 Advantages and Limitations of Optical 3D Range Sensing 77
3.5 Summary 79

Chapter 4: Three-Dimensional Occupancy Grid Algorithm 80
4.1 Introduction 80
4.2 Real-time 3D Modeling: Overview of Approach 81
4.3 Concept of 3D Video Range Modeling Algorithm 83
4.4 Requirements for Three-Dimensional Occupancy Grid Modeling 85
4.5 Algorithm Details 87
4.6 Step-By-Step Explanation of Algorithm Using an Example 91
4.7 Findings And Considerations During Algorithm Development 103
4.8 Summary 106

Chapter 5: Experiments, Results, and Analyses 107
5.1 Preliminary Work to define Algorithm Parameters and Experiments 107
5.1.1 Calibration of Sensor's Field-Of-View 108
5.1.2 Non-Ambiguous Distance and Voxel Size 111
5.1.3 Characteristics of Target Objects in Experiments 113

المنارة للاستشارات
www.manaraa.com

5.1.4 Characteristics of Experimental Field-Of-View 114
5.1.5 Summary of Indoor Experimental Setup 116
5.1.6 Summary of Outdoor Experimental Setup.....117
5.2 Experimental Research Methodology.....119
5.3 Experiments to Determine Position and Dimension Accuracy.....122
5.4 Experiments to Determine Direction and Velocity.....132
5.5 Outdoor Experiments with Static 3D Video Range Camera 137
5.6 Experiments with Moving 3D Video Range Camera 141
5.6.1 Indoor Experiment 141
5.6.2 Outdoor Experiment.....143
5.7 Results and Analyses of Experiments.....144
5.7.1 Time to process data 145
5.7.2 Discussion of Errors.....145

Chapter 6: Potential Application Areas 149
6.1 Target Identification and Tracking 149
6.2 Assisted Site Construction 149
6.3 Work Zone Safety 150
6.4 Asset Management.....151
6.5 Intelligent Transportation Systems (ITS).....152
6.6 Summary of Application Areas 154

Chapter 7: Conclusions 155
7.1 Review of Objectives.....155
7.2 Contributions and Conclusions 156
7.3 Recommendations.....157

Appendix A – Technical Details to 3D Range Camera and Interface 160

References.....165

Vita 173

المنارة للاستشارات
www.manaraa.com

List of Tables

Table 1: International visibility table (DeWeer and Gilbert, 2006) 18

Table 2: Ideal and existing attributes of range sensing solutions.....58

Table 3: Commercial range cameras manufacturers and products 60

Table 4: Limitations affecting the potential of 3D range cameras.....76

Table 5: User-defined parameters and thresholds.....92

Table 6: Tracking decision criteria 98

Table 7: Angular calibration of field-of-view.....110

Table 8: Analysis of field-of-view of SwissRanger 2b.....112

Table 9: Position measurement error of experiment #28.....125

Table 10: Position measurement error of experiment #28.....126

Table 11: Summary of position accuracy.....128

Table 12: Dimensional error from experiment #22 with a box.....129

Table 13: Dimensional error from experiment #22 with pipe.....129

Table 14: Dimensional error from experiment #28 with 5 boxes 130

Table 15: Summary of dimensional accuracy.....131

Table 16: Differences in directional measurements.....134

Table 17: Differences in velocity measurements.....136

Table 18: List of potential application areas for 3D range sensing.....154

المنارة للاستشارات

www.manaraa.com

List of Figures

Figure 1: Framework illustration for real-time 3D modeling of infrastructure..1

Figure 2: Dual fatal accident caused by missing safety features.....4

Figure 3: Research methodology flowchart 10

Figure 4: Wavelength images and Planck Curves.....15

Figure 5: Object scattered light signal throughput on sensor detector 19

Figure 6: Taxonomy for optical range imaging.....23

Figure 7: Illustration of three fundamental optical range imaging methods 25

Figure 8: Illustration of the modulation parameters. 27

Figure 9: Laser scanner range image and LIDAR image from WTC site.....33

Figure 10: Sparse Point Cloud Approach: (a1) Fitted and matched cuboid, (a2) Actual object, (a3) Fitted and matched Pipe, and (a4) Actual object. Bounding box generation: (b1) Actual objects, (b2) Bounded objects. Merging primitives can improve model display: (c1) Step 1: No object recognition process, (c2) Step 2: Object distinction process, (c3) Step 3: Object distinction and object reconstitution process and (c4) Picture of the scanned scene (Teizer, 2005a). 37

Figure 11: Sparse point cloud modeling approach using a Total Station.....39

Figure 12: Preferences of optical range data acquisition technologies 57

Figure 13: Working principle of video 3D range camera technology.....61

Figure 14: Video range sensing working principle 62

Figure 15: Simulated cross section and potential distribution of a 2-tap sensor (Büttgen et al., 2005). 63

المنارة للاستشارات

www.manaraa.com

Figure 16: Object, shadow area, and projection of range points on active sensor pixels 67
Figure 17: Standard deviation of 100 measurements on a range image shows diffractions of up to 3 cm (Gut, 2004) 70
Figure 18: "Ghost image" and noisy measurement in outdoor environments....72
Figure 19: Original, intensity, and range image information, with extracted convex hull of a parked car (Teizer et al., 2005) 72
Figure 20: Outdoor tracking of car and bus in bright sunlight 73
Figure 21: Spectral Filtering with burst of LEDs – Influence of the natural background light on sensor emitted light source (PMD, 2005) 75
Figure 22: Modeling capabilities of a laser scanner vs. 3D range camera 77
Figure 23: Three-dimensional modeling to detect static and moving object boundaries 82
Figure 24: Three-dimensional occupancy grid model 84
Figure 25: Flowchart of 3D Occupancy Grid Based Model 85
Figure 26: Data processing flowchart 88
Figure 27: Original color and range image (fascia board, box, wall) 92
Figure 28: 3D occupancy grid of one frame (sensor's FOV is Local Model, World Model encompasses entire scene) 93
Figure 29: Single-linked agglomerative hierarchical clustering dendrogram 95
Figure 30: Elevation and front view of clustered 3D occupancy grid model 97
Figure 31: Visualization the track of an object by plotting 52 consecutive Local Model frames in plane view 98
Figure 32: Typical view on summary results of one experiment 100
Figure 33: Analysis of Cluster #7 102

المنارة للاستشارات
www.manaraa.com

Figure 34: Light beams in Cartesian and spherical perspective 108
Figure 35: Sensor's angular field-of-view.....109
Figure 36: Elevation and plane view of camera's FOV 109
Figure 37: Sensor's maximum range limit to non-ambiguous range 111
Figure 38: Vertical and horizontal position of SwissRanger 2 and its output matrix 112
Figure 39: Components of indoor experiments 114
Figure 40: Overview of indoor experimental setup.....115
Figure 41: Geodesy to indoor experimental setup (sensor's FOV, boundaries, tracks) for different resources and angles.....117
Figure 42: 3D video range camera setup in outdoor environment 118
Figure 43: Overview of outdoor experimental setup.....118
Figure 44: Flowchart of experimental plan in an obstacle-avoidance process for heavy machine operation (extended, after Gonzales and Woods, 1992, and Kim et al., 2004) 119
Figure 45: Phases of range data collection, processing, and analyses.....121
Figure 46: Position and dimension measurement – Experiment #20 124
Figure 47: Position and dimension measurement – Experiment #28 128
Figure 48: Indoor experiment – Moving pipe 133
Figure 49: Indoor experiment - Moving box 133
Figure 50: Indoor experiment – Human climbing a ladder 133
Figure 51: Outdoor experiment #24-04-03: Human walking around barrel 139
Figure 52: Outdoor experiment #24-03-02: Two humans and object clutter...140
Figure 53: Outdoor experiment #18 with a moving skid steer loader.....141

المنارة للاستشارات
www.manaraa.com

Figure 54: Start and end position of moving 3D video range camera capturing static objects.....142
Figure 55: Sensor mounting on bobcat and outdoor experimental setup.....143
Figure 56: Processed range frame from moving sensor at a velocity of 2m/s .144
Figure 57: Sources of errors defining the accuracy of detection and tracking.146
Figure 58: Components of 3D camera.....160
Figure 59: 4-tap structure on sensor (after Gut, 2004).....161
Figure 60: (a) Screenshot of camera software interface with main tab options, (b) Start tab, (c) Movie tab (d) Range grabbing tab, (e) 3D display options tab, (f) Frame rate, sensor model, update display, and (g) Depth, intensity/amplitude, and histogram window.....164

المنارة للاستشارات
www.manaraa.com

Chapter 1: Introduction

This research seeks to demonstrate that emerging range sensing technology and range data processing methods are able to detect and track static or moving objects in real-time in the field of view of a static or moving three-dimensional (3D) video range camera. The following sections in this chapter explain the research background, motivation, challenges and hypothesis, objectives and scope, methodology, and the outline of this dissertation.

1.1 BACKGROUND

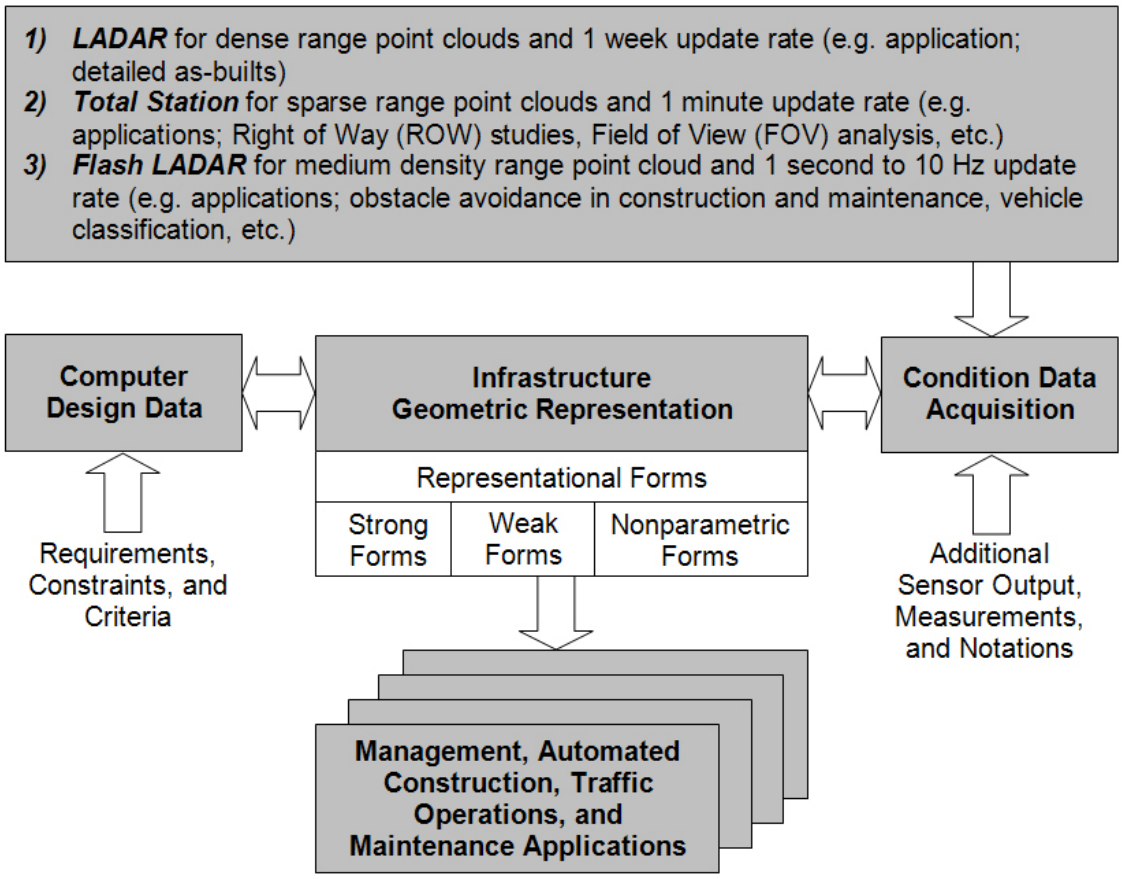


Figure 1: Framework illustration for real-time 3D modeling of infrastructure

www.manaraa.com

Having the ability to locate, describe in 3D, control, and track objects within a field of view has become an important factor in intelligent infrastructure systems, construction, maintenance, and asset management. Depending on the application, this usually requires that assets somehow be scanned and then modeled in 3D at varying frequencies. Asset management may require no more than bi-annual updates, whereas construction activities may require real-time or greater than 1 Hz updates (Kim and Haas, 2002). For these cases and the continuum in between, it is useful to have a framework for the process of data acquisition and 3D model building. An overall framework for doing this 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. In addition, 3D models can be represented in three basic forms.

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. Defined by Hirschberg (1996), "weak forms" often are related to strong forms, but previous design information was improperly documented or is missing. Examples are a rectangle which, over time, 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.

المنارة للاستشارات
www.manaraa.com

Thus, the need for fast and accurate geometric modeling requires using existing and emerging laser ranging technologies like laser scanning, LADAR (Laser Detection and Ranging), Total Stations, and 3D video range cameras (aka. Flash LADAR or Flash Laser Detection and Ranging). These new sensor technology innovations now allow addressing problems of the highest priorities in the transportation and construction area, such as real-time wide field-of-view 3D modeling to support an active obstacle detection and avoidance system for heavy machine operation.

1.2 MOTIVATION AND CHALLENGES

The commonly known "Three E's" in safety engineering refer to "Education" to train people as a preventive tool, "Enforcement" as a regulative step if education was not successful (e.g. through the Occupational Safety and Health Administration – OSHA), and finally, "Engineering", as last instance to develop new tools that can automatically assist, e.g. in safety matters. As a result, technology and methods are such tools that engineers can develop to prevent such intentional or unintentional misbehavior and thus avoid accidents, injuries, or fatalities. However, technology can only be a "third" eye.

This research tries to demonstrate that detecting an object in real-time could prevent accidents by properly installing existing or emerging safety devices. In this research the focus, however, will be on the detection of objects in the field-of-view of an emerging range imaging device, but not on proving a concept or study of the successful implementation of such a device to prevent accidents. In addition, safety, as a matter of fact, is only one application where this research and the used emerging technology may become a leading stimulus for research. Many other application areas exist and are not limited to construction, e.g. obstacles on railroad crossings. Some of these areas are currently under research and promise benefits even larger than the developed rapid three-dimensional (3D) object detection and tracking method might generate.

المنارة للاستشارات
www.manaraa.com

1.2.1 Need for Safety in Construction

A typical construction site environment is characterized by static and slowly moving objects (velocity of machine operation and humans of up to 15 km/h). Various modeling techniques and commercial range scanning products are able to create computer models of static construction scenes. However, these technologies are not able to detect and track moving objects or work in real-time and thus do not fit many applications.

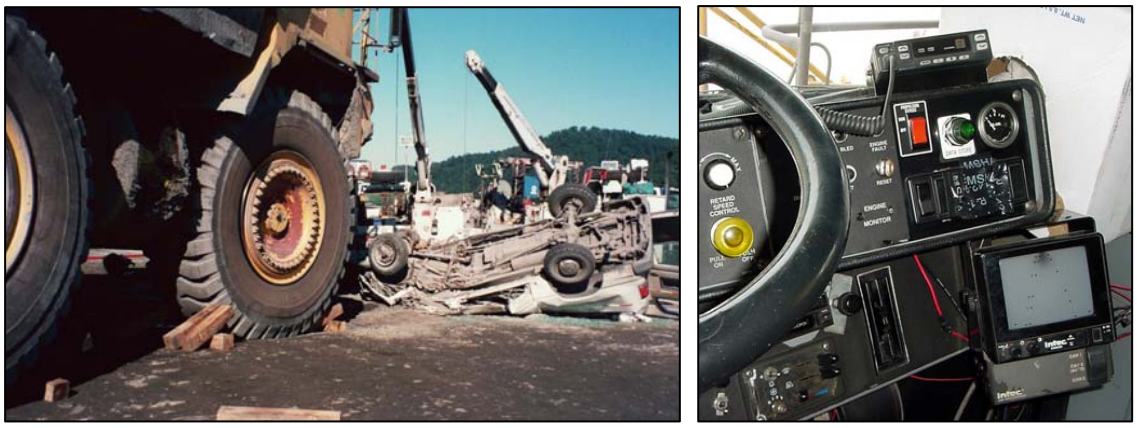


Figure 2: Dual fatal accident caused by missing safety features

Tremendous efforts have been placed by construction equipment manufacturers to improve the operator's path oversight by using video systems for bulky machines (Haas and Kim, 2002). This can make skilled operators become unnecessary or less attentive during the operation phase of their equipment as they heavily rely on current technology systems and its information value provided. Such a system could be a video camera and screen for the reverse gear of a dump truck. Figure 2 (MSHA, 2003) shows a dual fatal accident. The original figure caption states that a video system could have avoided the accident and the loss of two lives. Ultimately, awareness of the heavy equipment operator or blind spot control using a long range sensing device with obstacle avoidance system

المنارة للاستشارات
www.manaraa.com

may have avoided this accident. A signal as an early warning in visual and acoustical form of an object appearing or being in the project path could have been set off. In this research and in favor to assist operators in making fast decisions, an automated range data acquisition and processing method is proposed which allows detecting static and moving objects by using a real-time laser ranging device, called 3D video range camera or Flash LADAR.

The impact of technology in construction environments of workforce and machines is important for this research in two ways: First, technology can assist the workforce in situations where the human perception is unable to control the environmental influences (e.g. complex cluttered scenes, night work, etc.). Secondly, properly developed technology can assist the workforce and engineers in making better and faster decisions (e.g. real-time generated 3D models). Applying fundamentally innovative technologies, such as laser range cameras, and new methods, like algorithms previously defined in different engineering fields, can make construction processes faster, cheaper, and safer. Such technologies can have significant impact in today's construction world as the workforce is less skilled due to several reasons. This problem can be solved with a strategy focusing on maximizing the construction productivity by allowing the workforce using actively or passively new technologies. This may lead to leverage, de-skill or re-skill, or even replacement of workforce (Haas, 2004).

1.2.2 Need for Automated Assistance

Targets that in general are easy to detect and track by humans are often difficult to select for a computer (and vice versa). The evolution of the human kind and its body functions has greatly enabled it to rapidly analyze environments with different natural senses, e.g. the eyes. The combination of both eyes, as a range data acquisition tool, in conjunction with the human's brain allows making decisions in fractions of a second.

المنارة للاستشارات
www.manaraa.com

Accidents like in Figure 2, however, have not been avoided due to lack of attention from the operator or missing safety features like video cameras installed on the heavy mining equipment.

Naturally, humans tend to use global or learned pattern strategies (i.e. object orientation) to recognize objects and to track them. The human brain remembers how many targets were in the scene a moment ago and what their spatial relationship was. If a target is difficult to find, that information is used to cue a more sophisticated search (Sanders-Reed, 2006). Computers vision and their integrated search algorithms, on the other side, look for patterns by simple threshold detection (i.e. change in grey value) and may not use global recognition of object-to-object relationships. This enables computers to handle much more objects at the same time, but while humans are easily overwhelmed and confused in this case, they can usually read and interpret alphanumeric information with their trained eye more easily.

For this reason, there is a need to develop a real-time 3D workspace modeling method which intends to work independently from human perceptions, but allows assisting and influencing the entire process at any time. This research will be a preliminary step to demonstrate that emerging range imaging technology can detect and track objects automatically. Future research still needs to address any implementation system, e.g. commercialized range imaging for an automated obstacle avoidance system.

1.2.3 Need for Advanced Data Processing Methods

Once the sensor has acquired range data, in the ideal situation a computer (and its data processing programs and algorithms) is able to process this data to information to find all and only the relevant targets. Computer software may accurately extract features and calibrate the data appropriately. A computer program may build time stamps and

المنارة للاستشارات
www.manaraa.com

histories for all detected targets in range frames. In an optimal case scenario, the data is generated and processed automatically or is available on demand.

The reality of the computer and data processing world tells us, however, that objects or targets often are hard to find and separate from the background. In particular the construction environment is characterized as very object cluttered, hazardous, and hard to overview. This may require special data handling. In addition, proper calibration data is often not available or is ambiguous since scenes and its content may change rapidly from one moment to the next. To detect potential object motion in a scene results then in broken tracks or objects are only partially visible since they were miss-segmented and miss-assigned to other objects. As a result, operator interaction is often required to select the best set of handling the situation (e.g. avoiding objects by choosing a different travel path) or helping to select the best computer processing options for a specific task.

In conclusion, achieving the ideal result, requires good observations, good data collection and planning, good software that enables proper processing, and a smart analyst, whether human or computer.

The detection and tracking of objects can be performed manually or automatically. Manual tracking offers great flexibility with a few, identified, large or partially covered object movements, but is tedious, since operators easily get overwhelmed with the large number of objects. Manual object detection and tracking is hard especially in low contrast, low resolution areas, where objects are closely spaced or move fast from one frame to the next (Sanders-Reed, 2006). In a more ideal heavy equipment operation like in construction, human operators should fully be able to concentrate on their work task to achieve highest productivity and to ensure overall best safety of the operation.

المنارة للاستشارات
www.manaraa.com

Automatic tracking can easily handle a large number of objects and has the potential to detect them all and to assign them individual tracks. Computer-assisted detection and tracking can take advantage of the developed data processing methods in the recent years. however, it also can lose the object, select and extract background objects if target motion is two large between frames, use calibrated intensity and range measurement values on a quantitative basis, or treat each object individually and independently, thus making automatic detection and tracking of objects to a powerful asset in navigation assistance.

1.3 RESEARCH HYPOTHESIS

The research hypothesis is: Obstacles in the work space surrounding heavy construction equipment can be detected and tracked in real-time using a three-dimensional range acquisition camera and appropriate range data processing algorithms.

The emphasis in this research is on experimenting with emerging prototype range sensing technology, developing capable real-time range data processing algorithms, and evaluating the accuracy of the technology-algorithm system in laboratory and construction like environments.

1.4 RESEARCH OBJECTIVES AND SCOPE

The objective of the research proposed here is to perform technically feasible three-dimensional modeling of static and moving objects in unstructured or structured construction scenes in real-time. The goal of this rapid workspace modeling approach is to support automated construction performance monitoring and control for obstacle avoidance by continuously providing timely, accurate, and precise detection and location information of objects within a workspace.

المنارة للاستشارات
www.manaraa.com

Applications to be identified will be based on the provided 3D range camera technology without assuming any hardware modification. Objects of particular interest in these applications are materials, machine equipment, humans, and other uniquely identifiable objects. To support the main objective, the research effort was divided into four sub-objectives:

1. Determine the technical feasibility of real-time 3D modeling using a 3D range sensor. Find the advantages and limitations of the sensor, i.e. its accuracy level, and find where and when it can become useful to be applied in natural environments like on construction sites. Identify the kinds of construction objects the 3D range sensor is able to measure ranges from.
2. Develop the basic framework algorithms for real-time 3D modeling. Describe and implement a rough and primitive model of a construction like scene or workspace including several objects of different kinds of art, e.g. static and moving targets.
3. Verify the applicability of the detection and tracking systems through implementation in indoor experiments and if possible, in real-life environments on construction sites. Demonstrate the modeling approach can satisfy under certain scenarios and determine the typical errors of the system parameters.
4. Describe future improvements for sensing using 3D range camera devices.

The proposed framework was not based on object recognition or object identification as this may follow in a next research phase. This research was limited to the development of algorithms for an obstacle-avoidance system to enhance safety for heavy equipment operations on construction sites. Implementation and testing of the obstacle-avoidance system is expected to proceed beyond the computer-simulation stage. Thus,

المنارة للاستشارات
www.manaraa.com

the proposed research scope is defined by: Detect and track objects in static position or with moving velocity relative to the 3D range camera's field of view.

The research involves detecting and tracking objects of simple geometry, various shapes and sizes, static or moving, at different speeds (slow and fast), in different relative velocities and directions, and in a work space environment clean of other objects. Detection and tracking was based on useful computational range data processing methods. Parameters of the objects detection and tracking were object position, object dimension, object speed, object direction, and number of objects in a scene. Due to current sensor characteristics the validation part of the research focused mostly on indoor experiments. Outdoor experiments, however, were performed as well.

1.5 RESEARCH METHODOLOGY

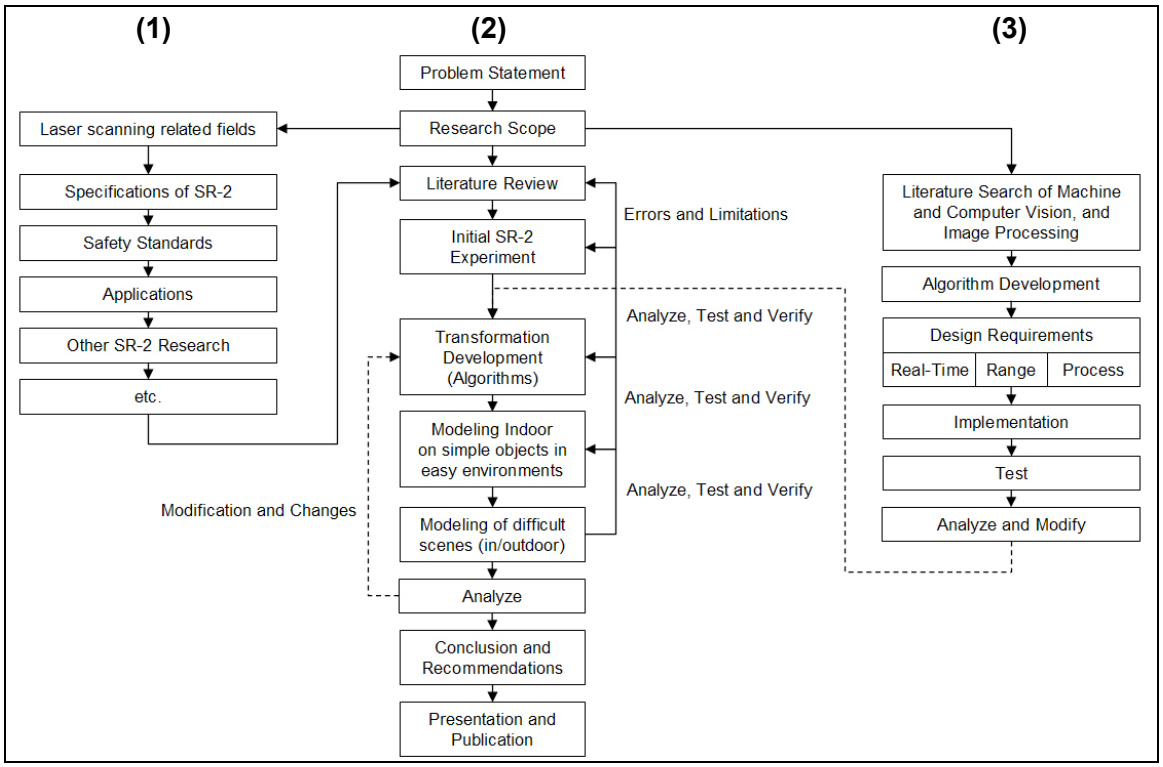


Figure 3: Research methodology flowchart

المنارة للاستشارات
www.manaraa.com

The research methodology focuses on three main parts. Figure 3 illustrates the methodology as the methods being used to comply with the research goals. (1) Review the existing knowledge in the literature to determine the background of range acquisition technologies, and needs for the target application in safety, (2) a literature review for range data processing methods and the development of suitable algorithms to meet the research objectives, and (3) Research to verify the developed algorithms in experiments.

1.6 STRUCTURE OF DISSERTATION

Research objectives, scope, and methodology have been discussed. The next chapters of the dissertation are structured in five main chapters. In the following chapter 2, a background literature review gives an overview and an evaluation of existing optical sensing methods and imaging technologies that can acquire range data reflectorless. In the same chapter, data processing algorithms and methods are reviewed. This chapter concludes in choosing the best fitting technology and processing method that have the potential to meet the research objectives. A significant part of this research has focused on learning which technology best fits the research objectives, thus, chapter 3 concentrates on evaluating the preliminary findings of the chosen technology and lists its advantages and limitations. Chapter 4 explains step by step and in detail the developed processing algorithm to convert the range data into valuable 3D model information. In chapter 5 the experimental plan is introduced. Experiments demonstrate the validation for selecting a technology and the developed processing algorithm. Analyses and results are presented. Potential errors sources are discussed. In Chapter 6 potential application fields in real-time 3D modeling in construction and transportation are introduced. In the concluding chapter, chapter 7, findings and contributions as well as future research is described.

المنارة للاستشارات
www.manaraa.com

Chapter 2: Literature Review

This chapter gives an overview of range data acquisition tools and range data processing methods. Based on this literature review, the selection of a range sensing technology and range data processing method conclude this chapter.

2.1 POTENTIAL RANGE DATA ACQUISITION TECHNOLOGIES AND METHODS

In the particular application of construction and elsewhere sensing technologies can become useful tools to locate, track, model, and visualize or simulate project or work task relevant processes. Such smart sensors can be categorized in the following sub-areas (Robotics Trends, 2004):

a) Global Positioning System (GPS): Includes conventional and differential GPS. Differential GPS where satellite information is compared to terrestrial fixed stations offers more accurate positioning in the lower centimeter range. GPS can also be closely linked with a Graphical Information Systems (GIS) to provide detailed terrain mapping. GPS manufacturers include: Applanix (Trimble), Garmin, Navcom Starfire, and OmniSTAR.

b) Inertial Navigation: Includes measuring translational and rotational distance, velocity, and acceleration through conventional macroscopic systems, MEMs or laser gyroscopes. Inertial navigation provides immediate short range feedback on the movement of a vehicle and can fill in the gaps where GPS information is not available or accurate enough. Often inertial navigation is integrated with GPS to offer a total solution. Manufacturers include: Applanix (Trimble), XSens, Crossbow, and Analog Devices.

المنارة للاستشارات
www.manaraa.com

c) Optical light and lasers: Lower capital intensive optical sensing systems include color cameras, stereo vision systems, 2D image analysis, 3D reconstruction, infrared, and low-light level sensing. Manufacturers include: Point Grey (Bumblebee stereo vision), SAIC (stereo vision), Kenyon Labs (gyro stabilizers for cameras), Indigo Systems (FLIR). Higher end light or laser based sensors are Light Detecting and Ranging (LIDAR) or Laser Radar Detection and Ranging (LADAR): Lasers are used to scan surfaces, provide range information, and perform 3D reconstruction. Manufacturers include: SICK, Riegl, Leica (formerly Cyra Technologies Inc.), Canesta, CSEM and PMDtec (for 3D range cameras, Flash LADAR), etc.

d) Radio Detection and Ranging (RADAR): Can be used for long and short range obstacle detection or as a radar odometer. Phased array technology warns of potential hazards, such as stopped or slow-moving vehicles. Doppler radar can be used to measure the velocity of moving objects. Manufacturers include: Dickey-John, Eaton (VORAD).

e) Radio Frequency and Identification (RFID): Can be used to identify and track materials and objects. Passive or active devices emit radio waves transmitted to readers in small or medium ranges of up to a couple dozen meters. Manufacturer includes: Identec, Intel, and Texas Instruments.

f) Ultrasonic: Can use single or multiple elements to gauge distance or as a basic point proximity sensor. Manufacturers include: Honeywell, EchoMaster, Polaroid, and Massa.

The sensor for this research is based on optical range sensing, since it delivers characteristics that offer a wide field-of-view, a reflectorless range dense point cloud at

المنارة للاستشارات
www.manaraa.com

high-frame update rates that are less error prone, eye-safe and in general safe to install and easy to handle. The following subchapters will explain the differences, advantages and limitations in optical range sensing.

2.2 Optical and Infrared Range Sensing

In the evolutionary process humans and animals have developed highly accurate sensing systems, e.g. eyes, ears, and noise. Equipped with these tools individuals have been able to survive a rugged nature, to find food, or a mate. Some animals, such as bats, aquatic mammals, and some electric fish, have made use of acoustic time-of-flight signals and electric fields to determine their immediate environment. Humans and animals, over time, predominantly improved the perceptual sense of the third dimension that allowed visualizing and navigating more effectively and easily. Today's living environment is comprised of three-dimensional objects in controlled or uncontrolled manner, which makes navigation more difficult and dangerous. Natural vision based sensing systems are often not able to compete with difficult tasks like to detect and track objects, to decide whether objects are safe to contact or a travel path that is the correct one to avoid hitting objects. These complex situations, like they are common for example in vehicular traffic, often overburden natural vision based sensing systems, causing accidents with collateral damage or even fatalities. The particular focus of this research is limited to construction typical environments. Construction environments can be characterized as one of the hazardous environments, since structures are unfinished, unstable, cluttered and often uncontrolled, and quickly changing.

The sensor system that are described next include passive infrared imagers, polarization imagers, and hyperspectral imaging spectrometers, and active laser radars (LIDARs, LADARs, laser range finders, laser scanners, Flash LADARs or 3D range

المنارة للاستشارات
www.manaraa.com

cameras) for sensing distributed or hard targets. These paragraphs introduce the basic concepts required for system-level understanding of optical and infrared sensor systems.

2.2.1 The Physics behind Light Waves and Photons

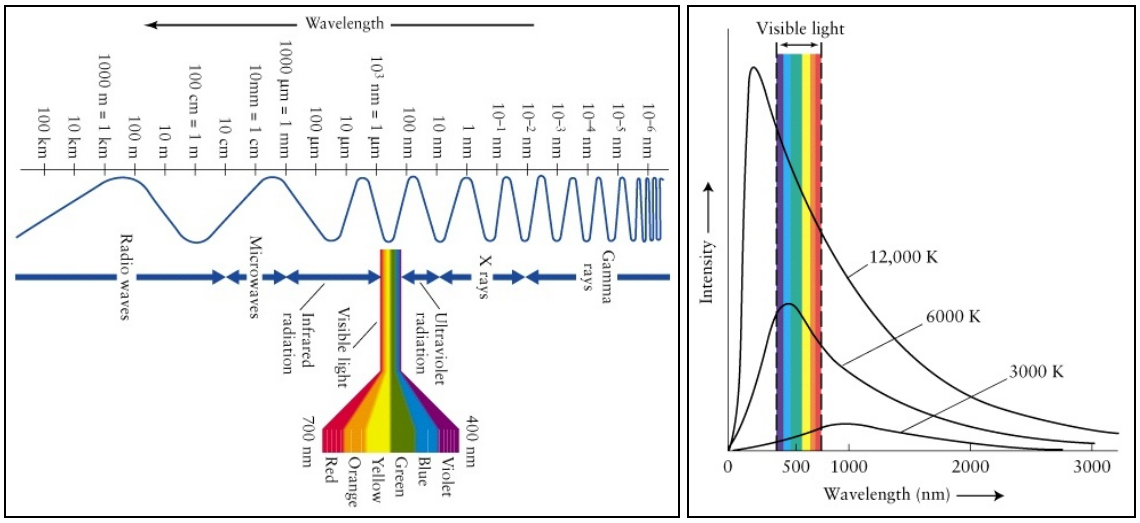


Figure 4: Wavelength images and Planck Curves

In general optical and infrared sensor use photon energy to measure distance information. In Figure 4 (Image courtesy: Freedman and Kaufmann, 2006, and Harvard, 2006), the wavelength of the optical and infrared spectrum is illustrated. Photon energy is proportional to frequency and inversely proportional to wavelength. The wavelength λ is defined by the formula

$$\lambda = \frac{c}{\upsilon}$$

or speed of light c divided by the frequency ν. The photon energy E is defined by

$$\mathrm{E} = h \cdot \upsilon = h \cdot \frac{c}{\lambda}$$

المنارة للاستشارات
www.manaraa.com

where h is the Planck's constant = 6.626×10^{-34} Js. Radiometric science allows to measure how much light is around. Quantity symbols and unit are energy Q [J], Power (flux) P or Φ [W], Intensity I [W/sr], Irradiance E [W/m^2], Exitance M [W/m^2], and Radiance L [W/(m^2*sr)]. Since photons are used quasi as a vehicle that carries timestamp data to measure the range information, it is important to measure the optical power collected by a receiver (sensor). To know the throughput G of photons on the pupil of the sensor system, it is necessary to define the projected solid angle. The projected solid angle defines at what angle photons can hit the pupil and thus, defines the space (a narrow cone) the pupil receives photons and then projects on the sensor. For small and narrow circular cones the projected solid angle Ω and the throughput G are defined by

$$\Omega = \frac{\pi \cdot r^2}{R^2} = \frac{A}{R^2}$$

$$G = A \cdot \Omega$$

(A is the flat projection of the spherical surface area). Mathematical details to derive the projected solid angle and throughput are presented in Shaw (2006). Finally, the optical power P, the amount of light gathered by an optical system is given by the product of source radiance and optical throughout,

$$P = L \cdot (A \cdot \Omega)$$

المنارة للاستشارات
www.manaraa.com

The received power can be increased with a brighter illumination source (L), a larger pupil area (A), or a larger Field of View (FOV, Ω), or all of the above (Shaw, 2006).

In thermal imaging (midwave and longwave sensors), for example, a sensor images the thermal emission (microwave, infrared, or whatever) to create an image where "bright regions" reflect warm parts in the scene or have higher emissivity. This approach is fundamentally different than intensity based images (2D) and heard to be implemented at Caterpillar Inc. for large dump trucks to help navigate them safely. In particular to avoid running over humans stepping too close to machinery it is necessary to know the Planck curves of blackbodies (see Figure 4). A "blackbody" is an object that emits a well defined spectrum of radiation solely based on its temperature. The hotter the blackbody, the more intense it is, and the shorter the peak wavelength (Harvard, 2006).

Several other physical parameters have effect on the throughput of optical power, e.g. Kirchhoff's Law (that all energy must be absorbed, reflected or transmitted), Lambertian radiation (that all radiation has equal radiance (intensity) in all directions), scattering in solar and terrestrial radiation, absorption of light by gas molecules (conversion of photons in thermal energy as they induce molecular rotations, vibrations, or electron orbital transitions alternate energy status), and atmospheric transmittance (conditions of atmospheres and molecules). The lower atmosphere, also called the troposphere, is at 0-11km height and contains N^2 (nitrogen) 78%, O^2 (oxygen) 21%, and Ar (argon) or other trace species of about 1%. Other influential components that can reduce the optical throughput include H^2O (water), O^3 (ozone), and CO^2 (carbon dioxide) (all up to 0.1%).

As a result, photons play a critical element in the measurement of range information using optical and infrared sensing methods. Photons are the carrier of the

المنارة للاستشارات
www.manaraa.com

range information and can highly influence range measurements, its accuracy level, if they at all can be registered at the sensor. This research focuses on the construction environment and machine, human, or other resource movement all generally takes part on the earth's surface ground (part of the lower atmosphere). Thus, atmospheric turbulences as random fluctuations of air temperature can cause the following phenomena: (Takarski, 1961):

- Scintillation (spatial and temporal variations in wavefront amplitude)
- Beam wander (random beam pointing caused by wavefront phase perturbations)
- Beam spreading (increase of beam divergence from phase variations)
- Image blurring (loss of high-spatial-frequency information from eddies
- Optical scattering (Rayleigh (e.g. atmospheric gas molecules), Mie (e.g. cloud droplets), Nonselective (e.g. rain drops, ice crystals), and Backscatter (e.g. aerosols, clouds))

Table 1: International visibility table (DeWeer and Gilbert, 2006)

Designation	Visibility [km]	Scattering coefficient σ_{ATM} [km^{-1}]
Dense fog	0 – 0.005	> 78.2
Thick fog	0.05 – 0.2	19.6 – 78.2
Moderate fog	0.2 – 0.5	7.82 – 19.6
Light fog	0.5 – 1	3.92 – 7.82
Thin fog	1 – 2	1.96 – 3.92
Haze	2 – 4	0.978 – 1.96
Light haze	4 – 10	0.391 – 0.978
Clear	10 – 20	0.196 – 0.391
Very clear	20 – 50	0.0782 – 0.196
Exceptionally	> 50	< 0.0782

The light propagation over a range R follows after Beer's exponential law with σ_{ATM} the scattering coefficient (see Table 1) (Jong, 2006):

المنارة للاستشارات
www.manaraa.com

$$L(R) = L(0)^{-\sigma_{ATM} \cdot R}$$ (Beer's Exponential Law)

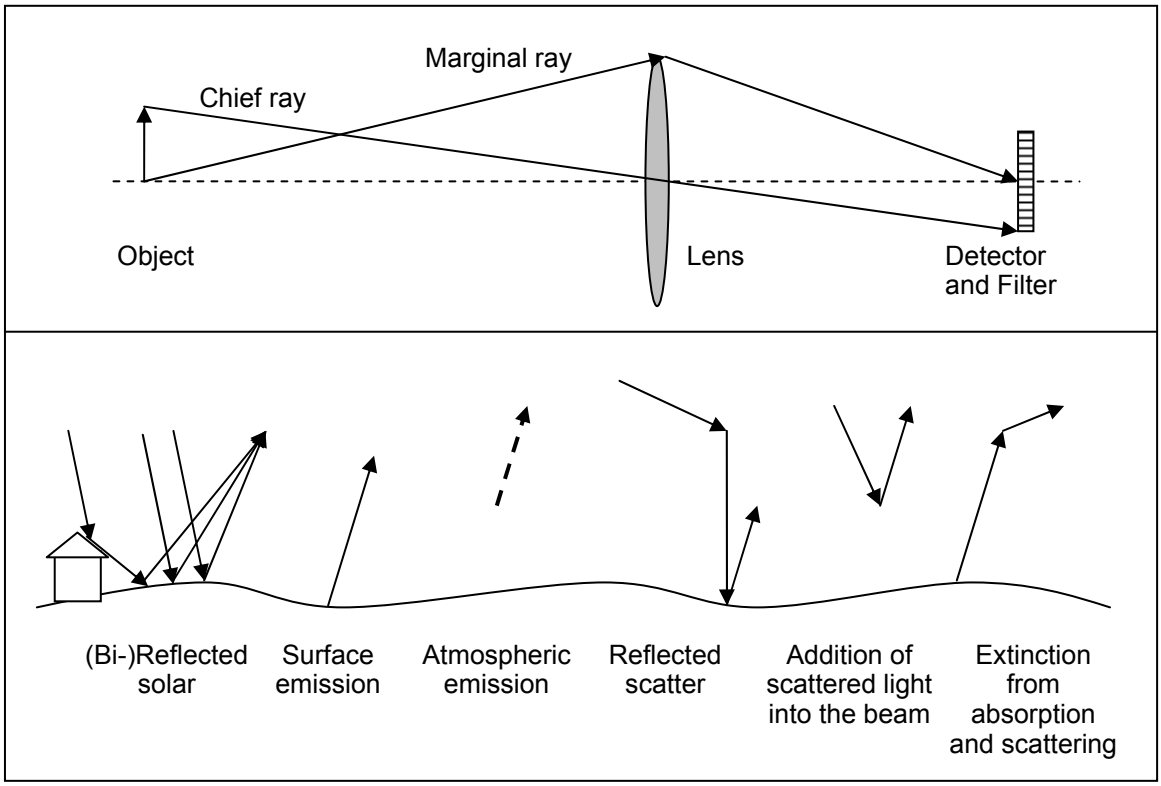


Figure 5: Object scattered light signal throughput on sensor detector

Thus, the estimated radiance (photons) the sensor receives to be able to calculate range values stems from the following considerations. As a beam of light propagates through the atmosphere, as seen in Figure 5 (after Shaw 2006), the sensor can gain light through emission and scattering and lose light through absorption and scattering. Important metrics can be used to measure the quality of the received signal:

- Signal-to-background ratio (strength of the measured signal relative to the background)

المنارة للاستشارات
www.manaraa.com

- Signal-to-noise ratio (strength of the measured signal relative to electronic fluctuations, also known as noise.

Typical noise values include:

- Johnson noise (thermal noise in an electrical device that changes the electronic signal through stimulations on the atomic level)
- Shot noise (the discretization of electron energy levels and the random nature of energy transitions causing a variation in the signal out of electronic devices)
- 1/f noise (also known as "flicker noise", a fluctuation that increases as the electrical frequency approaches zero, but so far this case is not completely understood, (Shaw, 2006))

The optimal distance information *z* based on the time-of-flight principle is:

$$z = \frac{c}{n(z)} dt$$

with a minimum range resolution of Δz

$$\Delta z = \frac{pulselength}{2} = \frac{c \cdot \Delta t}{2n}.$$

Including all noise factors, however, the optimal count of photoelectrons *n(z)* must be calculated including:

المنارة للاستشارات
www.manaraa.com

$$n_\lambda(Z) = n_L \cdot \Delta Z \cdot T_\lambda^2(Z) \cdot \beta_\lambda(Z) \cdot (\frac{A_r}{Z^2}) \cdot \varepsilon(Z) \cdot \varepsilon(\lambda) \cdot \delta_\lambda \qquad \text{(Shaw, 2006)}$$

$n_\lambda(Z)$ = Received photoelectrons at wavelength λ from range Z [# of photons]

n_L = Transmitted laser photons [# of photons]

ΔZ = Range bin [m]

$T^2_\lambda(Z)$ = Atmospheric round-trip transmittance to range Z for λ [unitless]

$\beta_\lambda(Z)$ = Atmospheric unit volume backscatter coefficient at range Z and wavelength λ [$(m\ sr)^{-1}$]

(A_r/Z^2) = Effective receiver solid angle Ω. A_r is the receiver effective area [sr]

$\varepsilon(Z)$ = Geometrical overlap factor [unitless]

$\varepsilon(\lambda)$ = Receiver's spectral transmission factor [unitless]

δ_λ = Quantum efficiency of detector [unitless]

In conclusion, this equation is valid for Light Detection and Ranging (LIDAR) sensors and shows how noise values can influence the time measurement needed for time-of-flight range measurements. In the following vision based range sensing techniques are reviewed in more depth.

2.3 VISION BASED RANGING TECHNOLOGIES AND METHODS

Vision based ranging can either be passive or active. Passive range sensors in general, e.g. single band, RGB, multispectral, hyperspectral, midwave or longwave infrared imagers, work similarly to single-lens-reflex (SLR) or film cameras, do not transmit their own energy into a scene, but use naturally present light to obtain information in single shots or multi-frame grabs. Passive sensors can be installed airborne, spaceborne, on a shore installation, or in sub-surface terrain (UUV-borne,

المنارة للاستشارات
www.manaraa.com

towed, bottom-mounted, diver-carried, etc.). Examples of passive sensor systems are Near-Infrared Passive sensor, solar-blind ultraviolet sensors, 3D color imagers, infrared noise-equivalent temperature difference sensor, hyperspectral imaging spectrometer, Fourier Transform InfraRed Spetro-Radiometer, and polarization imagers (DeWeert and Gilbert, 2006, and Shaw, 2006).

Advantages for passive sensor are their need for low power since day time light offers already a bright light source of $500W/m^2$ or more, they offer high spatial resolution, and a multi-band capability. Limitations are that the visible spectrum works best in daytime (nighttime requires sophisticated processing) and the contrast even in daytime is low (DeWeert and Gilbert, 2006).

Instead of using a passive measurement principle, active sensors transmit some form of energy (spectrum in Figure 4 into a scene to receive a return signal that may contain important information values needed to work on applications. Examples of active sensors are 3D laser scanners, dual-polarization LIDARs, micropulse Flash LADAR and IR imager, differential Absorption LIDAR, coherent LIDAR/LADAR, and Doppler Wind LIDAR (Shaw, 2006).

This further literature review in vision based ranging technologies concentrates on active sensors. The following paragraphs are a synopsis to the physical limits of optical range imaging techniques. Classifying optical range imaging techniques is based on how the spatial image is acquired and structured. Three different methods exist: (a) Triangulation, a method that uses different perspectives of cameras or light projectors, (b) Interferometry, a method that is based on standing light waves patterns, and (c) Time-of-flight, a method that emits modulated light waves in order to receive from impinged objects the backscattered range information.

المنارة للاستشارات
www.manaraa.com

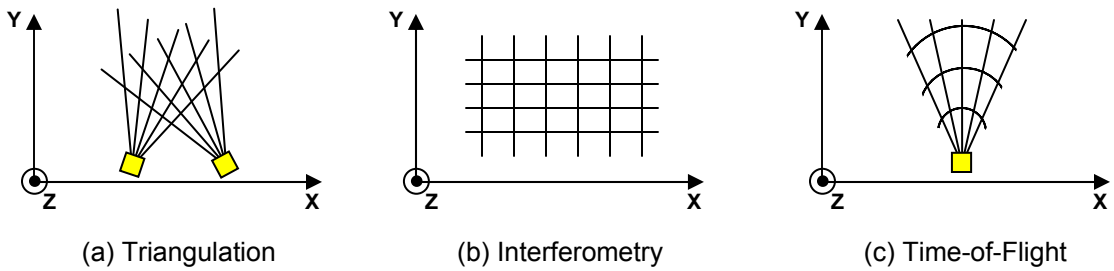


Figure 6: Taxonomy for optical range imaging

2.3.1 Triangulation, Photogrammetry, and Stereo Vision

The predominant method in nature and in three-dimensional vision based techniques is triangulation or often called "stereo vision", an indirect way to measure the third dimension. Geometrical relations between the object, a sensor and its known baseline allow to calculate distances. In passive triangulation, the position of an object in the third dimension can be determined in Figure 6(a) when each of the two or more sensors measures the 3D position of the object. In active triangulation one or more cameras functions as a projector of a patterned light that is emitted into the scene. The other camera(s) receive the returned light signals of objects surfaces that can be arbitrary, uniform and non-patterned (Besl 1988). At first, triangulation based on stereo vision requires two or more cameras and a larger space for alignment, and secondly, it may be too costly and difficult to set up in changing environments (Zywitza et al., 2005). The comparison of two distance maps requires time consuming processing (Schwarte et al., 2000). Since the triangulation base needs to be large to generate light patterns from objects at larger distances, miniaturization is almost impossible. On the other hand larger periods allow to measure ranges to objects at large distances. Commercial developments have driven this technology with high resolution, but requiring for background structure

المنارة للاستشارات
www.manaraa.com

in order to solve the correspondence problem. A more advanced or practical setup is shown in Figure 6(a). The sensor in a camera with focal length L measures the displacement x of the image of a sinusoidal pattern with a period Λ, projected by a pattern generator placed at the triangulation base b.

In summary, stereo vision based methods in general have problems in low contrast scenes or scenes that contain shadow. Stereo-vision requires elaborate computing power to make the measured data meaningful (two cameras plus powerful processing unit). These limitations cause significant delay in calculating the third dimension.

2.3.2 Interferometry

The interferometric method, see Figure 6(b), uses a standing wave pattern that is created by reflected monochromatic light (lasers) through a mirror to measure distance in the sub-micrometer range. The period of the wave pattern is half the wavelength. Plane wavefronts and mirrors lead to a rectangular coordinate system, but limit this range sensing technique to very small ranges. New developments replacing monochromatic light with lower-coherence, broad spectrum light in so-called white-light interferometry or optical coherence tomography (OCT) increase the range distance to several millimeters (Bouma and Tearny, 2001). Interferometry is commonly applied in the measurement of distances in the micrometer range and in addition requires the availability of matching wavelength and sensitivity of light emitters and receivers and at overall very high purchase cost.

2.3.3 Time-of-Flight (TOF)

Optical time-of-flight (TOF) imaging determines range from a central position as seen in Figure 6(c). This range measurement method uses the known finite speed of light ($c=2.9979 \times 10^8$ m/s in free space) to calculate the time needed to travel from the

المنارة للاستشارات
www.manaraa.com

measurement system to the object and back again at accuracies of 0.1% for distances between a few centimeters to up to several hundred meters. Time-of-flight thus works with time instead of spatial coordinates for structuring the environment (Hosticka, Seitz, Simoni. 2005). The modulation (emitting signals into scene) and demodulation (receiving and splitting signal) operates at very high frequencies making the time-of-flight measurement method very affective to noise measurement and signal processing. High frequency data, however, allows the investigation into multiple applications and areas which could not be tackled, e.g. detecting and tracking static and moving data from a single source static or moving sensor.

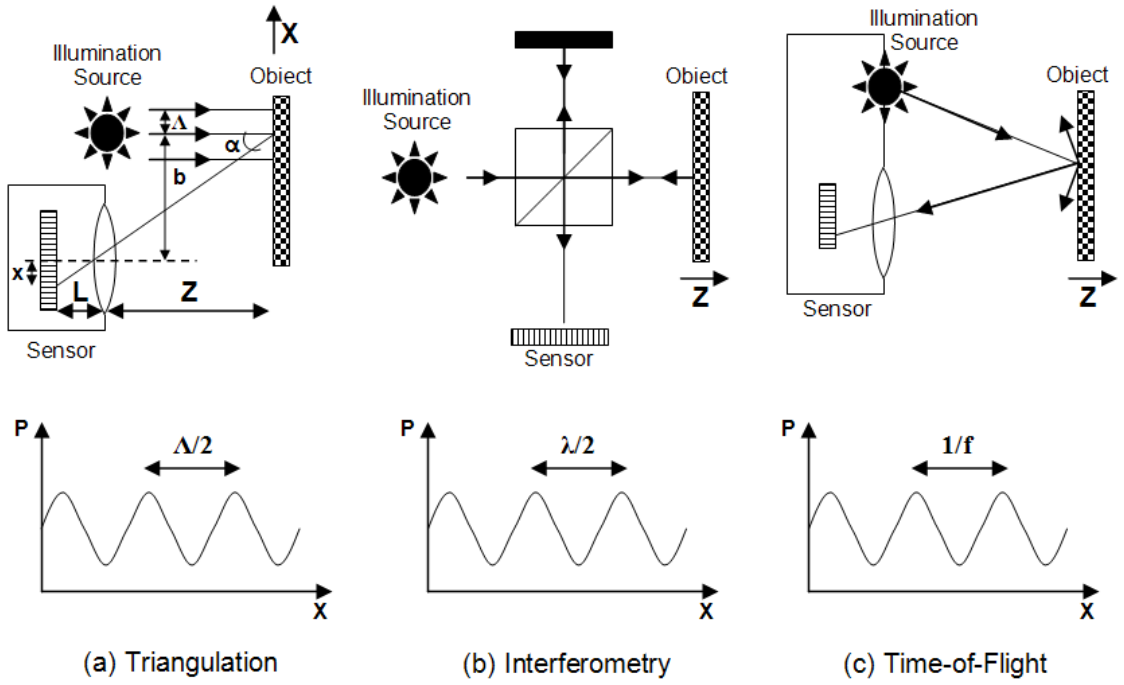


Figure 7: Illustration of three fundamental optical range imaging methods

In Figure 7 the schematic illustration of the basic methods for the three fundamental classes of optical range imaging are presented (Seitz, 2005). Figure 7a shows (active) triangulation, where a light projector generates a periodical, sinusoidal,

المنارة للاستشارات
www.manaraa.com

parallel light pattern in the sensor plane in the direction x, that falls at an angle α to the direction of displacement z of an image sensor.

Figure 7b illustrates interferometry after the Twyman-Green interferometer with a plane reference mirror (light splitter) for measuring the extent of an object in direction z. In Figure 7c, the time-of-light principle is shown. A time modulated light source is projected through a lens on the image sensor of a camera. The phase delay of each local light stream, reflected from the object, is used to calculate the distance.

In linear shift-invariant systems the phase delay of a harmonic function (emitted near-infrared signal) is required to determine the desired range information. The precision, distance resolution and the dynamic range of measurements is dependant on photon characteristics, behavior, and noise. Sinusoidal modulation (spatial or temporal) of the illumination source, e.g. near-infrared light emitting devices (LED), and a practical synchronous sinusoidal demodulation in the image sensor, e.g. CCD/CMOS lock-in pixels, however, are of special importance for all optical range imaging methods because the sinusoidal form of these function stay the same in the optical, the optoelectronic, and in the electronic domain (Lange, 2000, and Lange and Seitz, 2001). As a result, the range measurement based on coherent light is based on harmonic functions. In all three cases,

Triangulation $$d = \Lambda \cdot \frac{\varphi_{TRI}}{\tan(\alpha) \cdot 2\pi} \quad ,$$

Interferometry $$d = \lambda \cdot \frac{\varphi_{INT}}{2 \cdot 2\pi} \quad \text{, and}$$

Time-of-flight $$d = c \cdot \frac{\varphi_{TOF}}{2 \cdot f \cdot 2\pi} \quad ,$$

المنارة للاستشارات
www.manaraa.com

the formula to calculate range information is based on similar elements, like period Λ, the wavelength λ, phase shift φ, angle of incoming light α (Seitz, 2005). For each functional form, the phase shift needs to be determined. The precision of the phase shift, as stated before, is ultimately limited in the statistical fluctuations of the photon generation and measurement process.

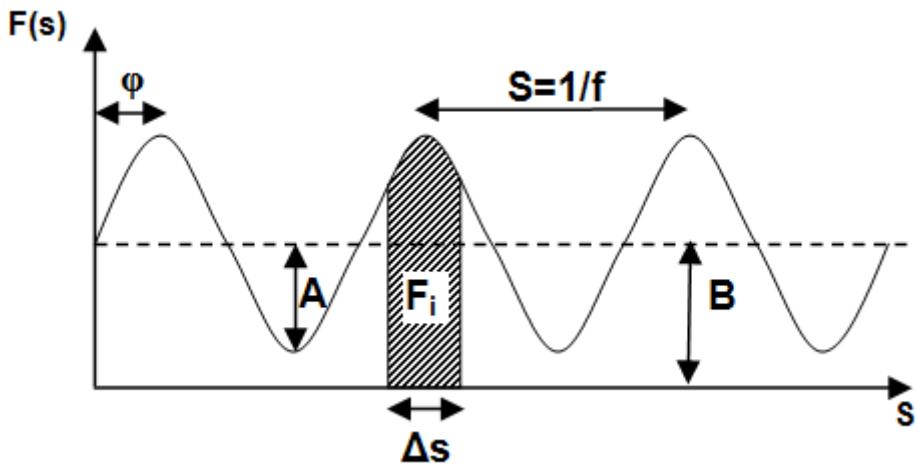


Figure 8: Illustration of the modulation parameters.

The sensor used in this research had a modulation frequency of 20MHz and emitted near-infrared wave fronts at a wavelength of 880nm. Technical details to the sensor used in this research will be explained in a separate chapter later.

As mentioned before, range values are measured based on harmonic functions. In the general case, picking or sampling signals from the backscattered harmonic wavefronts is very hard to control since the length of sampling period is too short to collect sufficient numbers of photons or photocharges for reliable range data sampling. Thus, integrating the flux of particles (e.g. photons, electrons, ions, atoms or molecules) over time for certain intervals allows capturing relevant range information. Moreover, sample signals over many periods are accumulated, e.g. 10,000 periods. Figure 8 (after Seitz, 2005)

المنارة للاستشارات
www.manaraa.com

illustrates in detail that during a sampling period Δs a flux of the returned photon wave front is measured over a minimum spatial area during the period S=1/f. Measured is a number of F_i particles (photons, electrons, etc.) for specific integration intervals. Using a four sampling technique divides an entire period S into four quarters of equal length. The amplitude A, the offset B, and the phase φ of the harmonic wave can be calculated.

$$A = \frac{\sqrt{(F_3 - F_1)^2 + (F_0 - F_2)^2}}{2}$$

$$B = \frac{F_0 + F_1 + F_2 + F_3}{4}$$

$$\varphi = \arctan\left(\frac{F_3 - F_1}{F_0 - F_2}\right)$$

Knowing the phase shift to each pixel on the sensor, the time-of-flight sensor is able to calculate the distance or range,

$$d_{PIXEL} = \frac{c \cdot \varphi_{TOF}}{2 \cdot f \cdot 2\pi} \quad .$$

Furthermore, as an important side effect of this measurement principle, amplitude and offset can be used to determine accuracy levels of the measurement, by calculating the standard deviation of the phase measurement σ

$$\sigma = \frac{\sqrt{B}}{\sqrt{2} \cdot A} \quad .$$

المنارة للاستشارات
www.manaraa.com

Example: At a modulation frequency of 4MHz, a sampling period (a quarter of the entire period) of 1μs, and a accumulation of periods of 10,000, the total integration time amounts to 10ms. Assuming to collect A=900 photons and B=10,000 background photons results in the standard deviation of σ_P=0.079 of the phase measurement, about 1.25% of the modulation period.

As a general result, the range measurement relies on the measurement of incoming photon waves. It is a particular problem if sunlight or background light overpowers the near-infrared light wave front that the sensor originally emitted to capture range information of objects in a scene.

2.4 STATE-OF-THE-ART OF OPTICAL RANGE SENSING USING TOF

Mainly there are three optical methods of acquiring range information of a scene after the time-of-flight-principle: (a) Using a stereo-system, (b) panning or rotating a 2D laser range finder or scanner, and (c) using a 3D laser ranger camera.

The acquisition of 3D data by optical sensors is preferred over alternative methods such as radar or ultra-sonic since optical techniques allow fast and (eye-) safe range acquisition at a high lateral resolution.

Emerging efficient 3D range cameras operate on lock-in pixel basis that allows acquiring and storing range and intensity data in matrices of points to entire scene in real-time. 3D range cameras do not use a scanning principle or stereo vision with complex filtering and correlation processing units that are needed for many real-time applications. The equipment itself does not require costly scanning components that require time and monetary investment, thus 3D range cameras easily can provide a distance map.

Within the 3D range camera a sensor chip is positioned to receive the incoming wave front that was reflected by the target objects in the scene or field of view (FOV). Each lock-in pixel on the sensor chip is able to demodulate the incoming optical wave

المنارة للاستشارات
www.manaraa.com

front and calculate intensity and range values based on the time-of-flight principle (TOF). Time-of-flight methods can be divided in two main categories:

- A pulse based measurement that measures the differences between the emitted and returned light directly.
- A continuously-modulated light wave that is emitted returns a phase delay once the signal is returned back on the sensor. The phase delay allows calculating the range information and to extract the range information to each pixel on the sensor.

2.4.1 Contactless Distance Measurement Approaches

The previous chapter has characterized the three main different optical range sensing methods. This chapter introduces in reflectorless range measurement approaches based on laser or light scanning using the time-of-flight principle. Laser or light range scanning can be classified in five main categories:

- Laser Radar or LADAR (Laser Detection and Ranging): Usual term for government-supported detection-related systems of hard targets, e.g. defense work.
- LIDAR (Light Detection and Ranging): Usual term for primarily commercial airborne mapping systems, typically measuring distributed scattering for environmental work.
- 3D Laser Scanner: Usual term for terrestrial/industrial use based on line the line scanning principle.
- Laser Rangefinders: Usual term for low end commercial/industrial use, e.g. Total Station.

المنارة للاستشارات
www.manaraa.com

- Flash LADAR or 3D Video Range Cameras: Usual term for an emerging technology and prototypes to be used in range imaging for real-time visualization, modeling, and simulation.

Each of these scanning methods uses the time light travels to calculate ranges. Its applications area has grown significantly in the past few years. The following sections will briefly explain each method and why and why not it may be suitable to be used in this research.

2.4.2 LADAR, LIDAR, and 3D Laser Scanner Approach

The description for LADAR, LIDAR, and 3D laser scanners is combined in one section since their technologies work very similar.

LADAR, LIDAR and laser scanning approaches find applications in the commercial field of airborne, terrestrial, and industrial sensing systems, and on the government level in airborne and terrestrial sensing systems. Applications include flood-plain mapping, forestry observation, digital preservation, cinema and video game effects, and object detection, recognition, identification, and tracking. Stone et al. (2004) listed hardware manufacturers and research prototypes in the area of optical range scanning.

Some commercial LADAR and LIDAR hardware vendor names are: Optech, Leica, Riegl, IGI, Laser Oprtonix, and TopEye. Software vendors include: ESRI, ERDAS Imagine, Intergraph, 3D Nature's World Construction Set, Applied Imaginery, Recognition, ER Mapper, Helical Systems Ltd., MapInfo, PCI Geomatics, Terrasolid Ltd., TruFlite's 3D World, and Visual Learning Systems.

A variety of industrial laser scanners are based on the TOF principle, triangulation, etc. Some hardware vendors include: Roland, Cyberware, Minolta, 3D Digital Group, 3D Scanners, IQvolution, Laser Design, ShapeGrabber, Surphaser, Visi

المنارة للاستشارات
www.manaraa.com

Image 3Dguru, and Vitronic. Software vendors are: InnovPolyworks, Raindrop Geomagic, Metrics, Rapidform, Technodigit Reshaper, and CAD/graphics software vendors like Maya, 3D Studio Max, Liuightwave, Cinema 4D, AutoCAD, Autodesk, and Bently. Terrestrial 3D laser scanner vendors are: 3rdTech, Leica, optech, Riegl, Trimble, I-Site, IQvolution, MDL, Quantapoint, and Z+F. Software vendor packages are: InnovMetricPolyworks, and I-Site Studio.

A 3D point cloud taken by a commercial laser scanner is shown in Figure 9 (Image courtesy: AMEC Earth & Environmental, Inc., and New York State, Office for Technology) in the top image. The goal of this technology is to capture a very detailed 3D point cloud, to manipulate the point cloud data for conversion into solid models. This can mean several applications, e.g. to accomplish best fit alignment, 3D mesh generation, creation of NURBS surfaces, or to allow exporting/importing files or objects. Terrestrial laser scanning has to main applications: (a) Point cloud editing and analysis, e.g. distance measuring, shape checking, analysis, and fitting, photo draping, scan merging and matching, plan and drawing integration, and decimation, and (b) in rendering, CAD model generation and viewing, e.g. 3D polygonal mesh generation, volume calculation, solid modeling, profile/cross-section generation, edge detection, standard shape extraction (pipe fitting, structural steel members, etc.), fly and walk-throughs, transparent/opaque clud/shape selection, or color detection display (Roth, 2006).

المنارة للاستشارات
www.manaraa.com

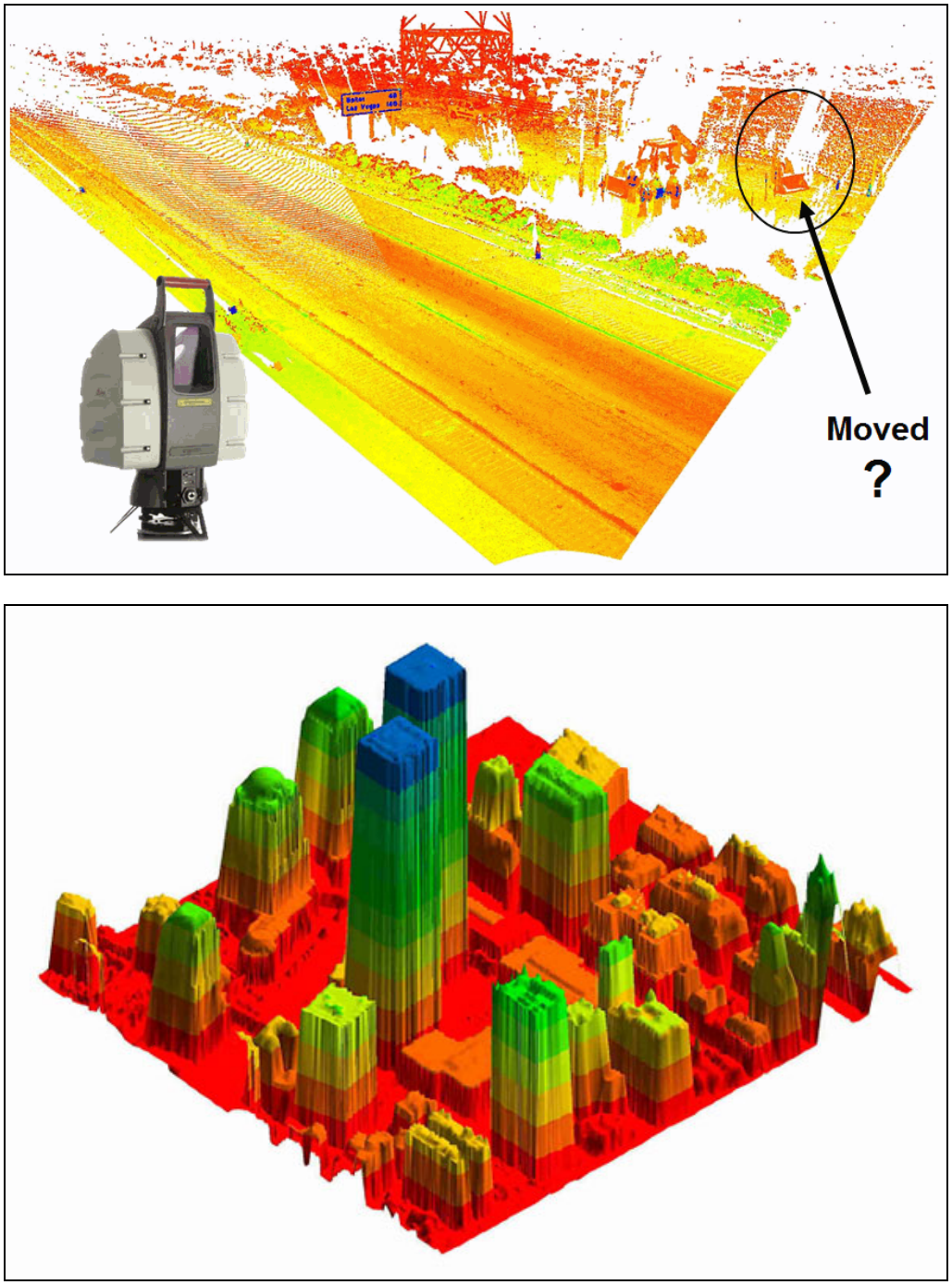


Figure 9: Laser scanner range image and LIDAR image from WTC site

المنارة للاستشارات
www.manaraa.com

In the bottom image of Figure 9 a Light Detection and Ranging (LIDAR) sensor collected geo-spatial information of the World Trade Centers in New York City before the collapse. Such an airborne based LIDAR systems use the laser beam as a photon source to collect 3D data after the phase shift principle. LIDARs can be classified by the size of the transmitted beam and received pixels, and the temporal resolution. The eye-safe operation altitude of laser class IV LIDARs is between 300m and 700m at a range data acquisition frequency of 1000Hz. The horizontal accuracy is about 2.5m and the vertical accuracy 0.25m respectively (DeWeert and Gilbert, 2006). The LIDAR approach can be used to acquire geo-spatial information. Geo-spatial applications are for example: Triangulation and distance measuring, data set merging, determination of bare-earth elevation for flood plane mapping, vegetation identification and growth, enhancement to orthophotos, urban modeling fly-throughs, photo draping, and decimation. These applications predominantly acquire distance information of static objects and require extensive data processing before distance values are transformed into 3D models.

In both cases, millions of points were acquired characterizing the scene in high detail, e.g. the street sign reads "Las Vegas 100 (miles)", however in the left image, the laser scanner was not able to capture the backhoe loader correctly, or in the right image, 3D data during the collapse could not be assessed. Focusing on the backhoe loader and asking for the error sources, the question asked is, whether the backhoe loader was moving during the range acquisition phase or the range measurement received a significant acquisition or data storing error. Similar to the count of heavy equipment in this scene (the backhoe loader may have been moved during the acquisition phase), it is not known how many workers are in the scene. If there are 1, 3, or 4 workers in the scene is unknown. Both example figures demonstrate the high accuracy of laser scanners, their work in fixed locations only, their large size, and very expensive to purchase. Needless to

المنارة للاستشارات
www.manaraa.com

say, the range data acquisition of millions of point in one scene can take up to several minutes, and data processing to generate a 3D model can take up to several days or weeks depending on the application.

2.4.3 Time Constraint and Clean Data

LIDAR and LADAR data often needs extensive post-processing of range data before using it to build a 3D model. They are very expensive to buy or rent and the use of the equipment often sees time and usage as a critical constraint. Lack of data, when the sensor is unavailable or non-existent and cost-prohibitive demands, as well as lack of "clean" data, when the level of data and scene complexity can not be set as desired, and processing tools that often are "home-brewed", all these are additional limitations of LADAR and LIDAR for real-time range sensing.

High resolution LIDAR data (1m or less post spacing) requires a precision geo-location with absolute accuracy levels of 20cm or less (Roth, 2005), thus increasing the cost of acquiring range images due to complex high pointing hardware requirements, e.g. multi-lock LIDAR (collect LIDAR data at multiple angles to have data registered, triangulated, block-adjusted and with a dense set of self-generated control points).

LIDARs in particular require precise alignment of received and submitted light beams. In general the field of view is limited (one spatial mode) that increases the system complexity. A critical part is to stabilize the laser beam and its characteristics. Gated cameras need fine tuning, such as choosing the right lens diameter, focal length, transmissivities, detector (size including number of pixels, pixel size and resolution), frame rate and exposure time, thus, making LIDAR approaches very expensive (Jong, 2006).

LIDARs or LADARs, in general contain moving parts and in general need a lot of other components to acquire range information. These lasers sensors provide range

المنارة للاستشارات
www.manaraa.com

images consisting of a set of point-measurements. Usually, range images are acquired from one view-point by "moving" the laser beam using rotating mirrors/prisms. The orientation of the laser beam can be easily measured and converted into coordinates of the image. Another possibility of acquiring a range image is moving the entire setup through an environment and measuring with a 2D laser orthogonally to the motion trajectory. As this description depicts, a laser scanning system requires complex efforts to collect range information.

2D laser scanners deliver a range by scanning one line after the other of a scene and adding it to an array. Since laser scanners do not deliver complete range images, the range sensors need to be panned or rotated to reconstruct an entire image. This scanning principle is basically a raster scan that scans through the scene, and requires complex mechanics to perform so. The challenging and time consuming part is to synchronize the acquisition of single lines with the motion of the sensor. Thus, laser scanners, are not only very expensive to purchase, but also are unable to rapidly asses range information of entire scenes in real-time.

3D laser scanners have a similar working principle and in general are very bulky and expensive. Most rangefinders employ a linear opto-electronic conversion to convert the received optical signal to the electronic domain. PIN-diodes and avalanche-photo-detectors (APD) are two examples that convert the impinging photon with a certain probability (the quantum efficiency) to an electron-hole pair, which then cause a microscopic current flow. The detectors, thus, produce an electronic signal that closely follows the received optical signal (plus some added noise) (Rohner et al, 2005)

To create rapid 3D models faster range data acquisition and processing is needed, especially once 4D CAD models are involved that require rapid data processing to meet the simulation criteria of being real-time. Collecting a few points to characterize objects

المنارة للاستشارات
www.manaraa.com

in a workspace, also called sparse point cloud approach, was a next step to create 3D models more rapidly.

2.4.4 Sparse Point Cloud Approach

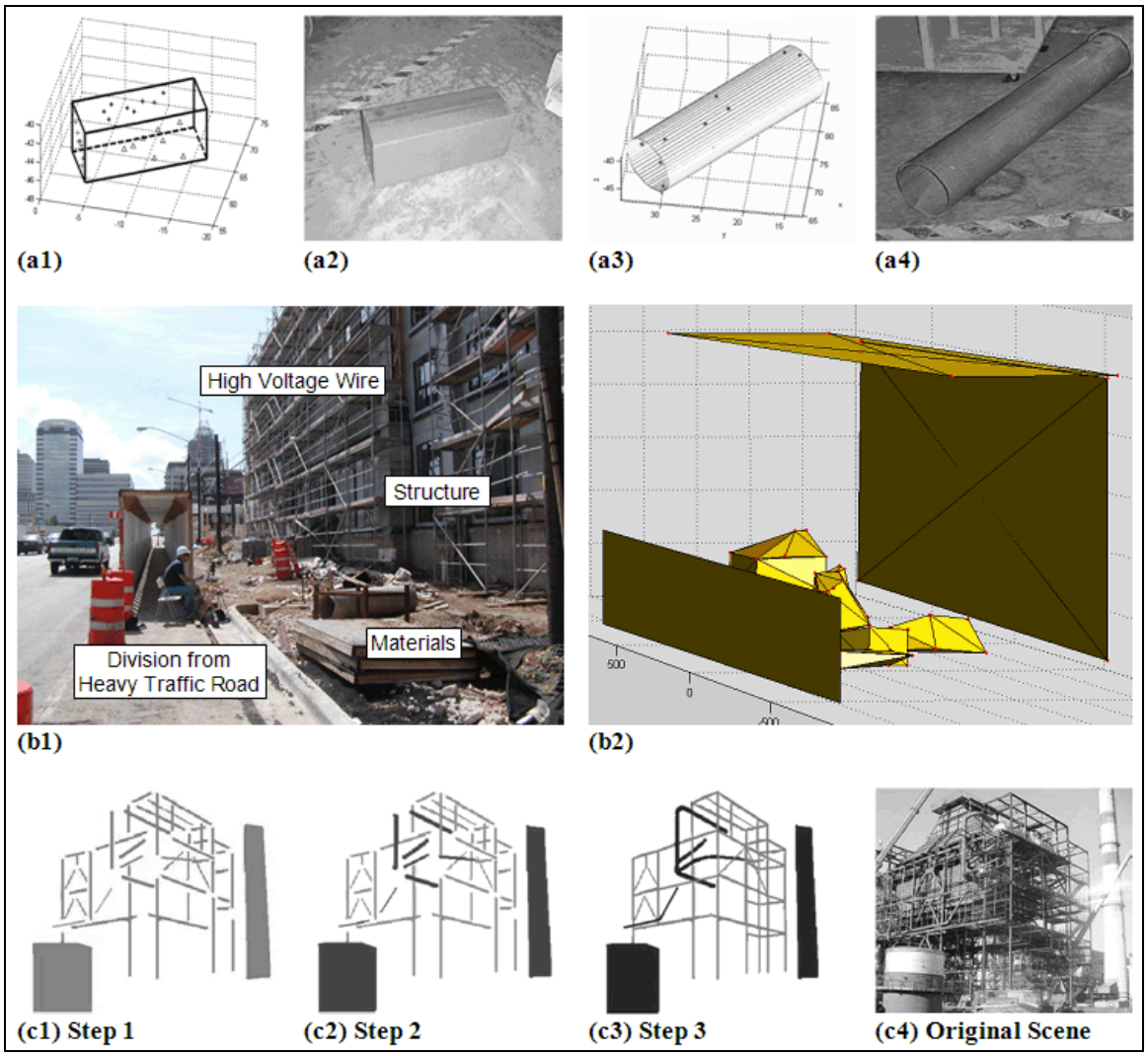


Figure 10: Sparse Point Cloud Approach: (a1) Fitted and matched cuboid, (a2) Actual object, (a3) Fitted and matched Pipe, and (a4) Actual object. Bounding box generation: (b1) Actual objects, (b2) Bounded objects. Merging primitives can improve model display: (c1) Step 1: No object recognition process, (c2) Step 2: Object distinction process, (c3) Step 3: Object distinction and object reconstitution process and (c4) Picture of the scanned scene (Teizer, 2005a).

المنارة للاستشارات
www.manaraa.com

The sparse point cloud approach focuses on selected points to avoid high computational costs of dense range point cloud information and therefore requires only a few minutes to model a scene. The sparse point cloud approach is based upon three basic transformations: (1) Fitting sets of range points to CAD primitives, (2) creating bounding objects, and (3) merging and compliance checking. It can produce strong, weak, or non-parametric forms. Human intervention is needed in all three steps to select meaningful points from a cluttered scene. Figure 10a demonstrates that an abstract model with a collection of primitive strong forms will suffice as building blocks for a description for many applications (Kown et al., 2004, and McLaughlin, 2004).

Bounding algorithms allow for grouping all range points into bounded objects such as convex hulls (Figure 10b). This is a process of abstracting or simplifying non-parametric and weak forms or cluster forms. It is useful for real-time applications because it minimizes computational burden. Use of obstacle avoidance calculations in the background in real-time can improve equipment operation and safety. When objects are related to tasks, object fitting, matching, and merging algorithms can be used to extract precise geometrical information from workplace scenes. Such spatial modeling can be applied in obstacle avoidance operations of heavy equipment (Kim et al., 2004).

A coherent view and idea of objects is based on relating individual parts to a world model. Figure 10c illustrates the use of developed algorithms for integrating merging and compliance checking capabilities into site modeling methods to improve the model's value for communication. The limited view of the range finding instrument from a single reference point makes multiple sets of range data and their corresponding model subsets necessary in modeling a workspace. Merging all subsets from different locations into a world model of the workspace requires that transformations and merging take place at the range point, geometric primitive, convex hull, and subset level.

المنارة للاستشارات
www.manaraa.com

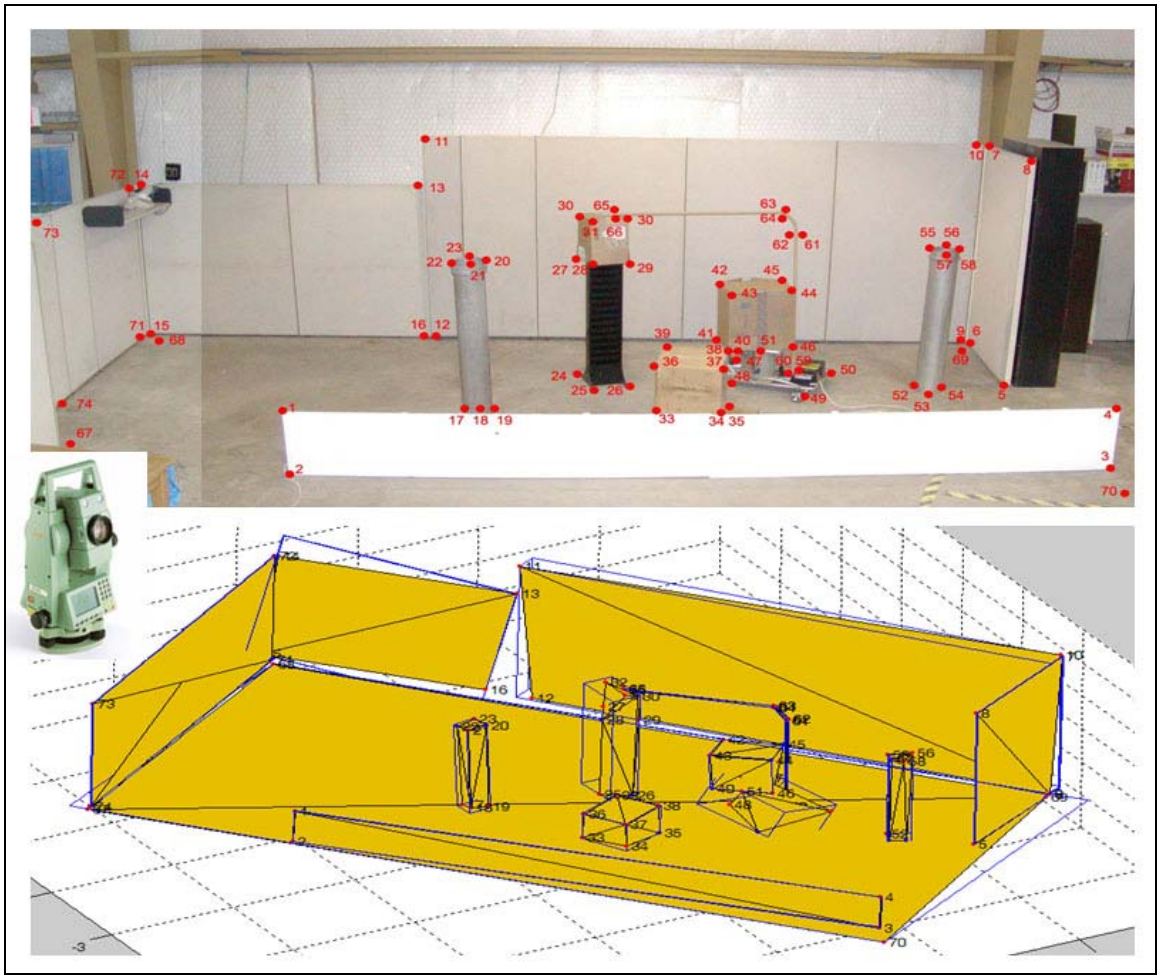


Figure 11: Sparse point cloud modeling approach using a Total Station

As Figure 11 illustrates, taking single range points that characterize objects in a scene can take up to several minutes. In this indoor laboratory example a reflectorless Total Station from Leica® TCR 1205 R300 with an angular horizontal and vertical accuracy of 5” (1.5mgon), and range accuracy of 3mm at distances smaller than 500m was used to measure one range distance after the other to a total of 74 scene and object relevant points. The entire range data acquisition step was performed by the author within 30 min (Teizer, 2006). Once all range points were taken a convex hull algorithm created in Matlab™ processed all range points within less than a second.

المنارة للاستشارات
www.manaraa.com

As a result, the major limitation (and power) of the sparse point cloud approach is the requirement for human judgment and the focus on static environments. Judgment is used in the process of acquiring distinct range point clouds. This enables rapid modeling of the static elements of work spaces, however moving objects can not be captured without distracting the operator. Thus, detection and avoidance of moving obstacles still requires full operator attention.

2.5 REVIEW OF RANGE DATA PROCESSING METHODS

The acquisition of range data technology and methods has been explained in the previous paragraphs. To ultimately build 3D models for various applications, data processing is needed to extract scene information such as number, position, dimension, direction, and velocity of objects in the field-of-view of the range sensor. The resulting information then can be applied, e.g. in the detecting and tracking of potential targets in an obstacle avoidance system

This paragraph reviews different data processing approaches that are important to process range information. At the end, this review leads to the basic framework for the preferred data processing method or also called "algorithm" of choice. Although the objective in this research is to use range data only to create 3D models, the review of existing data processing includes approaches based e.g. on intensity data, surface normals, 3D sampling, noise evaluation, clustering, or object merging and tracking. Such a review approach enriched the development of the real-time object detection and tracking algorithm. The algorithm which was developed in this research combines several ideas of existing data processing techniques mentioned in the following.

Advances in algorithms, e.g. clustering, are broad and come with advantages and limitations. To fully meet the research objective this review concentrates on range and on selected intensity data processing methods that both can be helpful in the development of

المنارة للاستشارات
www.manaraa.com

new algorithms and are compatible with the available prototype 3D range camera. Once the algorithm for this research is developed, it is expected to be advantageous in use for specific situations, and on the other hand, to be less useful in other applications.

2.5.1 Model-Based and Data-Driven Approaches

A model-based processing approach transforms data to geometrically simpler models. In construction this can be for example, building data converted to boxes with image textures, or curved surface data creating NURBS (Non-Uniform Rational B-Spline, a mathematical representation of a 3-dimensional object), or refinery piping scans transformed to a CAD pipe model. Usually the model-based approach requires computationally/manually intensive fitting and relaxed rendering, and is thus not very appropriate in applications like Automatic Target Recognition (ATR) or for the creation of video games, etc. where real-time photo-realistic visualization is required (Roth, 2006).

Data-driven approaches, like in this research, focus on enhancing the data and visualize it directly. Assumptions about the data is minimized and thus revealing information from the data is appropriate, e.g. to find and characterize objects in data.

2.5.2 Occupancy Grids

Moravec and Elfes (1985) developed the concept of two-dimensional occupancy grids, a well known technique for the detecting of obstacles in mobile robotic applications. Originally they used multiple single wide-angle sonar range measurements providing information where empty and occupied volumes were located in the FOV of a sensor. A rasterized map of the surroundings of an autonomous mobile robot was created from the processed range readings using probability profiles of where occupied and where empty areas were located. Range measurements from multiple points of view,

المنارة للاستشارات
www.manaraa.com

orientations, and at different time stamps were systematically integrated in the map. Accuracy of detecting obstacles in the sensor's FOV depended on the frequency of distance measurements. Frequent updating of range measurements improves the certainty of empty or filled spaces over time, since overlapping empty volumes reinforce each other, and serve to reduce multiple range readings for each cell to one occupied volume only. A final two-dimensional map shows regions probably occupied, probably unoccupied, and unknown areas. Murray and Little (1998) and many other research successfully built occupancy grids based on intensity data from real-time stereo vision for mobile robot navigation and an effective handling of clutter. Moravec (1996), for example, worked with evidence grids in 3D space based on stereo vision.

2.5.3 Data Clustering

The task of choosing the proper clustering algorithm is a significantly important step, because the large amount of range data needs to be segmented fast in order to give rapid feedback to the application (e.g. 3D CAD model generation). To ensure safety, the detection of all objects in the scene is necessary (e.g. obstacle avoidance system) and that each object is in its own cluster (e.g. detect and track movements in different direction). Five commonly known types of clusters exist (Tan et al., 2005):

- Well-separated clusters: Sometimes use thresholds to specify that all points in a cluster must be significantly closer to one another than to any other points in a different cluster.
- Prototype-based (Center-based) clusters: Often have a centroid, i.e. average (mean) of all the points in the cluster where each point is closer to the centroid than to any other centroid of a different cluster.
- Graph-based clusters: Connect points that have defined distances to each other.

المنارة للاستشارات
www.manaraa.com

- Density-based clusters: Are clusters of high density separated by regions of low density.
- Conceptual clusters: Share some general property that derives from the entire set of points.

Three simple but important clustering techniques exist following the above mentioned clustering types:

- K-means (prototype-based clustering finding user-defined number of clusters)
- Agglomerative Hierarchical Clustering (all closest points as singleton clusters are merged until an entire image is clustered, large computational cost on entire data sets)
- DBSCAN (Density-Based SCAN clustering algorithm producing partial clustering with automatic detection of cluster number, points in low density regions are classified as noise and omitted)

The agglomerative clustering linkage methods commonly used include the following metrics of linking points to groups (Guralnik and Karypis, 2001):

- Simple Linkage: Groups of the smallest inter-group distance between any pair of objects are linked into different shapes.
- Average Linkage: The two groups with the smallest average inter-group distance are linked. Groups may then have similar diameters in each metric direction, though each group can be a different size.

المنارة للاستشارات
www.manaraa.com

- Complete Linkage: Group of smallest maximum separation and maximum inter-group distance between all pairs of groups inter-group groups have same size but different shape.
- Wards Method: (links the pair of groups that produce the smallest variance in the merged group, most CPU intensive method)

Partitional clustering segments objects in non-overlapping subsets (clusters). Permitting subclusters refers to hierarchical clustering which is a set of nested clusters that are organized as a tree. Each node (cluster) in the tree (except for the leaf nodes) is the union of its subclusters, and the root of the tree is the cluster containing all the objects. Exclusive (vs. overlapping or fuzzy) clustering allocates each point to one object or one cluster. Complete clustering assigns every object to a cluster, whereas partial clustering does not (Tan et al., 2005). In construction environments the complete scene may be segmented and range points belong to one object only.

As a result, (graph and) density based clustering approaches are more suitable to segment range information. K-means, as a prototype and partitional clustering method, does not determine the number of clusters automatically, but requires a priori knowledge of the number of objects in a scene to find a meaningful computation for cluster centroids. Other clustering techniques, such as many agglomerative hierarchical approaches do not require the underlying nature of the data sets and strictly defined attributes.

To choose a clustering algorithm in this research, a look at the data set and main application is needed. The 3D range camera (SwissRanger 2) as range sensor, delivers 124 by 160 (or 19820) distance points at a maximum frame update rate of 30Hz. These large data sets (storage requirement of about 160 kilobyte per second) do not allow a

المنارة للاستشارات
www.manaraa.com

hierarchical agglomerative clustering technique to be applied directly on the raw range data unless time is not a critical issue (Tan et al., 2005). Filtering noise and outliers is one option, but in general, can be literally dangerous in this research approach, because the elimination of data points may erase objects that are at a remote area in space or are of smaller size. Overall, clustering must run successfully and produce clusters that represent the structure of the original scene (Teizer et al., 2005a). A next thought needs to be directed at what parameters describe the clusters best and how this information is later on used to detect and track movements.

2.5.4 Image and Model Registration

The inaccurate knowledge of platform attitude and position and the need to estimate either absolute or relative positions and orientations among all images and models, make object registration and tracking based on intensity values very challenging.

Two examples are image-to-model registration for accurate texture mapping (image matching to other intensity images before to the model, extract lines and search for matches, predict shadow positions, synthetic imagery and approximate sensor parameters in terrain) and model-to-model registration for building consistent 3D models from 2.5D range scans. Some solutions in the literature refer to:

- Iterative closest point (ICP): iteratively refine rigid-body transform parameters by repeatedly selecting pairs of corresponding points and minimizing an error metric, requires reasonable initial alignment, can perform well with little overlap (Besl and McKay 1992, and many variations afterwards). Surmann et al. (2003) reconstructed abandoned mines by aligning point clouds using a three-dimensional ICP algorithm.

المنارة للاستشارات
www.manaraa.com

- Feature-based methods match geometric primitives (e.g. points or lines) or more complex invariants (e.g. shape descriptors) to 3D model data. Methods exist where no initial alignment or significant overlap for lines is required (Stamos, 2000), for points (Haralick et al., 1989), using spin images (Johnson and Herbert, 1997), and 3D maps using Extended Kalman filtering (Jensen et al., 2005).
- Optimization with multiple models: Sequential model adjustment (Chen and Medioni, 1991).
- Simultaneous modeling (or iterative model adjustment methods that distribute error uniformily) (Pulli, 1999, and Hsu, 2003)

Since the correspondence among images and models often is determined manually, which is time consuming and generally inaccurate, automated methods to accurately estimate the correspondence are maturing (Roth, 2006).

2.5.5 Tracking Objects

Humans are usually easily overwhelmed with detecting and tracking several objects at the same time. Ultimately flexibility in a computer based approach is desired.

Tracking is the detection of similar or the same features of an object over time. The purpose of tracking is to build a time history of object features to help analysis scenes and react to special behaviors, such as: Position vs. time (position, velocity, and acceleration), orientation vs. time, shape vs. time, intensity vs. time, range vs. time, velocity vs. time, object dimension vs. time, and separation distance and angle between objects.

Different approaches to solve the detection and tracking assignment are identified: Single hypothesis approaches (Nearest neighbor, Greedy, Global Nearest Neighbor (GNN) that examines all possible objects and track pairs), PDA/JPDA (Assign all

المنارة للاستشارات
www.manaraa.com

detections to all tracks, is good for widely separated targets in heavy clutter), MHT (Assign all detection to all tracks, improved performance at greater computational cost). Sanders-Reed (2006) comments on these approaches that they have typical problems in detecting and tracking of objects that are: Multiple count of up to 100, appear, disappear or spilt, resolve or unresolved, objects that maneuver, crossing of trajectories, require correct temporal histories, or that have a erratic sensor line-of-sight.

An important question in object tracking is to define which feature should be tracked. In general several features exist which are worthwhile to be tracked. Some features are only available by using certain image processing methods, e.g. contrast detection. Depending on the application some features are: Centroid location (e.g. weighing all pixels equally, or based on intensity), peak intensity location, leading edge location, outline shape, minimum and maximum extents in x, y, and z dimensions, second moments (moments of inertia) to give simple shape description, peak intensity, total intensity (sum of pixels), peak signal-to-noise ratio (SNR), number of saturated pixels, number of "bad" pixels, range pixel position, range value, dimension of similar range values, velocity of range pixels, direction of range pixels, derivation of positions vs. time to obtain velocity and acceleration (smoothing potentially required), frequency analysis of position, separation, or intensity, and velocity and angle distribution for each object.

2.5.6 Classification of Objects

In segmentation and classification for LIDAR images for example, features can be extracted and classified, e.g. bare-earth extraction can filter the bare earth from the data points while rejecting vegetation and man-made objects such as buildings (Sithole and Vosselman, 2003). Another step further enables to separate between vegetation and building or natural and man-made objects. Building shapes, dimension, and position, tree volume, stem height and diameter and crown diameter were modeled by Holmgren

المنارة للاستشارات
www.manaraa.com

(2004), templates of building structures were matched based on intensity values by Iqbal and Aggarwal (1999), and Haala (1999), and a data driven method in the same field was developed by Södermann (2004).

The algorithm in this research has not the goal to identify or name objects. Since the algorithm is designed for detecting and tracking objects, however, it includes the partial availability of features (e.g. dimension of objects and volume) that can allow the implementation for identifying objects later on.

2.5.6 Modeling Approaches of Large Data for Large 3D Terrains

Another 3D modeling approach involves surface generation and feature extraction. The most common form of surface and object generation from irregular point data, such as the 3D video range camera provides, is the Delaunay triangularization method. To create a 3D model, the Delaunay method is according Figure 1 a non-parametric approach that chooses the largest angles to build triangles that encompass the object surface. Time-efficient Delaunay algorithms exist, e.g. “order n log n”, or QT Viewer speedup (Roth, 2006).

One challenging modeling approach is to present displays in a rapid manner such that objects are automatically revealed. This task asks for cutting through interferences and distractions, e.g. to pick out a tank underneath a tree. Geometric or radiometric filtering and culling is necessary to enhance the visualization and modeling.

Complex geometry in scenes can slow down the modeling speed significantly. Typical real-time application (e.g. video games) use texture maps applied to relatively simple geometry to create complex-looking scenes. Visualization of large geometric models usually can have over a few million vertices making it extremely difficult to operate near real-time and need the help of supercomputers. Several alternatives to large-model visualization and modeling have been proposed, such as level-of-detail culling,

المنارة للاستشارات
www.manaraa.com

image chips ("imposters") to substitute for distant geometry, and surface pixel ("surfel") sampling (Roth, 2006).

Similar to recent research in large digital elevation models (DEM's) more efficient coding methods are required in this research to reduce the need for memory processing, e.g. the structure of general 3D data occupies an order of magnitude more data than a 3D vertex and the indices for two angles. Approaches in new data structures are undertaken, such as Octree (divide rectangular regions into eight quadrants can help to visualize up to 200M samples of 3D data in real-time on regular PCs), or binary triangle tree (bintree, to divide triangles into two halves). Luebke (2003) discusses popular methods to find detail for terrain rendering and modeling is the Continuous Level Of Detail (CLOD) approach discussed by Lindstrom (1996), the Real-time Optimally Adapting Meshes (ROAM) by Duchaineau (1997), Real-time CLOD by Rottger (1998), View-Dependent Progressive Meshes (VDPM) by Hoppe (1998), Multi-Triangulation (MT) by Puppo (1998), and Visualization of Large Terrain Made Easy, by Lindstrom (2001). Wang et al. (2003) built 3D models of urban areas.

2.5.7 Localization for Three-Dimensional Maps

A 3D obstacle avoidance algorithm needs to work when the sensor is in static position or in motion. Self localization and continuously updated maps for autonomous vehicles require frequent and accurate updates of location, pose, and direction, to navigate vehicles without hitting obstacles. This is known as Simultaneous Location and Mapping (SLAM). Usually on-the-fly range information is collected (Jensen et al., 2005). Since a vehicle with a sensor might continuously move, sparse range information of wide angles is preferred over high-resolution dense point cloud approaches that need longer acquisition and processing time. Extended Kalman filters can be used to predict the robot pose along with its error.

المنارة للاستشارات
www.manaraa.com

Hähnel et al. (2001) build 3D models from a mobile robot equipped with a 2D localization system and vertically mounted laser scanners. Thrun et al. (2003) used large-scale 3D models for helicopter mapping. Jensen et al. (2005) and many others developed algorithms for large-scale 3D models from a set of 2D laser scans using rotating laser scanners and laser range finders. Globally registered range data from several viewpoints in indoor and outdoor scenes was collected and processed to estimate the pose and travel of a navigating robot.

2.5.8 Obstacle Detection and Motion Detection Methods

Singh and Keller (1991) describe obstacle detection as the determination of whether a given space is clear from obstructions for safe travel by an autonomous vehicle. It determines the existence of a certain type of objects and finds their locations. Three goals have been identified: (1) to detect obstacles in a timely manner, (2) to identify obstacles, and (3) to track objects which are in the motion path of the vehicle.

Hoover et al (2003) worked on the problem of estimating the motion of the camera between the sensing of two images of a range camera. The focus on indoor planar surfaces which remain stationary between the acquisitions of two range images is based on three steps: extract features, correspond features between views, and compute motion. Hoover et al. concluded that noise in the original range data is overcome through surface modeling and the imperfections in the segmentation process are overcome through explicit handling in model construction and motion estimation. Edges as secondary feature and a volumetric description of space (called space envelope or the volume of viewed empty space) where not examined.

In related work, surface features were extracted and corresponded them against a priori known models, for example for object recognition by Arman and Aggarwal (1992). Their review presents obstacle detection methods and algorithms. The authors describe

المنارة للاستشارات
www.manaraa.com

that range sensors generally provide vast arrays of range data, but as such "raw" data is not readily usable, low-level processing are necessary in order to extract a better representation of the data. In their study and in contrary to the proposed research, the methods used by the authors create models of objects using a priori information, such as the geometry and topology of the object, and provide this knowledge to the recognition system. However, many representations are possible, and many of them are used in computer-aided design applications.

Less of these works considers the problems caused by potential mis-segmentations (Johnson and Herbert, 1997). Even for planar surfaces, perfect segmentation across a large number of images is not attainable with currently known algorithms (Hoover et al, 1996). An overview of algorithms for computing motion estimation from 3-D point correspondences may be found in Blostein and Huang (1987).

2.5.9 Using Intensity and Range Information to Create 3D models

As the range camera used in this research provides intensity and range values, the approach of Stamos and Allen (2003) to model large buildings allows creating a true geometric CAD model. Their modeling process is describe: (1) Acquisition of multiple range scans and 2-D images of the scene, (2) The range scans are segmented into planar regions, (3) 3-D linear features from each range scan are automatically found, (4) The segmented range data from each scan is registered with the other scans, (5) Each segmented and registered scan is swept into a solid volume, and each volume is intersected to form a complete, 3-D CAD model of the scene, (6) Linear features are found in each 2-D image using edge detection methods, (7) The 3-D linear features from step 3 and the 2-D linear features from step 6 are matched and used to create a fully textured, geometrically correct 3-D model of the scene. However, Stamos and Allen do not mention the time needed for modeling.

المنارة للاستشارات
www.manaraa.com

2.5.10 Image Processing Techniques Based on Intensity Data

The following list briefly reviews some of the commonly known and used techniques in image processing based on intensity values. Knowing that 3D range sensors can collect both, range and intensity information, it is important to review these techniques, because some principles can be applied in the development of a range processing algorithms.

- *Frame Subtraction:* The subtraction of two frames allows eliminating the static foreground and background. As moving elements can be rapidly found, noise still exists and may be filtered with a threshold. The subtraction of similar range values in range frames that are timely very to close to each other may eliminate most range data, including the object's range information. Sato and Aggarwal (2004) successfully segmented and tracked humans from background.
- *Thresholding:* Thresholding means to apply a value on all range information and filters those values which do not meet the threshold value or condition. If the objects' size is known the thresholding parameters can be set more precisely.
- *Smoothing (Median Operator):* Smoothing is appropriate to reduce error measurements. Some dead pixels with no range value or outliers can be filled with the average range value of surrounding pixels. Thus, error measurements do not have any longer a high impact on the range image. However, if regions contain a lot of noise, good range values may be manipulated, and the range image may loose some of its accuracy, e.g. smoothed edges could be not as significant as before. Median smoothing offers a good practical solution if noise needs to be removed first.

المنارة للاستشارات
www.manaraa.com

- *Region Growing:* The goal is to find regions which have the same distance. Dividing the Grouping single points of almost same distances sitting next to each other can simplify range images. An area needs to have at least 10 or more points to be considered as separate area.
- *Classifiers:* It is doubtful that complex classifier methods are successful, as often training of the samples is required. However, the feature of objects on construction sites are multiple (but most likely repeating), and additionally training and evaluating of sample data does not meet the research criteria of time. ZuWahn's research (2001) in building detection from a vertical monocular viewpoint showed that neural networks, Bayesian classifier and Bayesian network had approx. 80 % detection success at false rate of minimum 10% at this level of detection. Both are unacceptable in the need for a simple but accurate and fast object detection system. Popular classifiers are Naive Bayesian classifiers and Bayesian networks. A naive Bayesian classifier is a simple classifier based on the conditional independence assumption among evidence, which shows good results for a number of problems despite the strong assumption.
- *Range Image Segmentation:* After cleaning the image in the pre-processing stage, segmentation is a central point in understanding the range image better. The goal is to separate the image into meaningful regions by using classification and then to enable the identification of single objects within each segment. Fast segmentation is successful, if (a) the object or timely very close frames shots of the object move in the same direction, or (b) segments changes minimal between two frame shots and can be used for correlation. Such segmentation is difficult as each frame relies on reduced dynamic. The better the pre-processing step is executed the better the segmentation (Gonzales and Woods, 1992). A

المنارة للاستشارات
www.manaraa.com

conventional method in segmentation is background subtraction in which each pixel is used to model the background. Any pixel changing in the next frame is then taken to be part of the foreground. Typical requirement is a fixed camera view with both depth and intensity image statistics (Gordon, 2002). Shum et al (1994) divide range image segmentation techniques three approaches: (1) Feature-based approaches such as Gaussian and mean curvatures are precise and used to label different regions before region growing, but sensitive to noise; (2) primitive-based approach focus on primitives and thus are more resistant to noise, however, the segmentation success gets worse the higher the degree of surface polynomial gets; and (3) statistics-based approaches which achieve high segmentation success rates, but are slow in computation.

- *Discussion of Edge Detectors:* 3D edge detection can be one step in the processing methods used to distinct the range image into several objects. Singh and Keller (1991) used edge operators in the detection of large obstacle on roads and found that edge findings were not very distinct in range images. The problem occurs when the model must be fitted to experimental data. Roberts operator produces thinner edges and tends to break up in regions of high curvature. Monadic, dyadic, and local operators can be implemented in k*m*n seconds. Disadvantage is the high sensitivity to noise, since fewer pixels are used in the calculation of the edge detection. There is also a shift in the image which is not produced by the Sobel operator. The Prewitt operator is similar to Sobel but more sensitive to noise and does not possess the same inherent smoothing. *Rank filters* operate to detect edges, reduce noise, sharpen edges and enhance images as well as *direction of the intensity gradient*.

المنارة للاستشارات
www.manaraa.com

- *Tracking:* Tracking is required to know the location of objects and to determine its relative velocity. To allow a fast processing method of range data, Singh and Keller (1991) propose that by projecting the path onto the image, a large portion of the range image can be ignored. This may be contrary to standard methods of using range data e.g., as stated in Daily et al (1987).
- *Reverse Engineering and Virtual Reality Techniques:* Various research approaches successfully concentrate on the digitization of the surface of real three-dimensional objects. Johnson and Herbert (1997) created spin-images to identify single non-flat objects such as sculptures, faces, and ducks in 3D. Karbacher et al (2001) described the general modeling steps as data acquisition, surface registration, including computationally intensive mesh reconstruction, data modeling and smoothing, and color reconstruction. Both approaches have a slow data processing speed and do not meet the need of modeling multiple objects quickly.
- *Analyze of Volume Growth:* Volumes of clustered range points to one object can be modeled by using bounding hulls. A bounding hull can be created by starting from the center point and building a convex hull from a point to the next closest point (or a lower incremental volume). The result can be a box, a sphere, or other forms. To track same objects the rate of volume growth can be used and measured from one range frame to the next.

2.6 SELECTION OF RANGE SENSING TECHNOLOGY & DATA PROCESSING METHOD

This summary reviews the selection of the of range data acquisition technology and processing method used in this research.

المنارة للاستشارات
www.manaraa.com

2.6.1 3D Range Camera for Range Sensing

Dense point cloud, sparse point cloud, and 3D range camera are methods for range data acquisition and have different characteristics to produce different results. A meaningful comparison of the three modeling methods can be made on the basis of the following criteria:

- Density of data used in modeling (a higher density offers a wider field of applications).
- Frequency of updating of the derived model (allows fast or real-time modeling updates).
- Precision and accuracy (how well the model reproduces the actual scene).
- Richness of the derived model (information quantity and quality incorporated into the model).

المنارة للاستشارات
www.manaraa.com

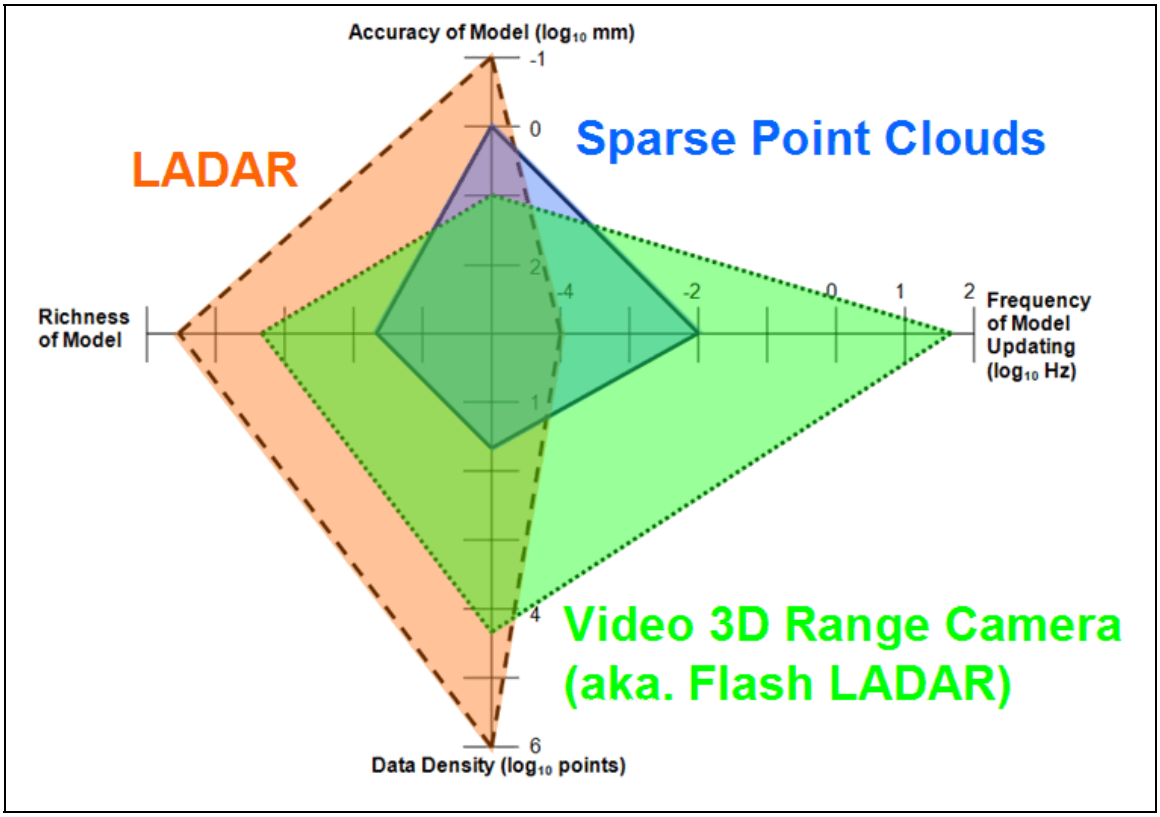


Figure 12: Preferences of optical range data acquisition technologies

Figure 12 presents a schematic diagram of the existing methods according to these criteria. While dense point cloud approaches are very precise but slow and expensive, the sparse point cloud approach tends to achieve a compromise between accuracy and speed that is useful for some real-time field applications and can be performed at a much lower cost (Kim et al., 2004). Based on the current technology development, the biggest and unique advantage of using a 3D range camera/Flash LADAR is to track moving objects in real-time of up to 50 Hz (approx. at TV or video image update rate). The given accuracy in ideal condition is in the range of millimeters and is expected to change in construction-like environments. 3D range cameras achieve an average to high data density (data array of 128x160 points) compared to up to 20 cumulative single

المنارة للاستشارات
www.manaraa.com

measurements taken by single axis range finders used in the sparse point cloud approach. Ultimately the array of measured points allows imaging complete objects in real-time. As applications define the requirements for choosing sensing equipment, the scale for richness of the model is not quantitatively described.

Table 2: Ideal and existing attributes of range sensing solutions.

Sensor Attribute and Units	Ideal Value (extended, after Stone et al, 2004)	3D Range Camera (ext. after Weingarten et al. 2004)	Dense Point Cloud (LIDAR, LADAR, 3D LS)	Sparse Point (Single Range Finder)	Importance			
					Obstacle avoidance	Navigation	As-built modeling	Rapid modeling
Example(s) of product	Next Generation LADAR (NIST)	SwissRanger 2 (SR-2)	Leica, Riegl, Sick, etc.	Leica Distometer, or (Total Station)				
Sensor Objective	Shape, Dynamic	Shape, Dyn.	Shape, Static	Dimension, Static				
Illumination source	1500nm (Eye-safe)	890 nm	232-904 nm	690 (670) nm	•	•	•	•
Field of view H+V	360°x360°	43°x46°	360°x310°	Sparse Point	•	•	•	
Range [m]	300	7.5	300	200 (300)	•	•	•	•
Range resolution	1-3mm@15-100m	4.6cm@7.5m	4mm@50m	1.5mm@100m	•	•	•	•
Angular resolution	<0.03°	0.28-0.35°	0.001°-0.658°	N/A	•	•	•	•
Update frame rate	>10 Hz	Up to 30 Hz	1-0.0111 Hz	< 0.00556 Hz	•	•		•
Size [cm]	"Coffee Cup"	4x13x3	38x34x43	17x7x5 (35x23x21)	•	•		
Weight [kg]	0.3	0.3	15	1.5 (7)				
Cost [1000 USD]	<1	5	20-1,000	0.6 (8)	•	•		•
System Operability	24/7	24/7	N/A	N/A	•	•		

Table 2 summarizes the characteristics of the sparse point cloud, dense point cloud, and 3D range cameras (Teizer, 2005a). Ideal values for the target applications in transportation and construction are also indicated in the table described as follows. The field of view of the sensor must not be too limited, but if limited, the sensor should come with a pan and tilt unit to reach all areas of the desired field of view. Range is defined by the application and must be able to deal with fast moving objects like cars on highways. Sufficiently high resolution and image quality provide reliable and robust obstacle/object detection and terrain feature information. The accuracy needs to be at the level of millimeter to precisely place and navigate objects, such as steal beams or vehicles (Lytle et al., 2003). As-built drawings need to be accurate to be further used in 3D or 4D CAD

المنارة للاستشارات
www.manaraa.com

applications. An update rate of 10 Hz or greater is desired to give the operators the chance to view the scene at any moment. The sensor and data processing equipment need to be small in size and low in weight. Both must be able to be mounted on autonomous or semi-autonomous vehicles operable in harsh environments which usually are common on construction sites. Sensor and all necessary equipment must be affordable as cost is a key issue.

2.6.2 Real-time 3D Occupancy Grid for Range Data Processing

The summary of the technology has preferred a 3D range camera that provides range measurements of an entire sensor field-of-view (a three-dimensional space) at frequent updates. 3D occupancy grid with above described parameters would be very suitable to solve the research needs processing acquired range data and for real-time 3D modeling at high range frame update rates of entire sensor field-of-views. In conclusion, occupancy grids, in a variation or extension to the existing occupancy grid techniques, were chosen to be the currently preferred method that would allow detecting objects and furthermore allow tracking objects over time of range frame measurements.

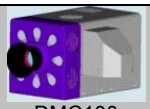
المنارة للاستشارات
www.manaraa.com

Chapter 3: Preliminary Study of 3D Range Sensing Technology

The next sections describe three-dimensional range sensing technology and its working principles. Typical characteristic features like advantages and limitations that are common for all optical range sensors are explained by means of the sensor used in this research, the "SwissRanger" model "2a" and "2b revised". Preliminary experimental findings demonstrate the usefulness of range sensors for different applications. A summary concludes what essential needs are necessary to advance the range camera in the field of static and moving obstacle detection and tracking.

3.1 THREE-DIMENSIONAL RANGE CAMERAS

Table 3: Commercial range cameras manufacturers and products

Commercial Products of 3D Range Camera (selection as of May 2006)	 DMC100	Development kit	SR3000	Prototype	A2
Manufacturer	3DVSystems Inc.	Canesta Inc.	CSEM AG	Matsushita Electronic Works	PMDtec GmbH
Wavelength [850nm]	N/A	785	850	N/A	870
Modulation Frequency [MHz]	N/A	52	20	N/A	4
Illumination Power [W]	N/A	<1	<1	N/A	20
Pixel Resolution (Horizontal/Vertical)	510x492	64x64	176x144	128x123	64x16
Distance Accuracy [cm]	>0.3	>2	4.6cm@7.5m	N/A	10@40m
Maximal Frame Update Rate [Hz]	60	30	50	15	60
Field of View (Horizontal/Vertical) [°]	45x35	30, 55, or 80	48X40	60x45	52x18
Non-ambiguous Range [m]	0.3-2.5	5	7.5	7.5	40
Power Supply [W]	50	10	12	N/A	10.8
Size [cm]	15x5x24	13x6x7	5x7x5	10x10x10	19x13x4
Weight [kg]	1.5	N/A	0.162	N/A	N/A
Connection	Firewire, RS232	USB 1.1	USB 2.0	USB2.0	Ethernet
Output	X,Y,Z, intensity	X,Y,Z, intensity	X,Y,Z, intensity	N/A	X,Y,Z
Website	3dvsystems.com	canesta.com	swissranger.ch	nais-e.com	pmdtec.de

An emerging range sensing technology, called 3D range camera or Flash LADAR, is based on the time-of-flight (TOF) measurement principle using light that

المنارة للاستشارات
www.manaraa.com

returns from an impinged object in a scene back on a sensor. Some manufacturers of these devices and their characteristics are listed in the Table 3. Stone et al. (2004) outlined more range sensors and research prototypes.

3.2 Working Principle of 3D Range Cameras

The range acquisition principle of 3D range cameras is based on Active Sensor Pixels (ASP) that deliver range data of an entire scene FOV in one frame (Oggier et al., 2003). The sensor takes advantage of the TOF-principle and can therefore be used to get depth information as well as intensity (grayscale) information. Infrared light of 800 to 900nm wavelength is used for active scene illumination.

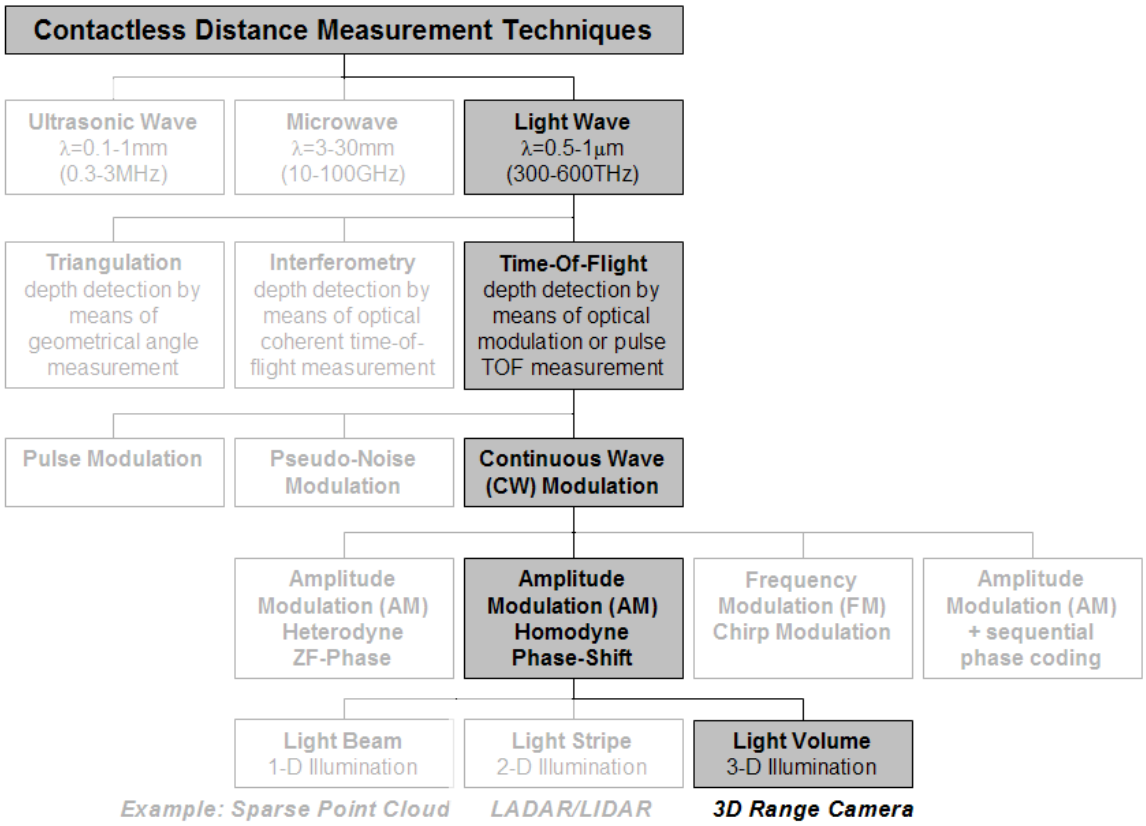


Figure 13: Working principle of video 3D range camera technology

المنارة للاستشارات
www.manaraa.com

This research used a prototype video 3D range camera called the SwissRanger 2 (Model B) that was manufactured by CSEM AG. This 3D range sensing technology is in the group of contactless distance measurement devices that are based on the time-of-flight (TOF) principle using phase shift measurement (see Figure 13).

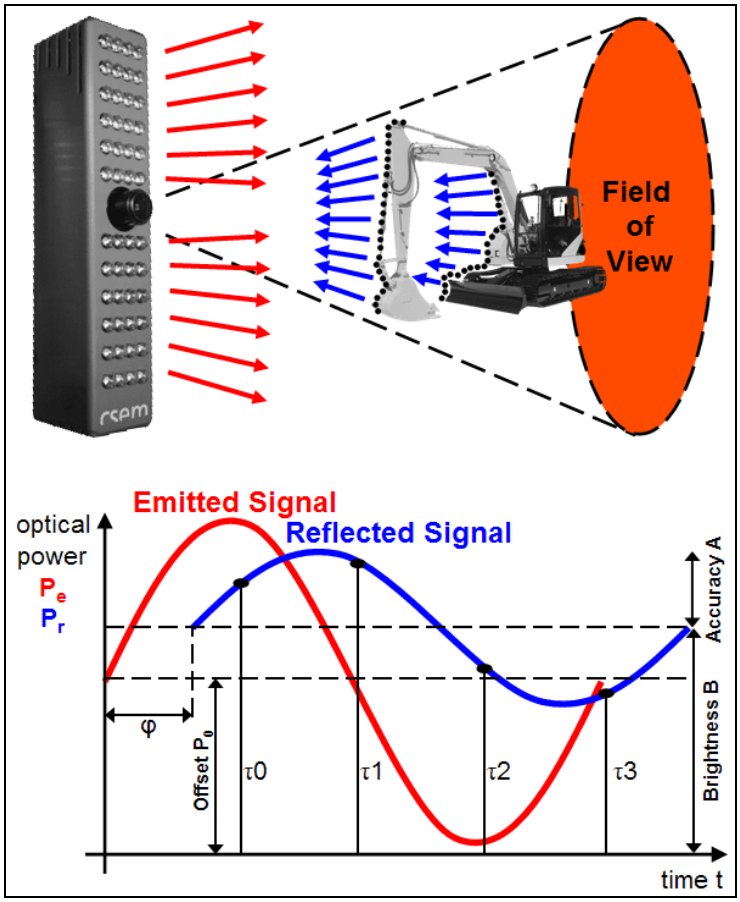


Figure 14: Video range sensing working principle

The sensor illuminates a scene with a synchronically controlled intensity modulated light wave emitted by a near-infrared (IR) light source using 48 conventional light emitting diodes (LED's). Figure 14 and Figure 15 illustrate the working principle:

المنارة للاستشارات
www.manaraa.com

Once the wave front is reflected by objects in the scene it is focused through a lens on a CMOS/CCD sensor chip.

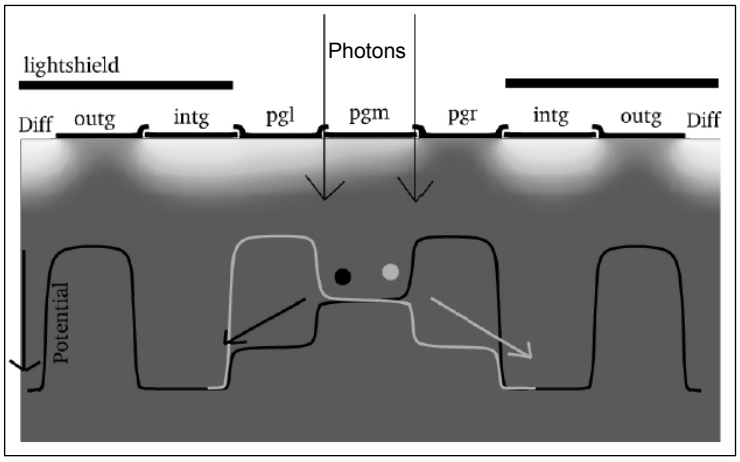


Figure 15: Simulated cross section and potential distribution of a 2-tap sensor (Büttgen et al., 2005).

The working principle of the 3D sensor used in this research operates on so-called lock-in pixels. A combined CCD/CMOS process is capable of demodulating intensity-modulated optical signals in parallel. Each pixel on the chip demodulates the incoming wave front and samples four discrete times within a period c(τi) (i=0,1,2,3), while each sample is delayed by a $\pi/4$ phase shift φ. Repeating this process for several thousand pixels on one chip makes real-time imaging possible. Lock-in Pixel for range imaging use the principle of time-of-flight. Both the detection and the complete demodulation are performed in the charge-domain using charge-coupled devices (CCD). That ensures almost noise free demodulation of the incoming light signal. Figure 15 shows the pixel architecture of overlapping poly-silicon gates that allows photon to enter the sensor. Büttgen et al. (2005) describe the working principle of the sensor as "complementary-metal-oxide-semiconductor (CMOS) technology implemented within the pixel (not in the figure) that diverts the photons to bins that count the photo-generated charge". Further

المنارة للاستشارات
www.manaraa.com

explanations in this section will briefly concentrate on how the distance information is calculated, but more information to the sensor and its working principle can be found in the literature or obtained from manufacturers (Büttgen, 2005).

The distance measurement is based on the phase shift. The phase shift is a result of the signal round-trip from the measurement system to the object in the scene and back. It is exploited for the extraction of the distance information by sampling. The received signal is off-set shifted by a mean optical power mainly due to additional background light and non-perfect demodulation. The offset component further contains the intensity information. Taking four samples from the returned wave front provides the phase φ

$$\varphi = \arctan\left(\frac{c(\tau_3) - c(\tau_1)}{c(\tau_0) - c(\tau_2)}\right)$$

as well as the amplitude and the intensity. The amplitude of the detected optical signal component is reduced by a factor k, depending on all optical losses. Its height defines the signal-to-background ratio and the achievable accuracy of distance measurements. D represents a direct measure of the distance of the captured target

$$D = \frac{\lambda}{2} \cdot \frac{\varphi}{2\pi} \quad .$$

The result is a complete 3D-map of the acquired scene. The LEDs emit IR light at a peak wavelength of λ_c=880nm. This carrier wave is modulated in amplitude with a frequency f_{mod} of

$$f_{\text{mod}} = \frac{1}{T} = 20MHz \quad ,$$

ultimately limiting the unambiguous distance *D* to

المنارة للاستشارات
www.manaraa.com

$$D \leq \frac{\lambda_{\text{mod}}}{2} = \frac{c}{f_{\text{mod}} \cdot 2} = \frac{3 \cdot 10^8 \frac{m}{s}}{2 \cdot 20 \cdot 10^6 Hz} = 7.5m \qquad .$$

The restriction to the non-ambiguous or maximum range to half the period of the modulated light is not the ultimate limit. Applying a second source with different modulation frequency, period, Λ_1 and Λ_2, generates a so-called synthetic wavelength, that results in

$$\Lambda_{12} = \frac{\Lambda_1 - \Lambda_2}{|\Lambda_1 - \Lambda_2|} \qquad .$$

Any a single modulation frequency will result in a limited maximum distance since the raw range data does not carry information about the ambiguity. One practical way to limit unambiguous data is to place a background wall before the maximum distance. A second way is to apply two modulation frequencies at the same time as a more promising approach in the future. In general, lowering the modulation frequency allows reaching multiple distances, but reduces the reflection of the emitted light from impinged objects to the extent that they may often not be detected with their correct distance measurement, but with

$$D_{non-ambiguous} = 0 < D - n \cdot \lambda_{\text{mod}/2} \leq \lambda_{\text{mod}/2} \qquad , \qquad n \in N$$

thus the 0m to 7.5 m current limits. The wavelength itself can be significantly extended in all optical range imaging techniques. Furthermore, as previously stated, the number of accumulated photons or photo-charges is directly responsible for precision of range measurements. Given the state-of the art of photo sensing, analog electrons and

المنارة للاستشارات
www.manaraa.com

analog-to digital conversion, a typical range is 1:1,000, thus, allowing determining the phase better than one thousandth of a period (CSEM, 2005).

The video 3D range camera (SwissRanger 2) acquires of range and intensity information for each of the 160x124 pixels in one frame organized as a grid. The brightness information is the average of all four amplitude samples. The resolution (frame) refresh rate is up to 30 Hz. The basic software acquisition interface coming with the camera allows different camera parameter settings in the registry (see camera interface description in Appendix A). All experiments used the standard camera settings achieving a frame acquisition frequency of 15.2 Hz and a field of view of (experienced) 41.7° horizontal and 44.6° vertical. A pixel at 7.5 m in the distance map represents a volume pixel (space volume also called voxel) of about 4.6 cm in each axis. Since this emerging technology emits near infrared light at a 880nm wavelength, the approach is eye-safe and is particularly useful for night operations where static and dynamic objects can hardly be recognized.

3.3 PRELIMINARY EXPERIMENTS AND FINDINGS

To understand the emerging field of prototype 3D range cameras better, this section focuses on laboratory experiments and field trials under realistic conditions for the development of obstacle detection and tracking algorithms for construction machinery using a high frame-rate 3D range sensor (SwissRanger 2a and 2b revised). The focus of the preliminary experiments was to find the benefits and limitations of the range camera. Once these were known, it allowed determining the characteristics that were needed to develop the range data processing algorithm. In a final step these findings were used to determine and prepare a physical indoor and outdoor experimental environment that housed the final experiments to validate the developed algorithm.

المنارة للاستشارات
www.manaraa.com

Tests were conducted to analyze and discuss the sensor performance and evaluate the trade-off between scene detail and processing speed. Preliminary experiments and tests with the sensor mainly targeted the current hardware limitations of the prototype sensor and the influence of an ambient environment. Various researchers and groups such as Dietiker (2003), Lytle et al. (2005), Kahlmann and Ingensand (2005), Teizer et al. (2005b), and Weingarten et al. (2004) have prepared and conducted preliminary experiments that characterize the minimal and maximum sensor performances (of a SwissRanger 2a and b revised). To know most of these boundaries was necessary to create an experimental setup that can still solve the research objectives. The preliminary findings are:

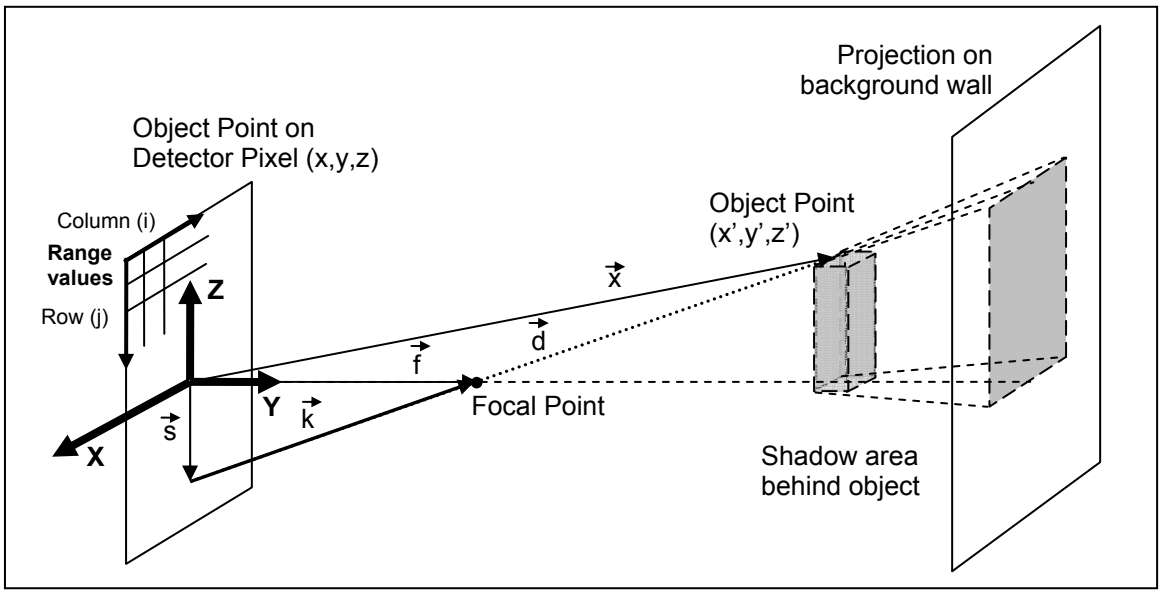


Figure 16: Object, shadow area, and projection of range points on active sensor pixels

- Range data allows creating a *two-and-a-half dimensional model* of the object in a 3D world. The backside cannot be measured unless the sensor moves to multiple

المنارة للاستشارات
www.manaraa.com

locations and allows computational and time intensive fusion or merging of sensor data (Teizer et al., 2005b).

- *Line-of-sight* is required which is typical for all laser scanning devices. As seen in Figure 16, the range camera typically performs a "perspective" projection. A 3D point is projected along a (theoretically) straight line toward the origin until it intersects the focal plane. Photons from multiple 3D points can be mapped to a single pixel on the 2D plane of the sensor. The sensor typically captures range points from the face of objects which is directed towards the sensor. The shadow area behind the LOS will not return any range values. Multiple sensor arrangements may solve this issue.
- Range data is limited by the *sensor's field of view (FOV)* and may not allow to model object in its entire volume or dimension. The FOV of one sensor is too small for entire 360 degree obstacle detection and makes obstacle detection suboptimal.
- *Dead sensor pixels* in the sensor manufacturing process occur. On the used range cameras a total of five pixels were not returning any range measurement. A processing algorithm needs to filter dead range measurements (Teizer et al., 2005b).
- The *design for an experimental setup* needs to reduce jitter. Since the camera has a small size and weight it can be very vulnerable to earth vibration or wind velocities. A fix sensor mount can significantly reduce undesired sensor motion during image product formation that degrades the range image quality (e.g. during camera exposure time, from frame-to-frame).
- Certain *object surfaces, edges and corners* may not return true distances as the emitted light reflects on surfaces and can backscatter from multiple points (see

المنارة للاستشارات
www.manaraa.com

Figure 5). Light beams can be diverted, extended, or reflected from object borders and extend the length of the travel of photons back to the lens and detector. Because of the angle of incidence the detector can receive no light return or smaller amplitudes returns from the emitted signal.

- *Dark colored objects*, e.g. black tires, function as a "black hole" and do not reflect the emitted illumination power well. On the opposite side, some objects can function as reflective mirrors.
- *Several light sources* in the near-infrared area exist in ambient environment that can influence the dynamic area of the detector, i.e. permanently overdrive the dynamic area causing saturation of pixels, e.g. bright sunlight. Scene points that receive small power from the sensor light source, may return a stronger than intended signal, thus giving an error measurement. Computational filters can reduce this problem (Dietiker, 2003). The CCD uses an interference filter to at least filter out all visible light. Infrared and ultraviolet light, as well as light forms that are close to the emitted photon source wavelength can influence range measurements. In general all light that chooses a direct path once it impinged the target objects back to the camera contribute to a good measurement, however, all light sources that choose an indirect path, will increase the error. Light sources can increase the error, when they are additive to the original direct signal, e.g. background sunlight or incandescent light bulbs.
- At an increasing *angle of incidence* a longer distance measurement was observed by Kahlman and Ingensand (2005). A radial increasing angle of incidence appears to be towards the center. This effect is partially influenced by the illuminations system (unsymmetrical mount of LEDs on camera or not enough available potential to illuminate the entire scene properly) and the used optics (lens). In an

المنارة للاستشارات
www.manaraa.com

averaged range image of 100 frames, seen in Figure 17, however, the standard deviation improves towards the center to less than 1cm. These pixels that lie in the corner or edges of the range image have a higher standard deviation and were extracted from the range data processing.

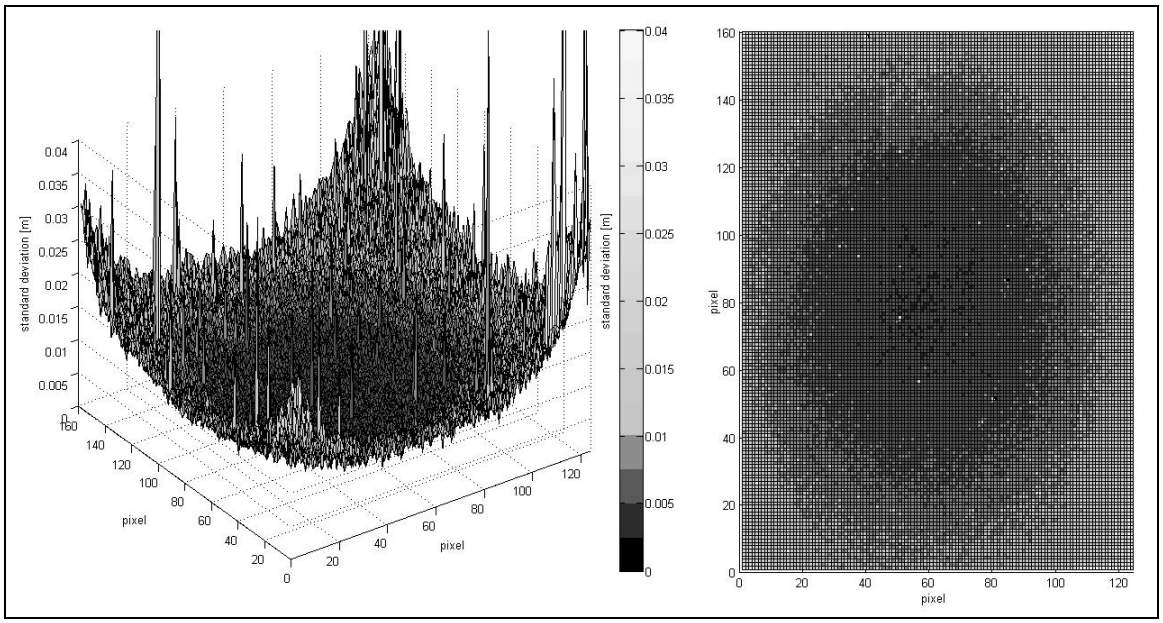


Figure 17: Standard deviation of 100 measurements on a range image shows diffractions of up to 3 cm (Gut, 2004)

- The linearity of TOF measurements is based on the constant speed of light. In experiments the linearity of distance measurements was appropriately correct, (Dietiker, 2003, Kahlmann and Ingensand 2005). One of the biggest problems is to use a light source emitting a synchronous and continuous wave front into the scene. The optical power decreases quadratically with the distance, thus further objects may be harder to recognize or the accuracy of the distance measurement may drop.

المنارة للاستشارات
www.manaraa.com

- *Optics*, e.g. the lens, plays a key role, since single light beams can affect several pixels on the CCD if the lens does not work properly, thus causing several wrong distance measurements.
- *Range accuracy* is limited by the illumination level of the available light. Short light pulses at higher modulation frequencies allow transferring enough power to detect the emitted photons at an eye-safe level, however make it difficult to generate light pulses for longer ranges with high accuracy range measurements (Dietiker, 2003). Range measurements can be fast and accurate enough for construction-type environments. A frequency of up to 15.2Hz was experienced during range frame acquisition and a maximum possible range with 7.5m (using one wavelength) with a range deviation at 7.5m of maximum of 4.6cm.
- *Visible effects of errors in areas of high velocity and distance gradients*: Per definition of the 1-tap distance measurements a serial method calculates range values for each pixel. For very fast moving objects it can happen that that only one or two of all four taps are covered with the distance to the object. The last two of the four tap measurements could focus on a background wall further away. This phenomenon can lead to a distance error that is inversely proportional to the frame grabbing frequency.
- *A "ghost measurement"* is shown in Figure 18 in an outdoor environment. The concrete truck appears in several frames after it already left the scene. This phenomenon of not resetting oversaturated pixels values from one frame to the next was observed with the SwissRanger 2a only and is due to a sensor manufacturing defect particularly critical in outdoor environments. By waiting a several few seconds the sensor want back in normal status.

المنارة للاستشارات
www.manaraa.com

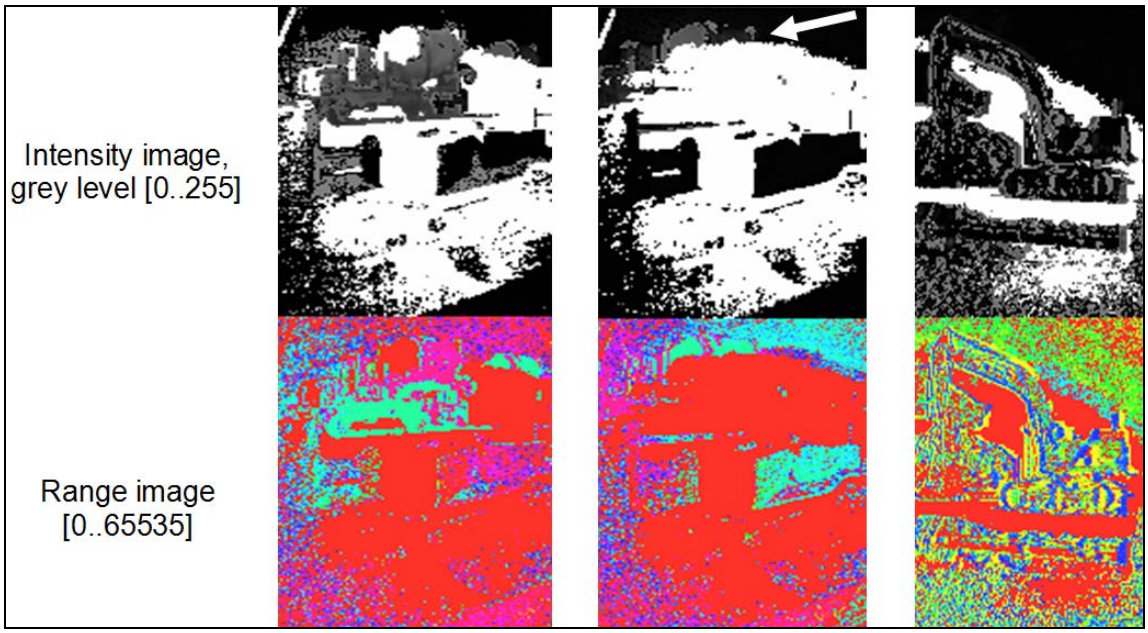


Figure 18: "Ghost image" and noisy measurement in outdoor environments

- *Calibrating 3D range camera* is important. The sensor does emit a non-ideal sine wave into the scene and the received wavefront behaves similar. Range values are based on the arctan-function corresponding to the ideal function and thus, do not perform any calibration of distance values. As a result, the camera has to be calibrated appropriately in order to obtain more accurate distance measurements.

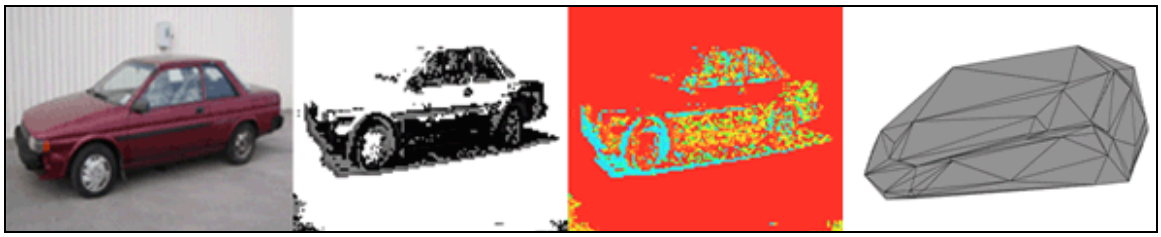


Figure 19: Original, intensity, and range image information, with extracted convex hull of a parked car (Teizer et al., 2005)

- Figure 19 shows the sequence of a preliminary experiment with the SwissRanger 2a version in an outdoor environment: A range and intensity measurement of a car

المنارة للاستشارات
www.manaraa.com

that is exposed to full sunlight. Only the parts of the car that were not exposed to full sunlight, e.g. parts in the shadow, returned the emitted light wave back to the sensor. The entire shape of the car was still modeled since many range values from different locations on the car were returned. An initially developed algorithm built a convex hull. This early on experiment demonstrated that objects in outdoor environment could be successfully recognized and a primitive 3D model could be built (Teizer, 2005b).

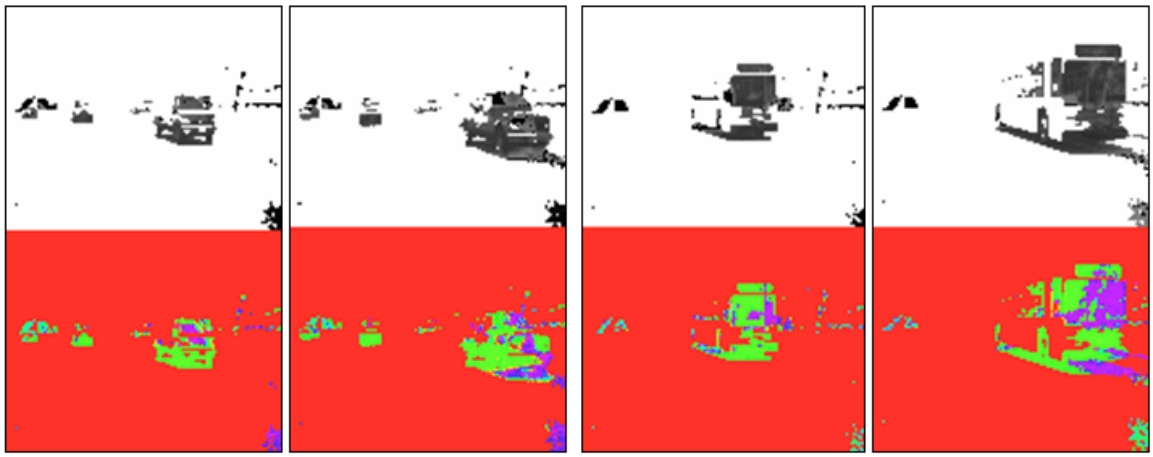


Figure 20: Outdoor tracking of car and bus in bright sunlight

- In a second preliminary experiment, ongoing traffic was monitored. Range points of moving objects, like a car and a bus at time stamp 0 sec. (left image in Figure 20) and time stamp 1sec. (right image Figure 20), were collected. Similarly to the previous experiment with the parked car, only shadow parts returned range information values. However, all moving objects were recognized (not tracked), even at distances greater than 100m (Teizer, 2005b). In addition, the known non-ambiguous distance of the range sensor of 7.5m was experimentally proven. Since no second modulation frequency could be applied to this prototype sensor, the

المنارة للاستشارات
www.manaraa.com

maximum distance of the sensor's FOV was limited to 7.5m to avoid ambiguous range measurements.

- In order to better understand the *photon limitation phenomenon* further tests were conducted by PMD Technologies GmbH and CSEM AG. Their findings are demonstrated in Figure 21 as sunlight entirely overpowers the scene with its wide spectrum light and makes the sensors capability to recognize its own emitted near-infrared light impossible. Even an applied spectral filtering and burst control of the LEDs did not help to recognize significant optical power that is needed to calculate range distances accurately. From the sensor LED came only 0.2 W/m^2 to illuminate the scene, but had to compete with $82W/m^2$ of the background light, e.g. sunlight. The chance of getting useful signal return from the LEDs' emitted near infrared light is minimal and thus the accuracy of range measurements is significantly lowered. Their latest sensor releases have solved this phenomenon by applying specially developed filters.
- *Objects close to the sensor* cover the FOV and cause a similar phenomenon. At object distances less than 1.2m from the sensor, the sensor acts similar to the sun and overpowers the scene with its illumination source, thus, receiving no valuable range return.
- *Internal and External Temperature (*manufacturing of camera): Initial experiments demonstrated that the internal temperature level of the 3D range camera (SwissRanger 2a) caused increasing range measurements over time. This effect stabilizes once the camera runs for several seconds. It is recommended to start and run the sensor for about 10 minutes before taking range measurements (Gut, 2004). In addition, research found, that the higher the surrounding temperature, the longer the range measurements get. This partially depends on the

المنارة للاستشارات
www.manaraa.com

sensors prototype status, its internal temperature control of self induced heating, and as well on outdoor ambient temperature levels. Both factors can influence range measurements of up to 8mm (Kahlman and Ingensand, RIM 2005).

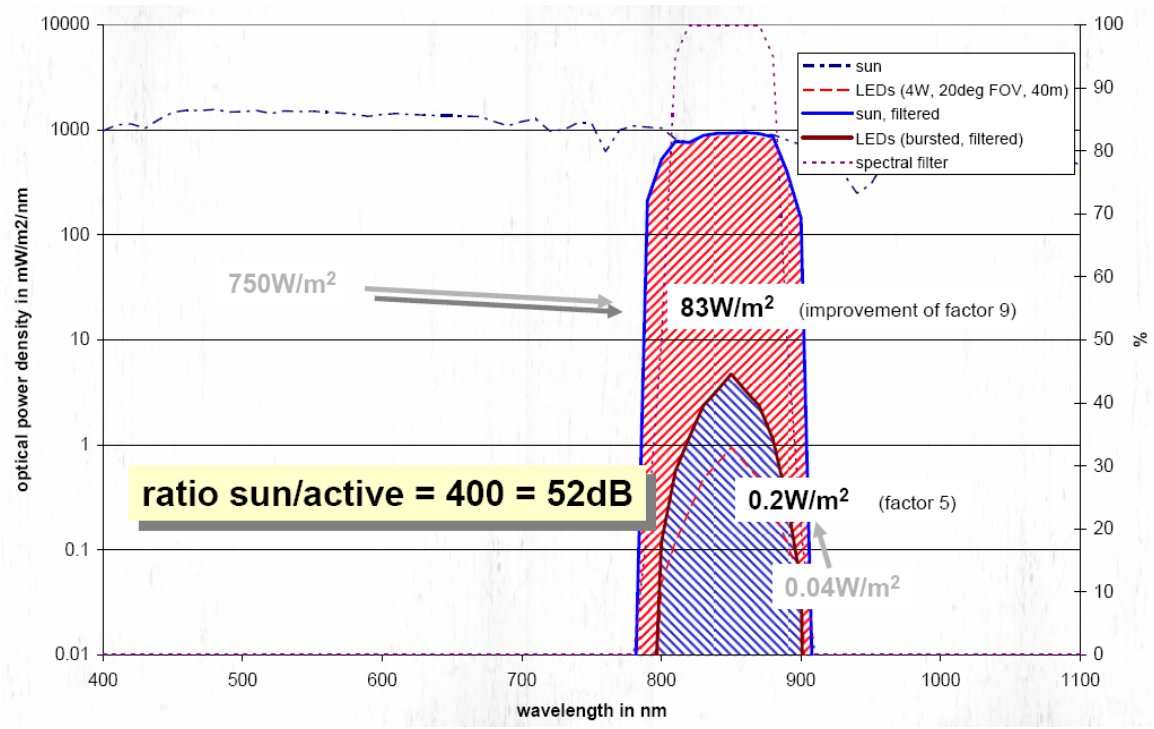


Figure 21: Spectral Filtering with burst of LEDs – Influence of the natural background light on sensor emitted light source (PMD, 2005)

- A *frame grabbing interface* comes with the purchase of the camera Details can be read in Appendix A. Challenges will continue to exist, however, in the availability of robust and fast frame rate image processing algorithms.

Table 4 summarizes the findings of the preliminary experiments as well as from other researchers, e.g. mainly documented by Sanders-Reed (2006).

المنارة للاستشارات
www.manaraa.com

Table 4: Limitations affecting the potential of 3D range cameras

Background and object radiance

- Contrast
- Signal-to-noise ratio (SNR)
- Signal to clutter
- Background light suppression: Reducing the impact from other light sources, e.g. sunlight must not significantly reduce the accuracy of distance measurements. Oggier et al. (2005) and PMDtec (2005) have proposed workable solutions to this problem. A longpass filter may solve this problem.
- Light absorption (objects get dimmer, black objects do not return light back very well)

Ambient environment

- Temperature
- Humidity
- Turbulence
- Spectral attenuation
- Dust
- Atmospheric propagation and phenomenology effects
  - Absorption
  - Emission
  - Turbulence
  - Scattering (when light is scattered by a moving object, the wavelength changes by a small amount. Measuring the "Doppler shifts" in laser light may be require to receive accurate measurements that is backscattered from airborne particles) (object too close)

3D range camera

- Optical and geometric distortion (e.g. lens distortion, typically radial about a specific point, is a problem from mounting the lens on the camera and can be very computationally expensive to fix. Lens distortion is more severe with wide range angles (short focal length) lenses, negligible with telephoto (long focal length) lenses).
- Optical diffraction
- Emitter
  - Signal variation: Is strong over the image scenery
  - Signal source (conventional LEDs vs. laser diodes)
  - Mounting symmetry of LED arrangement
- Field of view (FOV) (limiting horizontal or vertical space)
- Line-of-sight (LOS) and shadow effect
- Camera motion or jittering
- Detector
  - Internal FOV (number of pixels)
  - Focal plane non uniformity
  - Wavelength (operating at one or multiple wavelengths to extend range)
  - Saturation of pixels: Enough optical power emitted into a scene needs to reach short and distant objects. Close objects an receive too large optical power, thus saturating the sensor easily at distances less than 1.4m (Büttgen, RIM 2005).
  - Dark current: Produces a constant offset of the demodulated signal.
  - Noise: Speckle phenomena, Flicker noise, thermal noise, kTC-noise
  - Quantity of electrons (QE) fill factor
  - Distance accuracy: At a modulation frequency of 20MHz the distance accuracy is better than 1cm when more than 12,600 electrons per sample have been accumulated.
- Electronics
  - Noise
  - Temporal drift
  - Variation between readout ports
- Calibration: Several research institutions (NIST, ETH Zürich, etc.) are working on the calibration of 3D laser sensors. This research did not intend to calibrate the range sensor which was used to acquire range information from static and moving objects. It was assumed that this calibration problem exist, its importance is significant to determine the precise and accurate position of objects, however, the research focus relied on developing the algorithms to detect and track static and moving objects from a static or moving sensor. As a result, the sensor is assumed to function as a black box, where the calibration problem is accepted to exist, and the existing sensor and the developed algorithm work simultaneously to produce position, dimension, direction, and velocity values as accurate as possible. As a matter of fact, this research and the developed algorithm are valued by its performance of those features, e.g. comparing the objects position, dimension, direction and velocity from the reality to the model. Calibration efforts with 3D range cameras have been conducted by several researchers, e.g. at the National Institute of Standards and Technology (Lytle et al., ISARC 2005), the ETH Zürich (Kahlmann and Ingesand, 2005), Kubacki and Pfeiffer (2005), and Steitz and Pannekamp (2005).

المنارة للاستشارات
www.manaraa.com

3.4 ADVANTAGES AND LIMITATIONS OF OPTICAL 3D RANGE SENSING

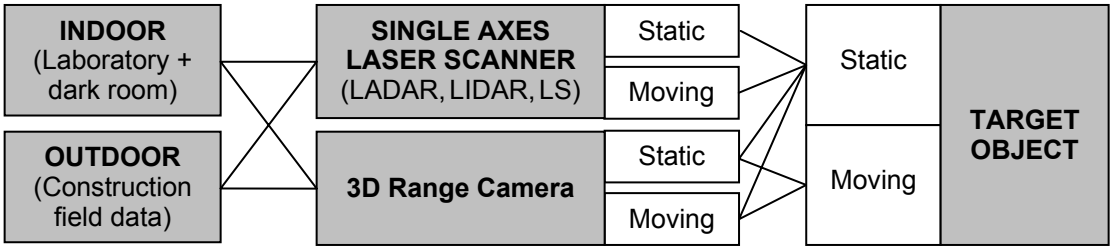


Figure 22: Modeling capabilities of a laser scanner vs. 3D range camera

As a result of comparing range scanning alternatives, Figure 22 demonstrates the biggest advantage of 3D range sensors is the ability to collect range data of a dense point cloud in real-time, of static and moving target objects. Thus, the 3D range point cloud approach is preferred to non-real-time methods like the previously mentioned sparse point cloud approach (using e.g. a Total Station) which uses single point measurements or any laser scanning approach (e.g. LIDAR, LADAR, 3D laser scanners) which need complex raster range scans using a single axes laser beam. The 3D range camera technique makes it feasible to supply real-time three-dimensional information to provide 3D real-time modeling results for obstacle detection and tracking in construction-like environments.

The ease of generating, manipulating, and detecting light is the reason why optical 3D sensing techniques have become the favorite approach in acquiring the 3D shape of our environment quantitatively. Continuously-modulated time-of-flight measurement has lower requirements to the sensing unit, since it (currently) operates on one bandwidth and at one modulation frequency. This allows reducing the manufacturing cost of the sensor. The biggest benefits from 3D range cameras are:

المنارة للاستشارات
www.manaraa.com

- Short range and intensity data acquisition times and at the same time
- High frame rate for immediate range feedback
- Deliver 3D range image and distance maps in one frame
- Capturing static and dynamic scenes and thus not conceivable to laser scanners
- Ease of use at day and night
- Insensitivity to background light (solved by companies PMD and CSEM)
- Handheld like small sized and compact devices
- Competitive prices (currently around $5,000, as of May 2006)

The applicability of optical 3D sensing techniques restricts its use to areas where line-of-sight is the preferred alternative. Light as a carrier wave to collect range data is sensitive to ambient environments. Physical effects deserve detailed consideration because they may also limit the performance of 3D vision methods. The main limitations to 3D range cameras are:

- Missing range data processing algorithms and applications
- Ambient environment influencing measurements (e.g. atmospheric noise) requiring post-filtering
- Optics or physical camera effects (lens or detector) causing inaccuracies in distance measurement performance
- No standardized calibration technique exists for laser range measurement methods
- Non-optimal manufacturing of camera device and parts
- Line of sight produces shadow effects
- Range data is unprocessed and needs (instant) post-processing

المنارة للاستشارات
www.manaraa.com

Each of the optical range sensing methods has its own practical and theoretical difficulties and limitations, but all range imaging approaches are following the same functional relationship, ultimately limited by the quantum noise of the light generation and detection process (as one limiting factor to the performance of optical range imaging systems). Optical diffraction, speckle phenomena, and other physical effects deserve detailed consideration because they may also limit the performance of 3D vision methods.

3.5 SUMMARY

The preliminary experiments allowed determining a list characterizing the advantages and limitations of optical 3D TOF range sensor in general, and in particular of the "SwissRanger 2b revised". To meet the research objectives, this gained theoretical and practical knowledge needs to be taken in consideration in the next steps of the research. Thus, before the development of range data processing algorithms, final experiments, and analyses begins, either these previously discussed limitations are first solved or their existence must be assumed and tolerated to be able to continue. Theoretical and practical limitations of optical time-of-flight range sensing are mostly known through preliminary experimental tests and the literature review. The limitations are documented, some potential solutions were explained, and this research acknowledges the existence of these limitations. The following chapters describe the algorithm development, experimental testing, and evaluation of its results.

المنارة للاستشارات
www.manaraa.com

Chapter 4: Three-Dimensional Occupancy Grid Algorithm

To accomplish the research objectives (real-time resource detection and tracking in construction environments incl. the position, dimension, velocity, and direction of static and moving objects in the field-of-view of a range sensor) and based on the literature review and preliminary experimental findings

- a 3D Range Camera (SwissRanger 2) and
- a Three-Dimensional Occupancy Grid algorithm based on a Single-Linkage Agglomerative Hierarchical Clustering technique

were selected. Both enable to acquire and process range data, extract noise measurements, and successfully cluster the remaining points data into objects, so that real-time 3D models of scenes can be created.

This chapter provides an initial overview of how the reviewed range sensing technologies and modeling approaches can be exploited their advantages to build 3D models for obstacle detection and tracking in construction. This chapter mainly concentrates on the detailed explanation of the developed three-dimensional occupancy grid algorithm.

4.1 INTRODUCTION

To meet the research objective, any kind of representation of the problem in the reality, e.g. detection of workers' unsafe distance to heavy construction machinery, in an experimental research model needs to be practical. The choice of algorithm and experimental setup defines how well the research objective can be solved. Intensity and range based data processing algorithms, including the theory of the sparse point cloud

المنارة للاستشارات
www.manaraa.com

approach and of occupancy grids based on two-dimensional rasterized range maps, have been reviewed in chapter 3. Technology that makes range sensing feasible in the entire three-dimensional field-of-view of a sensor was reviewed in chapter 4. Advantages and limitations of technology and processing techniques were evaluated and helped to find answers for acquiring and processing real-time 3D range information.

The strength and weakness of clustering methods have shown that processing the range information using Hierarchical Agglomerative Clustering has advantages to some of its competitors (K-means and DBSCAN). Its biggest weakness is the potentially expensive computation of range points. Thus, reducing the data first to a smaller set than the entire original range map can help decrease the computational cost to segment range images.

Real-time spatial scanning and modeling is defined in this research approach as when both range data acquisition and processing are successful at frequencies greater than 1Hz or match the update frame rate capacity of the sensor. Fast range data acquisition requires processing these dense range point clouds at high frame update rates to detect and track static and moving targets while the sensor was in a static or moving position. It is assumed that a sufficiently ergonomic developed algorithm could be devised that would make use of this information to warn equipment operators or modify the machine movement in a manner that would improve safety by resulting in collision avoidance.

4.2 REAL-TIME 3D MODELING: OVERVIEW OF APPROACH

The fundamental approach of this work is illustrated in a simplified form in Figure 23. Safety on job sites can be improved by detecting, modeling, and tracking three-dimensional boundaries around hazardous zones, classifying them and separating them from the active construction workspace. In heavy equipment operation, for

المنارة للاستشارات
www.manaraa.com

example, objects or zones that might cause or have proven to offer hazardous potential, such as contact with scaffolding, walls, power poles and lines, and reaching into vehicular and pedestrian traffic lanes can be modeled in 3D using the Sparse Point Cloud approach, as described in Kim et al. (2004). This approach was initiated to allow machines to navigate and operate safely using an automated obstacle avoidance support system (McLaughlin et al., 2004).

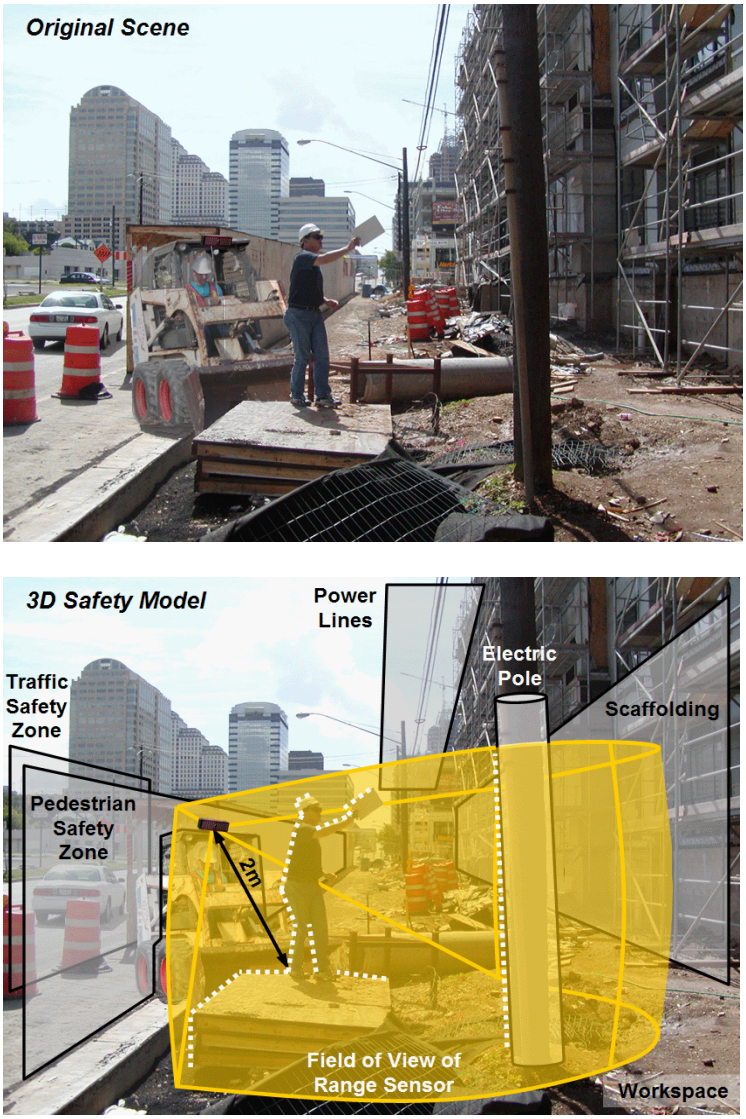


Figure 23: Three-dimensional modeling to detect static and moving object boundaries

المنارة للاستشارات
www.manaraa.com

Based on a few manually selected range points the sparse point cloud modeling approach encapsulates hazardous obstacles in primitive geometric models such as bounding boxes, cylinders, or planar surfaces that separate the job site into safe and unsafe work zones. Figure 10b simulates the modeling results of the spars point cloud approach for permanent spaces (e.g. vehicular and pedestrian traffic, and power poles and lines) and semi-permanent structures (e.g. scaffolding). This spatially efficient demarcation of protected from dangerous workspace, along with real-time obstacle avoidance subsystems embedded in equipment control systems, potentially reduces collisions and consequently injuries and fatalities from accidents.

The sparse point cloud approach, however, does not detect moving resources like Figure 23 demonstrates in machines, workforce, or materials that frequently change their position within the workspace. Thus a new methodology was sought to complement the sparse point clouds approach. The goal was to detect, track and model moving objects for input into an automated obstacle avoidance system. After investigating in the previous chapter the potential sensing and information technologies for acquiring and processing three-dimensional data in real-time, the proposed methodology focused on video rate range imaging (Teizer et al., 2005a).

4.3 CONCEPT OF 3D VIDEO RANGE MODELING ALGORITHM

The research methodology for a 3D video range modeling algorithm included a literature review of range data processing methods that had the potential to solve the needs in the algorithm development. The algorithm development was then split up in four steps: (1) Define the requirements for real-time 3D modeling in construction, (2) design and code algorithms using helpful range image processing techniques, (3) integrate, test and debug in preliminary experiments to measure processing speed, object position, range, detection and tracking validation, and (4) conduct and analyze experiments in

المنارة للاستشارات
www.manaraa.com

indoor and outdoor situations. The developed real-time 3D modeling approach is based on three-dimensional occupancy grids that facilitate modeling and detection of features without having the a priori knowledge of how many objects are within a scene (Moravec and Elfes, 1985). Details of the occupancy grid model are illustrated in Figure 24.

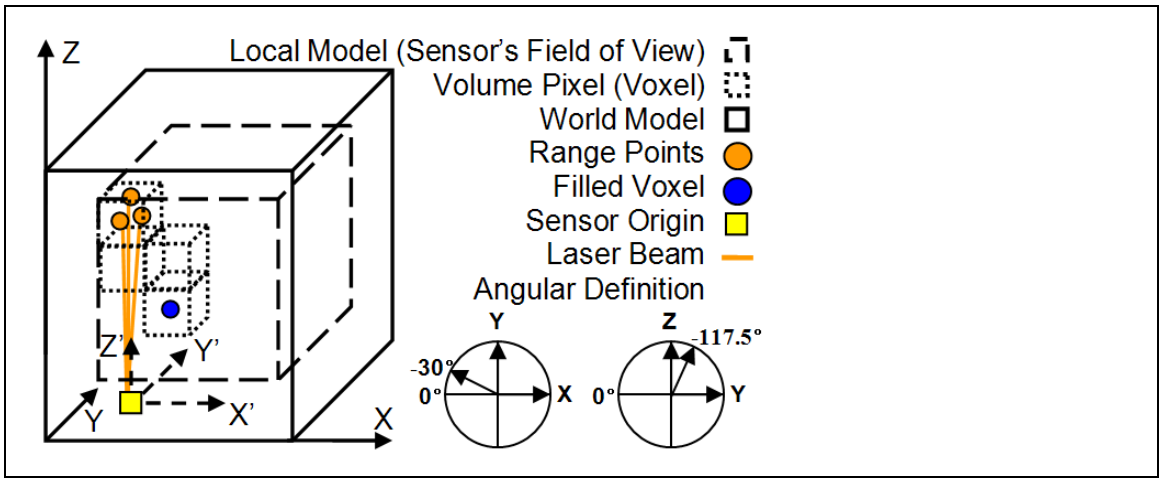


Figure 24: Three-dimensional occupancy grid model

Three-dimensional occupancy grids have cubic units called cells (or volume pixels). Range measurements falling within a cell's coordinates are accumulated, and at some predefined threshold, the cell is considered occupied with mass (value "1") or not (value "empty"). This data reduction method offers an effective way to significantly reduce the original data volume and lower the computational cost. Another advantage of occupancy grids is that the world is modeled within a simple Cartesian coordinate system. The range camera's field of view (called Local Model) is part of a World Model that includes the entire scene. In the following explanations a Local Model refers to a single processed range frame of the sensor's field of view, whereas the World Model encapsulates the processed information of all Local Models. In the World Model not all range data is necessarily needed to detect and track features in a 3D model. Thus, only

المنارة للاستشارات
www.manaraa.com

important information about cluster locations, directions, dimensions, and speed might be stored and be enough to characterize a scene for an application, e.g. related to safety. Figure 25 illustrates the occupancy grid modeling approach. Details will be described in the following sections.

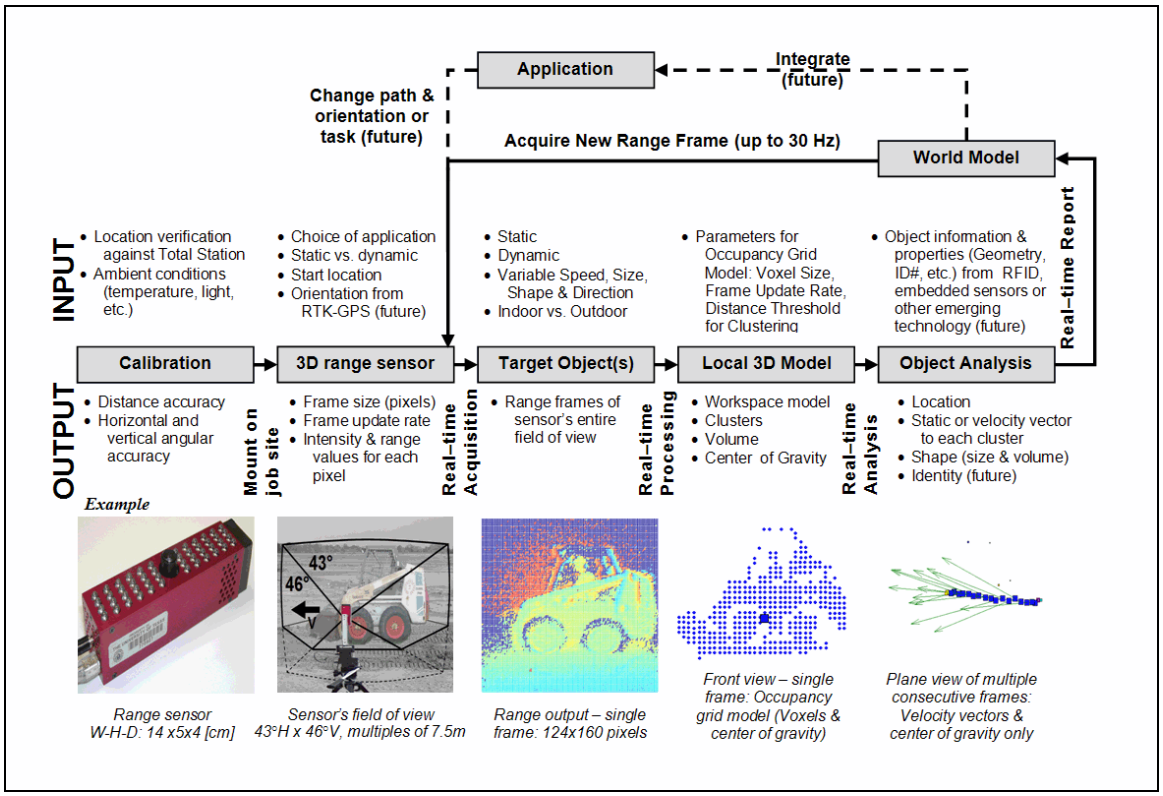


Figure 25: Flowchart of 3D Occupancy Grid Based Model

4.4 REQUIREMENTS FOR THREE-DIMENSIONAL OCCUPANCY GRID MODELING

Some steps in the designed algorithm to process 3D range values in real-time require input parameters which once entered at the beginning of experiments allow the automated generation of outputs, in real-time. These steps include the preparation of the (experimental) sensing environment and preparation of the sensor as well as specific parameters in the data processing algorithm itself:

المنارة للاستشارات
www.manaraa.com

- Calibration: Before range data can be accurately acquired the range sensor needs to be calibrated to the ambient conditions. Kahlman and Ingesand (2005) and Lytle et al. (2005) have developed calibration approaches that were used to determine the distance accuracy as well as the field of view of the 3D video range camera. Distance and horizontal and vertical angular measurements of a Total Station (a laser range finder) to known object locations were compared to the respective location in single frame shots from the video 3D range camera. Differences in measurements were used to calibrate the video 3D range camera. This is an important task in particular for applications related to safety since a well calibrated sensor enables detection of objects at the spot where they are actually located.
- Range Data Acquisition: The camera was mounted on a horizontal gantry or on a tripod and either in static or in moving position at known speeds of up to 0.2m/s. The camera's start point and its orientation of the field of view (FOV) including the location or travel path of target objects were determined before streams of range images at update rates of up to 30Hz were acquired. Static or moving target objects had variable speeds, sizes, shapes, and directions.
- Range Data Processing: Since range data acquisition, processing, and information analysis all work simultaneously, the developed algorithm asked for fixed input values (described in algorithm coding) before experiments (acquisition, processing, modeling, detection and tracking, and analysis) started. This allows some flexibility in adjusting the developed range image processing algorithm to the needs of the particular application, e.g. high accuracy in as-builts vs. speed in machine safety operations. Range frames were processed individually and stored

المنارة للاستشارات
www.manaraa.com

in a Local 3D Model. A Local 3D Model is a temporary bin that contains the processed information and model of one range frame only.

- Range Information Analysis: The objective is to detect and track objects from a stream of range frames. Information within each frame such as cluster identification, cluster dimension and volume, and location of the clusters' volume centroid, speed and direction were stored in a Local Model (individual frame).
- Update World Model: The developed algorithm compares the previous analyzed frame to the current frame result and integrates only important scene information in a World Model (all frames). The information provided in a World Model facilitates tracking objects' location, determining their speed and direction, and can be utilized in applications such as obstacle proximity and avoidance sensing for safety applications. New range frames from identical or different orientations can be acquired.

4.5 ALGORITHM DETAILS

Figure 26 illustrates the details to the data processing. Orientation and position of the sensor is then transformed in the World Model. Since in indoor experiments the ranger sensor was moved on a horizontal gantry, sensor start point, speed, and previously recorded gantry coordinates from a Total Station were sufficiently precise to specify the sensor location in the World Model. Once the experimental setup is established and range data is collected, the first step in the occupancy grid algorithm is to convert the original range data in Cartesian format into spherical format.

المنارة للاستشارات
www.manaraa.com

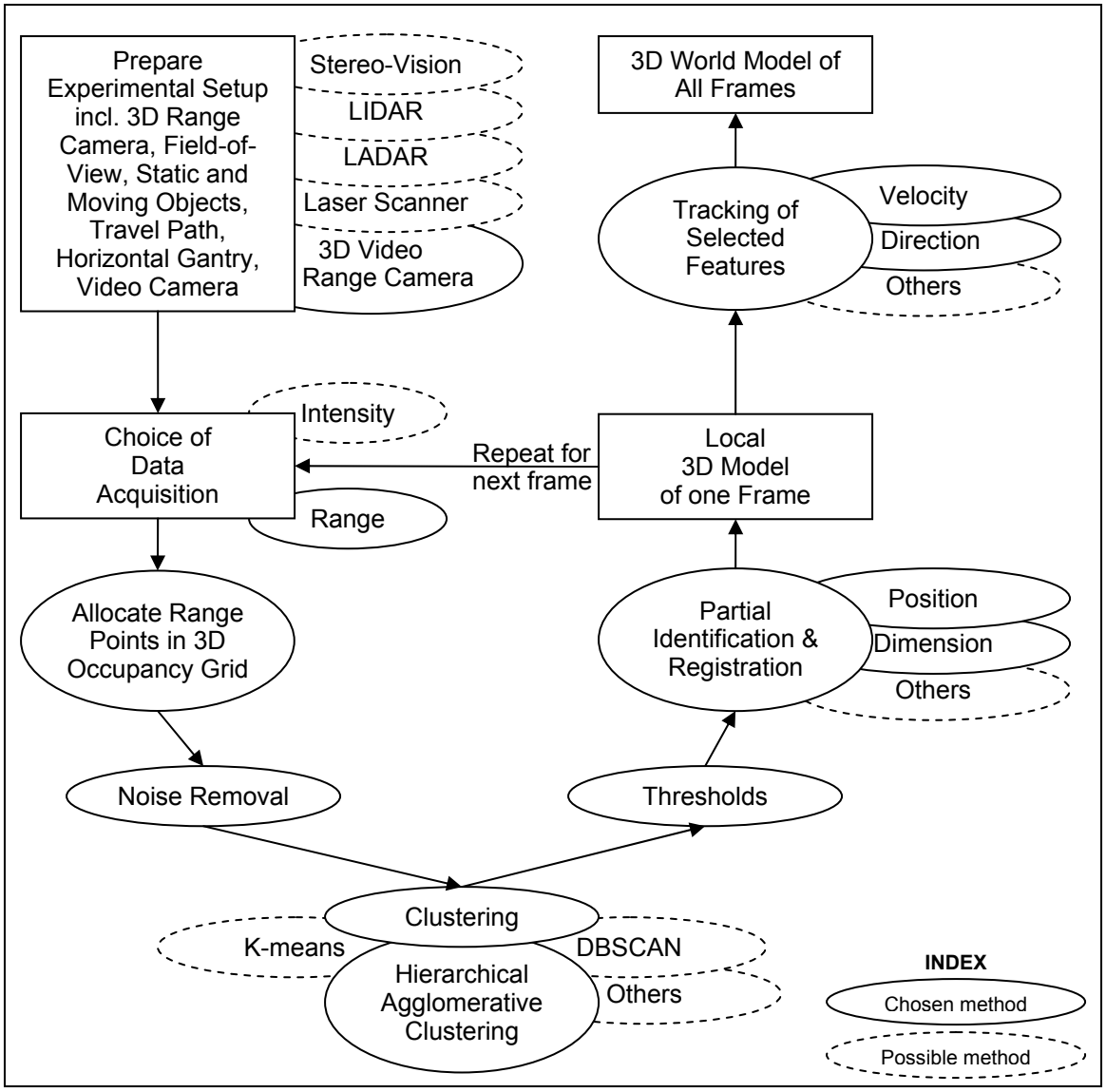


Figure 26: Data processing flowchart

Occupancy grids are based on the principle of allocating range points in a prepared world which is divided in a grid system of variable or fixed voxels (volume pixels). Based on initial empirical tests, the occupancy grid size was fixed to 0.1m in the X, Y, and Z axis. In contrast, larger grid sizes would blur the image and reduce the quality of the model, but can improve the processing speed. In a next step, a voxel then

المنارة للاستشارات
www.manaraa.com

can be filled according to a user-defined threshold value, e.g. if two sensed range points fall into the voxel, it is filled. This procedure by itself could reduce the resolution of the range image and can reduce the number of small sized features contained in a scene. On the other hand, experiments have shown that using a threshold value to fill voxels allows reducing the memory requirements by at least 50%. Experiments have shown also that objects in construction are large enough that the loss in resolution is not significant. As another side benefit, this allocation method also rapidly allowed reducing the initial amount of useful data from random single event measurements, also called "salt and pepper noise".

Before clustering, a second threshold was applied to identify filled cells that were isolated. Only those cells were kept which had at least n filled neighbors out of a surrounding set of m^3-1 cells (e.g. n=6 filled neighbors out of m=3 or 26 surrounding cells). This is a type of median filter. The result was typically a reduction in the complete data set to about a twentieth of the original memory (range and intensity are each 160x124=19,840 points). Clustering techniques with threshold values reduced the data set to roughly one to two thousand remaining range points and allowed achieving rapid results in segmentation and classification.

Each cluster corresponds to a static or moving object in the scene. The algorithm determined clusters' characteristics, such as the dimensions, the position of volume centroid (VC) based on averaging the position of each voxel (X, Y, and Z) belonging to each cluster, as well as the direction and speed in case of a moving object.

Data clustering relates to the segmentation of the 3D data obtained in one 3D image, with the aim of detecting all different features in an environment. The clustering method chosen is not based on a-priori information on the number of clusters, like k-means approaches typically require. Since the number of objects to be detected is

المنارة للاستشارات
www.manaraa.com

unknown, agglomerative hierarchical approaches were preferred because no initial guess is needed about the number of objects in a scene. A single linkage agglomerative hierarchical clustering algorithm was used (Jain and Dubes, 1988, and Jain et al., 1999).

The basic clustering working principle relies on merging clusters if they are close to each other. A threshold value allows the user to define at what maximum distance clusters should be merged or separated. Since the angular accuracy of the sensor was determined 4.6cm (or roughly the voxel size of 0.1m), the cluster separation threshold was set to the minimum size of two cells. This value overall proofed to achieve a good separation of features.

Due to the high noise ratio in the range data acquisition of the prototype sensor, very small objects could appear with only a few filled voxels. Thus, on the remaining clusters another threshold was applied to filter smaller clusters with less than f filled cells (e.g. f=10 cells). As a result of this specific threshold setting, objects of smaller size may not be recognized with large filtering threshold values. Empirical data has shown that split or isolated clusters should be attached to nearby clusters or once computational power increases, threshold values should be removed entirely. For safety applications, for instance, it can be literally dangerous to exclude smaller clusters.

Tracking environment features over time (over a stream of frame images) based on intensity values and background information has been extensively conducted by Sato and Aggarwal et al. (2004). In this algorithm, an empirical approach based on thresholds was used to determine whether two clusters in two consecutive frames are the same. If the size of and the distance between the cluster (cluster volumes) did not change by more than a defined threshold they were considered the same. To define if a cluster is static or moving the algorithm compares two consecutive frames and verifies whether the cluster's VC is within a minimum and maximum distance threshold. According to ongoing

المنارة للاستشارات
www.manaraa.com

experiments, this threshold value was set to a maximum speed limit of 3 voxels per second or respectively the (maximum allowed) speed of heavy equipment of about 10-15 km/h. Of course, at some velocity an object may be moving quickly enough to be mistaken for two different clusters (or objects). This would be a dangerous mistake, since the velocity information would then be lost. Future research will determine at what point this might occur and if it is beyond the capability of a machine to react in any case because of momentum for example.

In summary, a range data processing method was developed which reduces the collected information rapidly and is still able to develop meaningful 3D models containing relevant static or moving features of entire scenes while the sensor is in a static or moving position.

4.6 STEP-BY-STEP EXPLANATION OF ALGORITHM USING AN EXAMPLE

An illustrative example is used to guide through an experiment while explaining the algorithm. Figure 27 shows the experimental setup. In the field-of-view of the 3D range camera are a fascia board, a box with horizontal velocity to the negative x-axis, and a background wall. Since this prototype version of the sensor provided non-ambiguous range information of maximum 7.5m, its FOV needed to be limited to this range. Applying a lower modulation frequency or a secondary illumination source at a different modulation frequency allows to extend the range of the FOV, however, for this research non-ambiguous data is expected. Ambiguous data might be filtered from a scene with probabilistic range data processing approaches (the value of the same voxel flips frequently from one frame to the next).

المنارة للاستشارات
www.manaraa.com

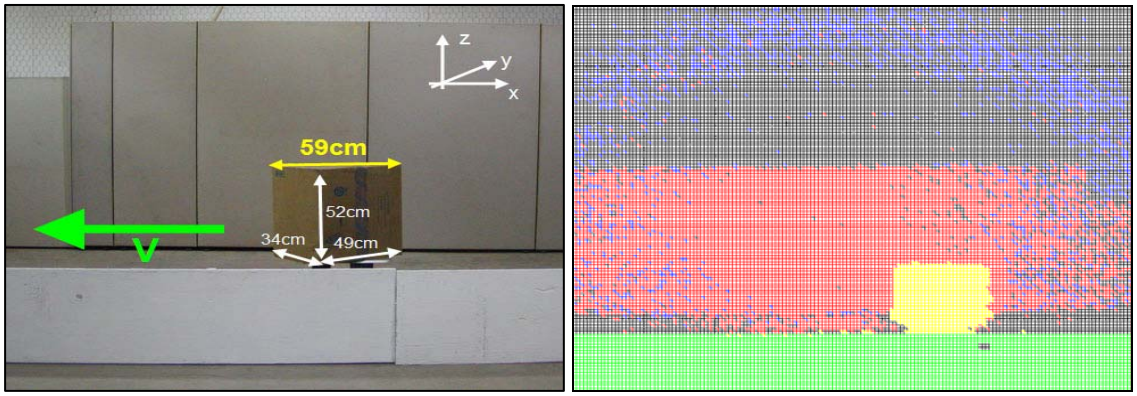


Figure 27: Original color and range image (fascia board, box, wall)

Table 5: User-defined parameters and thresholds

Parameter	Value for experiment
01. Length of side of occupancy grid voxels	0.1m
02. World Model size	14m x 14m x 14m
03. Sensor's initial position (X, Y, Z)	(7,3,2.3)
04. Data reduction value	10 pixels
05. Voxel fill factor threshold	Minimal 2 range data points
06. Filled neighbors threshold	6 voxels
07. Minimal number of filled voxels in one cluster	10 voxels
08. Hierarchical clustering maximum merging distance	2 voxels
09. Tracking decision based on volume difference	25% from one frame to the next
10. Tracking decision based on VC position difference	Minimal 1, maximal 3 voxels

The 3D range camera illuminates the entire scene with a near-infrared wavefront and gets in return the distance map to the entire field-of-view. The amount of data especially when working in 3D easily becomes huge. 3D scans can be composed of several thousand points. In experiments, the 3D range camera "SwissRanger 2a" and "SwissRanger 2b revised" both provided dense point clouds of 19840 or 124x160 pixels per frame per every 15.2Hz.

المنارة للاستشارات
www.manaraa.com

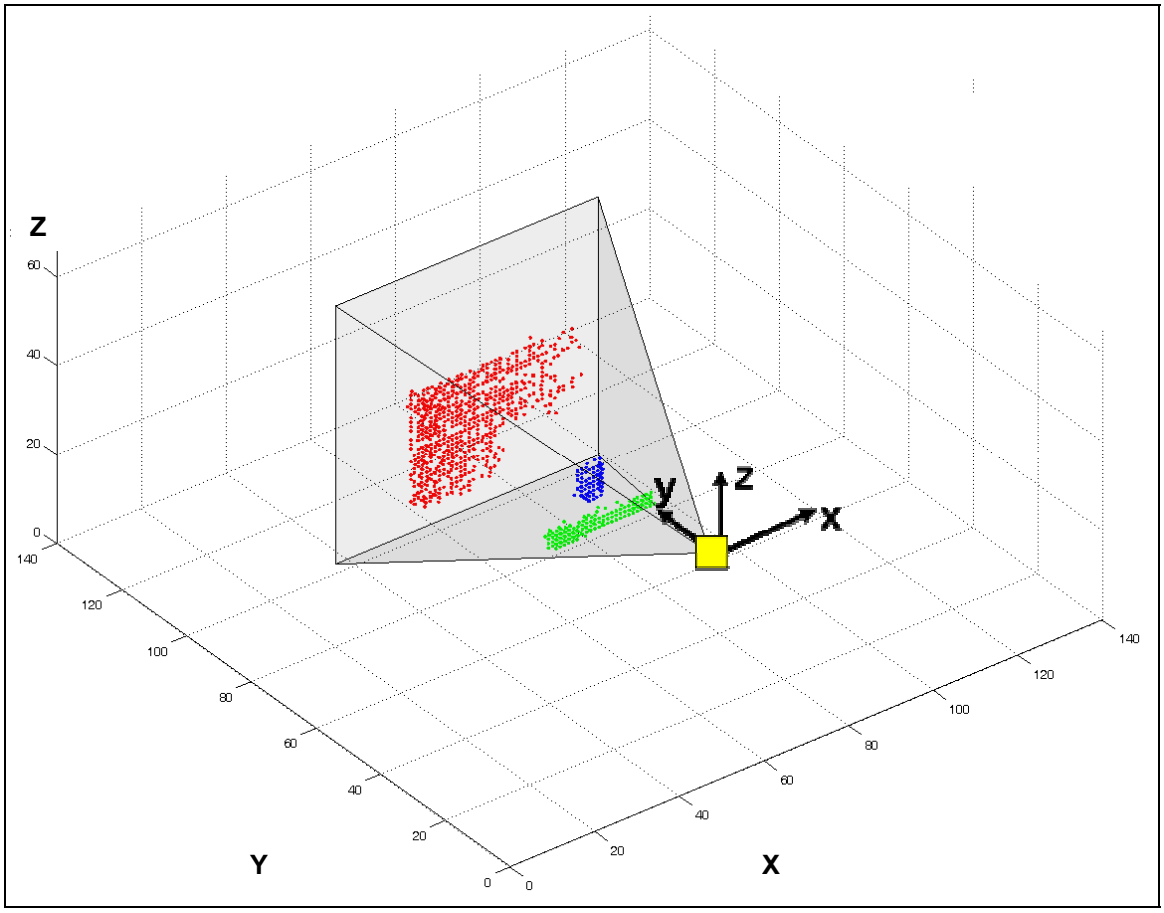


Figure 28: 3D occupancy grid of one frame (sensor's FOV is Local Model, World Model encompasses entire scene)

Table 5 lists several algorithm parameters that define an experimental setup for the purpose of obstacle avoidance in heavy construction equipment operation.

The occupancy grid decomposes the space into regular volume pixels (voxel, cell, or cube) which in this work are chosen to be cubes with a side length of 0.1m (parameter 01). The entire experimental lab hall (physical comparison to the World Model) has the defined dimension of 14m in each axis direction (see parameter 02 and axes grid values in Figure 28). The sensor's start point is defined at (X,Y,Z)=(7,3,2.3) [m] (parameter 03).

المنارة للاستشارات
www.manaraa.com

Preliminary experiments with the sensor demonstrated in Chapter 3 had shown that edges and in particular corners of range images are affected by highly inaccurate and noisy measurement (see Figure 17). These reasons justified to set a region of interest cut 10 pixels on all sides of each frame image (parameter 04). An alternatively method to reduce these noisy measurements is equal to set a region of interest in the camera's FOW. The removed pixels were not considered any longer for the next data processing steps.

To speed up the data processing, as mentioned earlier, the occupancy grid can function as a filter. Only those voxels in the Local Model remain filled and are considered for further processing that had at least two range points contained and in addition each of these voxels needed to have at least six (out of 28 potential) filled neighbors in the immediate surrounding (parameter 05 and 06).

The next step in the algorithm was to process the remaining voxel information so that it was possible to differentiate between objects and extract object information. The voxel cloud was searched for groups that contained at least 10 filled voxels (parameter 07). These voxel groups of 10 or more voxels still required an intelligent clustering algorithm to create separate objects.

المنارة للاستشارات
www.manaraa.com

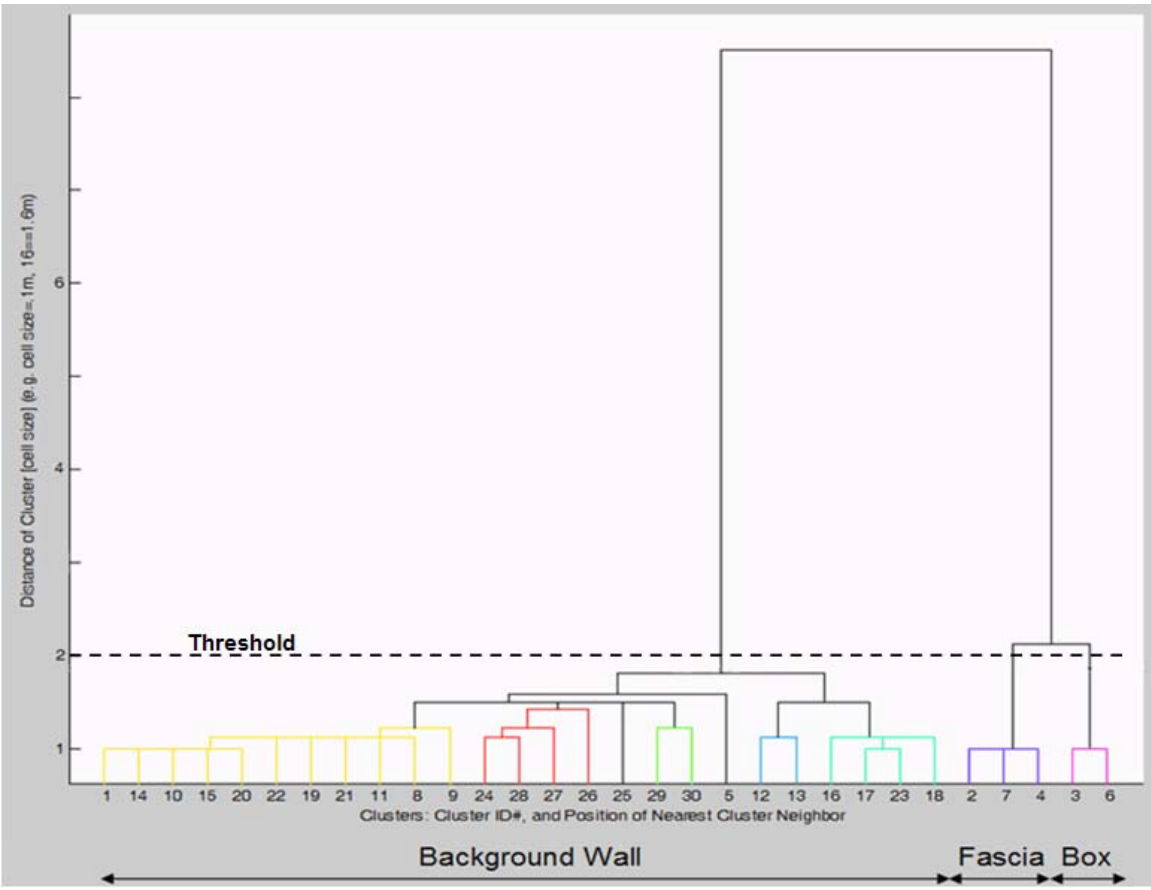


Figure 29: Single-linked agglomerative hierarchical clustering dendrogram

The remaining dense range point cloud, now converted to voxel groups of at least 10 voxels, needed to be processed to individual 3D models using a single-linked hierarchical agglomerative clustering algorithm.

Hierarchical clustering is a set of nested clusters that are organized as a tree. Each node (cluster) in the tree (except for the leaf nodes) is the union of its subclusters, and the root of the tree is the cluster containing all the objects. The two closest clusters (starting as points as individual clusters) are merged when the single-link criterion is met (defined closest distance of two cluster-sets) until all points are clustered.

المنارة للاستشارات
www.manaraa.com

The dendrogram in Figure 29 allows moving a threshold criterion up and downwards. A user-defined threshold value of 2 voxels (threshold value on left axis in Figure 29) groups only those voxel clusters that are closer than 0.2m (=2x0.1m). Otherwise this single distance criterion separates clusters and defines a new object. In Figure 28, for example, three clusters were found: A large object far away from the sensor, a small one in middle distance, and one medium sized object very close to the sensor. Although the algorithm does not provide any classification (naming or identifying objects after certain recognizable features), it is still evident for the human eye that background wall, box, and fascia board were modeled correctly.

The algorithm automatically calculates a couple characteristic values that can be exploited for (partial) object identification. At this stage of the research, these values are only used to determine how well the accuracy of the algorithm worked. The accuracy of the position of the volume centroid to each cluster after the averaging principle and the dimension by counting the furthest apart x-, y-, and z-values within each cluster was measured and then compared to the reality measurement with a Total Station. In Figure 30 this method is illustrated for measuring the dimensions of the box in the 3D model.

The remaining unanswered question is when objects are in static or moving position? An object is defined in “static position” when the volume centroid varies from one frame to the next varies by less than 1 voxel. The factors the algorithm uses to classify a moving object (cluster) are expressed in parameters 09 and 10.

Objects are defined in “moving mode” when the location of the volume centroid changes between one and three voxels from the previous to the current frame. A location change of three voxels per frame has a reason.

المنارة للاستشارات
www.manaraa.com

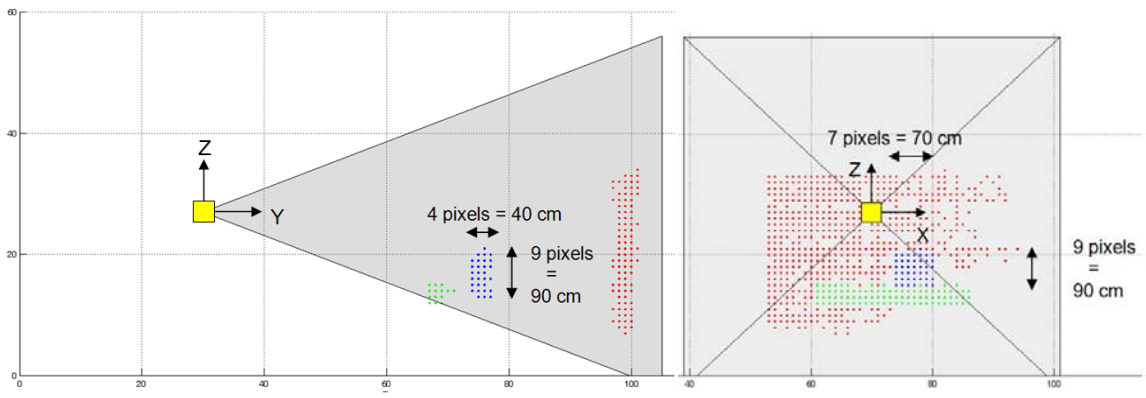


Figure 30: Elevation and front view of clustered 3D occupancy grid model

Signs at construction projects of the University of Texas' main campus indicated a maximum speed limit of 10mph for construction operating on site. Collecting roughly 15 frames per second means an object could travel 45 voxels in one second and still be tracked. Since the size of a voxel is user defined at 0.1m (parameter 01) the object's maximum distance is 4.5m per second which is relative to a speed of 16.2km/h (or approx. 10mph).

A primitive obstacle avoidance system depends on the accuracy of the position and tracks of objects in the FOV of the sensor. Since the sensor generates a noisy dense range point cloud from a 2½D view (not 3D), the accuracy of the position of objects can vary and thus the algorithm can generate 3D models bigger or smaller, or shift the position and direction values of objects.

In this example, the object (box) was propelled perpendicular to the sensor's FOV. Clusters contain the information value of how many voxel they include. For an obstacle avoidance system that allows objects moving in any direction within the sensor's FOV, however, significant voxel count changes can become an unsafe issue. As a result, the change in voxel count needs to be considered in tracking objects using an algorithm based on 3D occupancy grids.

المنارة للاستشارات
www.manaraa.com

Table 6: Tracking decision criteria

Δ in Volume$_{Frame(i)-Frame(i-1)}$	Δ in Position$_{VC(i)-VC(i-1)}$	Tracking decision	
		Object	Classification
>25%	any	Different	New object
≤25%	Δ position >3 cells	Different	New object
	1 cell ≤ Δ position ≤ 3 cells	Same	Dynamic
	Δ position < 1 cell	Same	Static

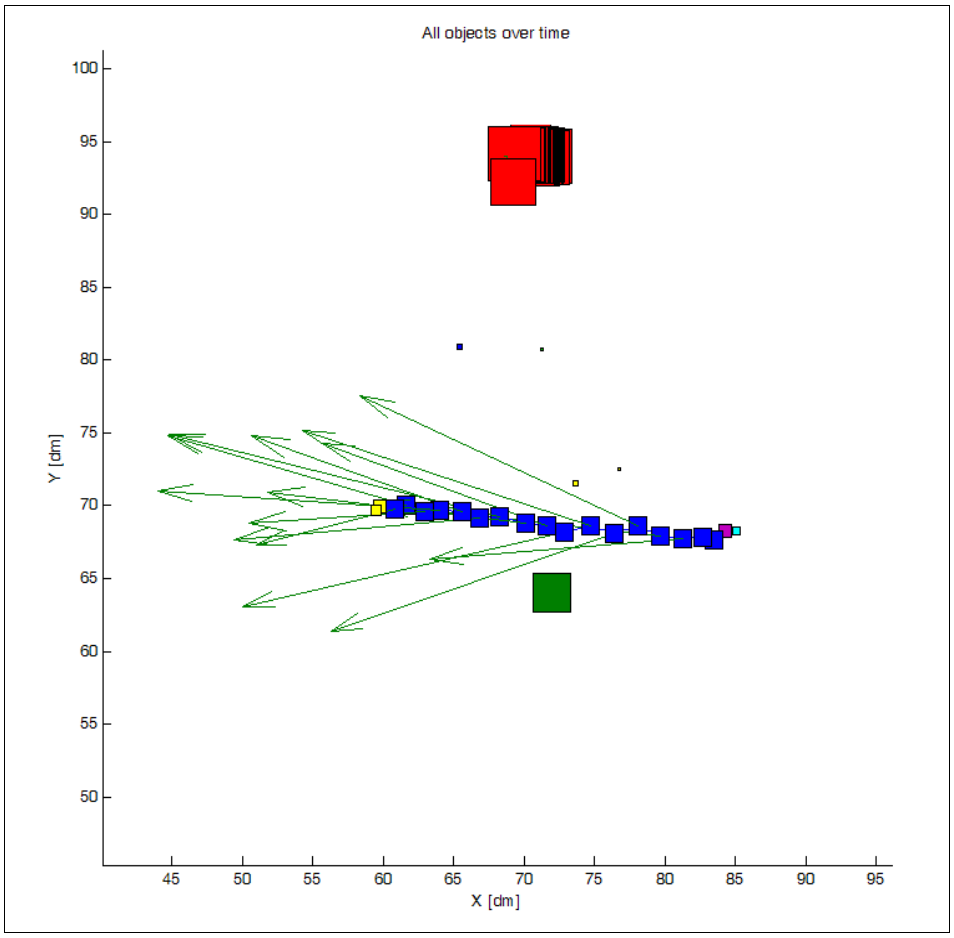


Figure 31: Visualization the track of an object by plotting 52 consecutive Local Model frames in plane view

المنارة للاستشارات
www.manaraa.com

To define boundaries when objects are static or need to be tracked because they move, the user can define specific tracking parameters. The clusters (or objects) that increased their voxel count, also called volume, by more than 25 percent were considered to be new objects (parameter 09). Experiments with fast moving objects towards or from the sensor away could change their voxel count by a lot. At a volume decrease/increase of 25 percent most objects (boxes, pipes) were still recognized. Objects that changed their volume less or equal than 25 percent were treated after Table 6. If a cluster changed its volume less than 25 percent and its location of the volume centroid changed buy more than 3 voxels (from one frame to the next) it was considered a new object. On the other side, a cluster was considered a static object if it moved less than one voxel between two frames. In between these two values a cluster was considered moving.

In Figure 31 the center to each square represents the volume centroid (VC) of one cluster to one frame. The size of each square illustrates the number of voxels in that particular cluster. Since the background wall covers the largest face in the sensor's field-of-view it has the biggest square. The fascia board has more filled voxels than the box and thus has also a bigger square. The VC of the fascia board appears in the projection in Figure 31 only once and can be classified as "object in static mode". Since the box is in the foreground it covers partially the background wall. The object's shadow causes temporarily disconnected voxels within the cluster of the background wall. As a result, the VC of the background wall slightly moves. To calculate the VC differently can be a solution this problem, e.g. using the extreme object points (edges or corners) to calculate the VC.

The projection of 52 consecutive frames in the plane view demonstrates the track of the box over time. Each VC gives the position of the box from one frame to the next. This and the knowledge of the duration between the frames (the invert of frame

المنارة للاستشارات
www.manaraa.com

acquisition frequency of 15.2Hz) allow calculating the velocity vectors and direction. The direction of velocity vectors varies since the curve is an artifact of the exposed surface influence on the VC.

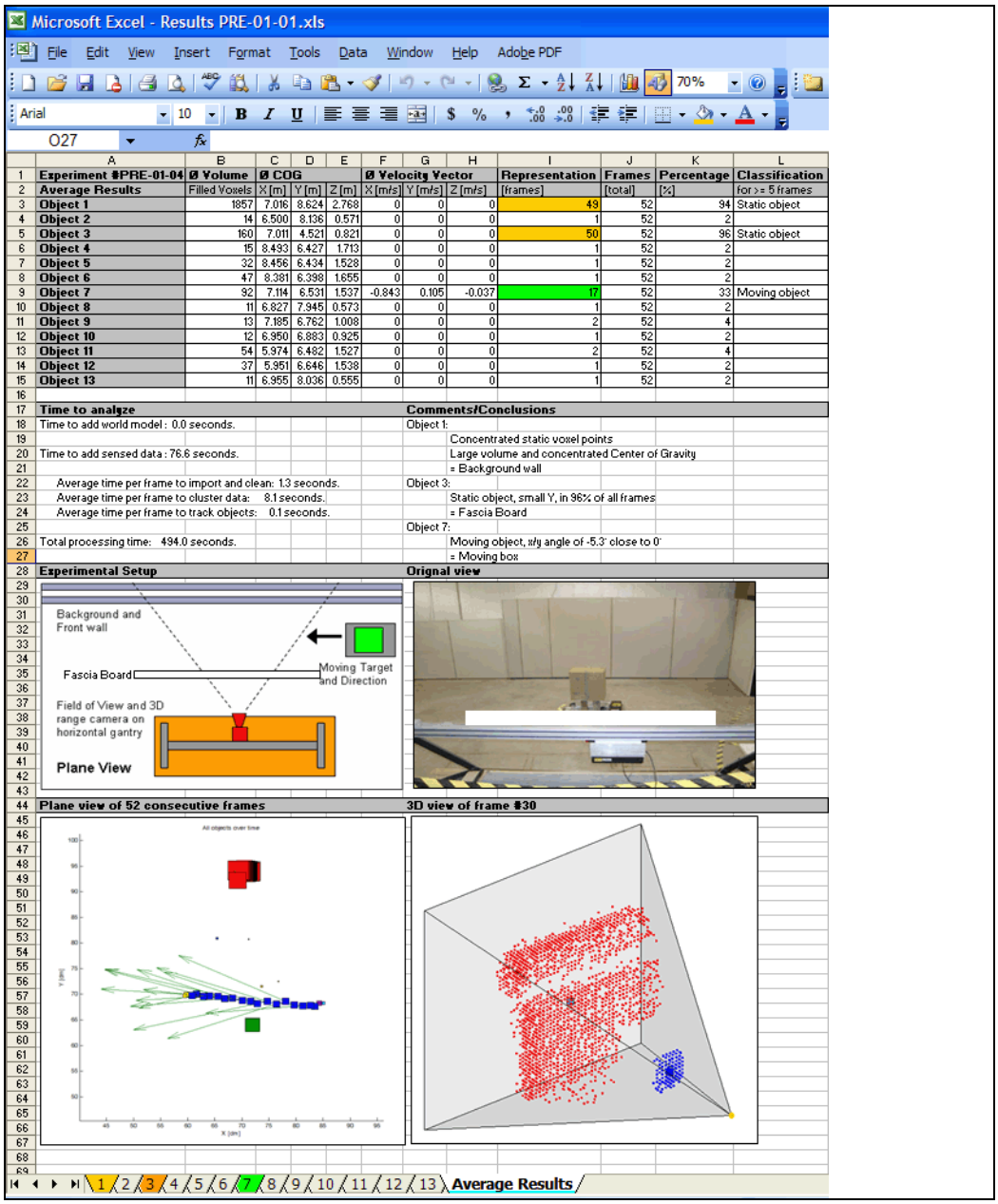


Figure 32: Typical view on summary results of one experiment

المنارة للاستشارات
www.manaraa.com

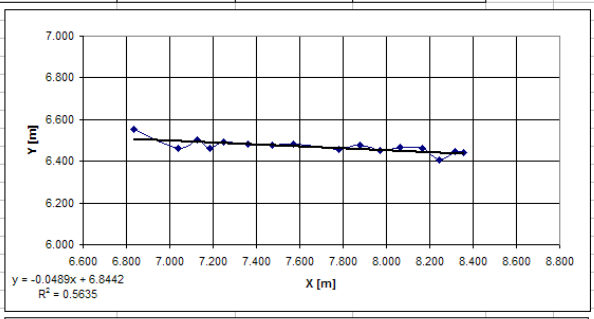
As seen in Figure 32, the first and the last two squares in the trajectory of the box are standing alone and are separated from the cluster, since it the object was only partially within the sensor's FOV (note: color difference of the squares). Obviously the box entered (and exited) the FOV and thus significantly changed the detectable size more than 25 percent than the previous frame at entry (and at exit). Changing the voxel count belonging to one cluster by more than 25% and comparing it to the same cluster in the next frame separated both as independent cluster.

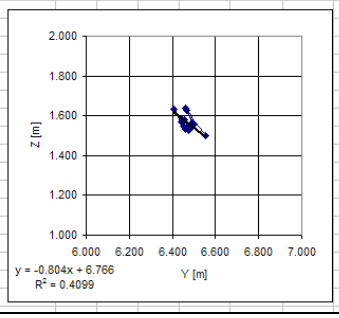
Reporting the results of detection and tracking of objects in static and moving mode completes the explanation of the algorithm. Once the processing is complete, the final results are automatically exported to an Excel® file. In Figure 32 (manually enhanced with colors and pictures for a better perception of each experiment) all clusters and their detailed information that were found in this particular experiment are listed. In this example there are 13 different objects. Object (cluster) number 7, for example, includes 92 filled voxels. Its cluster dimensions (width/length/height) can be read from the Matlab® program. The position of center-of-gravity of cluster 7 is X/Y/Z=7.114/6.531/1.537 [m]. Its averaged (since only linearly moving objects were assumed in all experiments) velocity vector is X/Y/Z=-0.843/0.105/-0.037 [m/s]. This cluster occurred in 17 out of 52 total frames and thus was represented in 33 percent of all frames in the entire experiment. All clusters that appeared in at least 5% of the entire frame grabbing sequence were analyzed in-depth (a quality control and review element to each cluster). All other clusters were considered to be a random event, e.g. cluster that describes noise (generally clusters with a voxel count of less than 60) or clusters that appear only once or twice in the entire frame sequence. Cluster number 7 had a velocity vector and was classified as a "moving object" (this algorithm does not identify objects, see research objectives Chapter 1). Below the data to each cluster other information is

المنارة للاستشارات
www.manaraa.com

presented. The time is measured for importing/cleaning and processing the range data. Depending on the experiment the time to process frames ranges between 8.5 seconds to 0.4 seconds per frame. This is mostly due to the Matlab™ program and the visualization (see Chapter 5 and 6 for recommendations). Images are presented to easily understand and recognize the characteristics of this experiment.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Experiment #PRE-01-0	Volume	COG			Velocity Vector			X min/max/delta	Y min/max/delta	Z min/max/delta	Angle X/Y, X/Z, Y/Z
2	Details for Object 7	Filled Voxels	X [m]	Y [m]	Z [m]	X [m/s]	Y [m/s]	Z [m/s]	[m]	[m]	[m]	[°]
3	Frame 27	66	8.358	6.439	1.568				8.358	6.439	1.568	-5.4
4	Frame 28	81	8.315	6.446	1.581				6.005	6.660	1.493	1.8
5	Frame 29	94	8.246	6.405	1.633				2.353	-0.221	0.076	-18.9
6	Frame 30	120	8.167	6.461	1.533	-1.186	0.833	-1.507				
7	Frame 31	107	8.062	6.469	1.535	-1.575	0.125	0.031				
8	Frame 32	88	7.972	6.450	1.550							
9	Frame 33	97	7.877	6.476	1.525	-1.414	0.394	-0.379				
10	Frame 34	88	7.781	6.457	1.581	-1.450	-0.292	0.839				
11	Frame 35	88	7.781	6.457	1.581							
12	Frame 36	89	7.573	6.480	1.530	-3.115	0.344	-0.755				
13	Frame 37	96	7.475	6.479	1.555	-1.471	-0.009	0.373				
14	Frame 38	92	7.359	6.484	1.543	-1.745	0.068	-0.176				
15	Frame 39	92	7.251	6.490	1.564	-1.614	0.098	0.310				
16	Frame 40	80	7.185	6.460	1.638	-0.991	-0.453	1.101				
17	Frame 41	91	7.127	6.502	1.559	-0.863	0.633	-1.172				
18	Frame 42	75	7.040	6.464	1.629	-1.312	-0.573	1.050				
19	Frame 43	94	6.835	6.555	1.498	-3.073	1.370	-1.972				
20	Average					-1.100	0.159	-0.141				





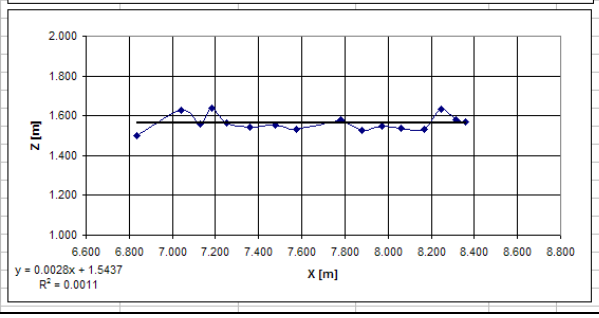


Figure 33: Analysis of Cluster #7

Each tab contains the information values to each cluster (see Figure 33). Cluster #7 started in range frame number 27 and continued to appear until range frame 43 (Column A). The count of filled voxels is listed in column B. In addition and equal to the VC position and velocity vector values mentioned above, each cluster tab contains the angular values describing the directional movement. The angles are based on the triangularization of entry (X,Y,Z) and exit (X,Y,Z) value of the VC. Since the box moved perpendicular to the sensor's FOV it directional angle should have been zero degrees.

المنارة للاستشارات
www.manaraa.com

The results of the experiment, however, measure an angle of 5.4 degrees (X/Y value in row 3, column L). The slight drifting of the box away from the sensor can be seen in Figure 31 and also has been recognized by a video camcorder mounted under the roof of the laboratory hall. The angular direction was plotted separately in elevation (Y/Z), front (X/Z), and plane (X/Y) view (each node reflects one VC of the cluster in one frame, fitted linear curve added, and gradient value is calculated).

4.7 Findings And Considerations During Algorithm Development

Several critical comments aroused during the algorithm development. It is commonly known in the data mining or image vision community that finding a well working clustering algorithm that achieves excellent segmentation results is very important and at the same time very hard. The literature review in particular highlighted agglomerative hierarchical clustering and clustering using DBSCAN. The latter one needs to be explored and most likely other algorithms as well.

The threshold parameters that were used to run the algorithm (see Table 5 and Table 6) meet approximately the safety needs on construction sites (e.g. maximum speed). It was defined that an object is recognized in static position when its location does not move more than 1 voxel. With the range frequency update rate of 15.2Hz a moving object can still be considered a static object unless it exceeds approximately 3mph. Especially at low speeds between 1 and 3mph, when workers may pay less attention, many injuries and even fatal accidents can happen. Thus some defined parameter values may not fit all desired application areas and need to be adjusted, e.g. 3D defect detection may ask for a small occupancy grid size in the millimeter range.

The calculation of the VC of clusters by averaging the position values of its voxels has limitations (precision vs. accuracy) and advantages (reflect changes quickly). A priori knowledge can help finding, for example, semi-permanent or permanent object

المنارة للاستشارات
www.manaraa.com

in space. Although the developed algorithm works well, it was not intended to find its optimal value. For instance, in order to reduce the computational cost significantly hardware and software processing as well as the development of sophisticated algorithms could have improved experimental results.

This research has demonstrated the fundamental principle of modeling dense range point clouds in (near) real-time (single range data frame processing between 8.5 and 0.4 seconds with a mean of 5.4 seconds). One of the challenges in programming this algorithm was to avoid computational costly loops and reruns of calculating the summary of voxel sizes and groups.

Another challenge was that by applying simple thresholds to separate objects, the algorithm forced some objects to split in two smaller clusters (e.g. as the demonstrated examples of the background wall). This disadvantage in handling the data set was still preferred since it allowed a simpler and less computational time consuming approach of removing small clusters without losing too much scene or object information.

One of the challenges, however, was that some clusters of moving objects were separated in two smaller objects making it impossible to track the entire object under the current parameter setting. A priori knowledge and the close position of the VC of clusters can help merging objects that split from one frame to the next. On the other side, objects that need to be separated, e.g. human walking on floor, can be handled similarly since the appearance of the person and the influence and shift of the VC of floor including the person can change the (experienced over multiple frames) static location of the VC of the ground. Including a probability parameter will improve the precision of the VCs.

A final challenge was to demonstrate the effect of the algorithm in the global map (World Model), when the sensor is moving. This effect can hardly be seen in the existing experimental environment. One solution is to extend the FOV and range of the sensor.

المنارة للاستشارات
www.manaraa.com

Other important future research needs for the algorithm development and testing can be:

- Track clusters by comparing several previous frames, not frame-to-frame.
- Real-time range data acquisition and processing above 20Hz.
- Path estimation.
- Deployment on heavy equipment for obstacle avoidance and performance of long-term tests to demonstrate less accidents occur. Use in real situation with real background.
- Combine advantages of several sensor to eliminate limitations (e.g. RFID or ultrawide-band sensing to retrieve a priori information with non-line-of-sight identification of objects and their dimensional (voxel) size).
- In the experimental setup use non-linear moving objects.

The clustering method in this research is based on hierarchical clustering and is called agglomerative because it merges clusters iteratively. The main weaknesses of agglomerative clustering methods are (Johnson, 1967):

- They do not scale well: time complexity of at least O(n2), where n is the number of total objects.
- They can never undo what was done previously.

The main advantages of agglomerative clustering methods are (Jain et al., 1999):

- They do not know the final number of clusters as K-Means and Fuzzy do.

المنارة للاستشارات
www.manaraa.com

- Single-linkage allows combining close sub-clusters (complete-linkage combines the farthest) in a group.

4.8 SUMMARY

This chapter has introduced an effective algorithm for modeling 3D environments in real-time with the particular application for construction safety. Experiments, results, analyses, conclusions, and an outlook for future research or other application areas in related engineering fields conclude this dissertation.

It can be expected that this area of research is expanding rapidly in a diverse field of application. In construction several significant impacts are foreseen in the near future from obstacle avoidance system, 3D as-planned vs. as-built comparison, rapid spatial models assisting project simulation, all of them ultimately decreasing cost and schedule, while increasing communication under project stakeholder, safety in machine navigation, and productivity in merging several emerging sensing technologies to a combined benefit.

المنارة للاستشارات
www.manaraa.com

Chapter 5: Experiments, Results, and Analyses

This chapter introduces two phases of experiments, its results and analyses. Preliminary experiments were conducted to determine some of the specific features of the developed algorithm, e.g. angular field-of-view of the 3D video range camera to precisely allocate range points in a predefined occupancy grid size. Once these initial test experiments of camera and algorithm were conducted to successfully complete the occupancy grid algorithm development, a secondary phase of experiments testing the accuracy of the algorithm in indoor and outdoor laboratory and field trial tests was started. The experimental setup including used target objects and design of the target space, the experiments procedure itself, and the results and analyses of all conducted experiments are presented and discussed in the following paragraphs.

5.1 Preliminary Work to define Algorithm Parameters and Experiments

To detect objects at their accurate position the angular as well as depth resolution of the FOV of the 3D video range camera are required to be known. Initial literature review efforts led by the ETH Zürich and the National Institute of Technology (NIST) indicated that so far limited calibration techniques exist for 3D range imaging cameras (Kahlman and Ingensand, 2005, and Lytle et al., 2005). A preliminary experiment was designed in this and other research to define the FOV of the 3D video range camera. Once the angular resolution of the camera in horizontal and vertical FOV was known these two parameters were implemented in the algorithm.

In a next step, to develop an experimental setup for detection and tracking in optical range sensing mainly two elements need to be determined before any experimental work can begin: (a) What are potential target objects that can characterize the application and what are its determining characteristics?, and (b) What features in the

المنارة للاستشارات
www.manaraa.com

field-of-view (FOV) of the range camera can limit or influence a physical experimental setup and what is needed to overcome any limitations? Answers to these questions are presented next.

5.1.1 Calibration of Sensor's Field-Of-View

As seen in Figure 34, to allocate range points into a three-dimensional occupancy grid, the initial step of the developed algorithm needed to convert range points from the Cartesian grid to a spherical coordinate system. This purpose required to know the angular as well as depth field of view that were determined in the following preliminary experiments.

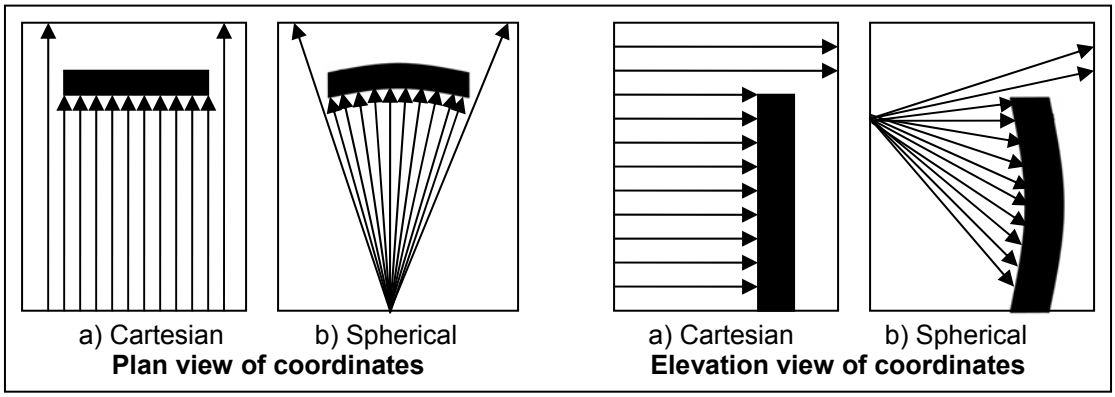


Figure 34: Light beams in Cartesian and spherical perspective

المنارة للاستشارات
www.manaraa.com

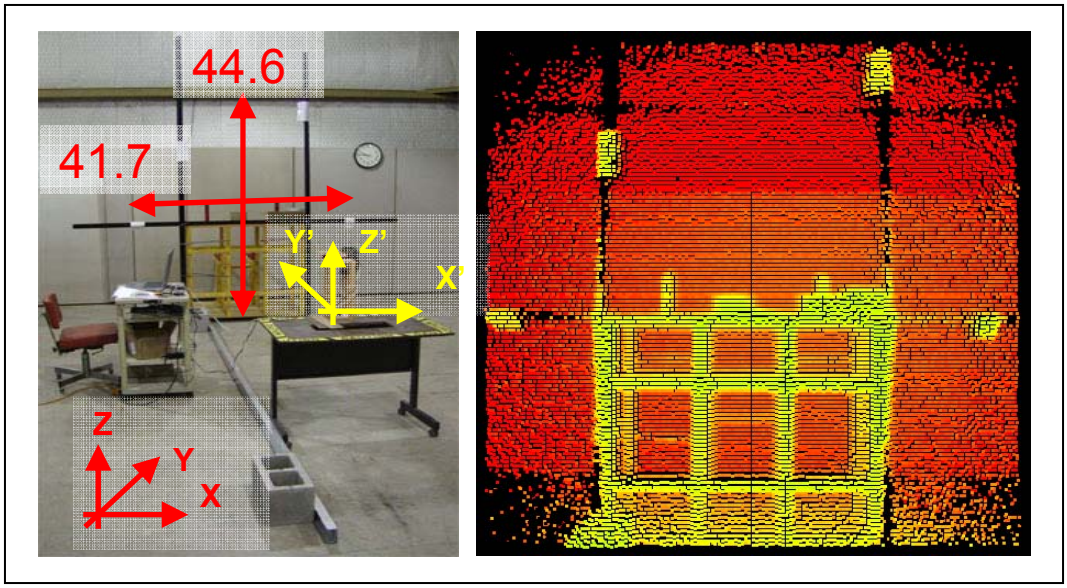


Figure 35: Sensor's angular field-of-view

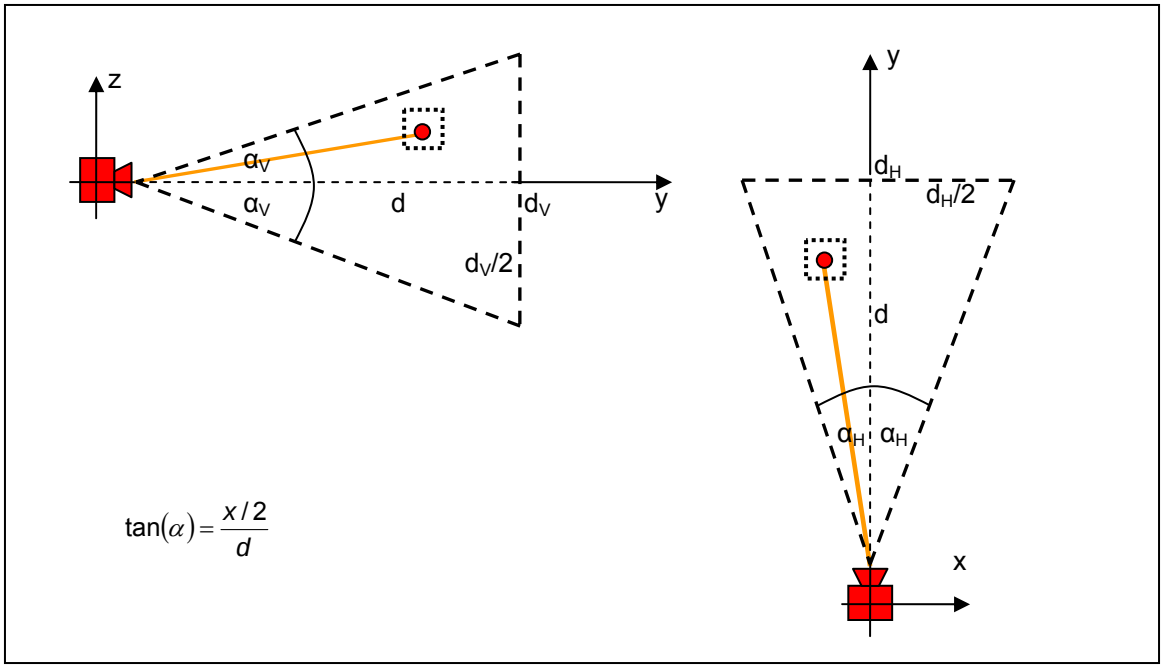


Figure 36: Elevation and plane view of camera's FOV

المنارة للاستشارات
www.manaraa.com

Table 7: Angular calibration of field-of-view

Sensor	SwissRanger 2, Version A				SwissRanger 2, Version B			
d [m]	d_H [m]	d_V [m]	α_H [°]	α_V [°]	d_H [m]	d_V [m]	α_H [°]	α_V [°]
0.50	0.404	0.388	44.020	42.434	0.394	0.410	43.031	44.610
1.00	0.669	0.751	37.009	41.183	0.686	0.807	37.883	43.970
1.50	1.123	1.297	41.066	46.785	1.147	1.238	41.868	44.871
2.00	1.514	1.695	41.484	45.953	1.580	1.784	43.130	48.098
2.50	2.013	2.105	43.882	45.685	1.880	1.952	41.233	42.673
3.00	2.395	N/A	43.543	N/A	2.332	2.384	42.500	43.361
3.50	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
4.00	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
4.50	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
5.00	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
6.50	5.230	N/A	43.853	N/A	5.040	N/A	42.403	N/A
Average			**42.122**	**44.408**			**41.721**	**44.597**

As seen in Figure 35, moving flags on two crosses towards the boundaries of the video range camera's FOV allowed measuring the horizontal and vertical distances of the field of view. Knowing the distance of the camera to the object and the distances (d_H and d_V) of the flags to each using a laser range finder (Leica® Distometer) other resulted in the angular FOV (α_H and a_V) (see Figure 36). The final FOV was calculated by taking the average of the angular measurements to multiple distances d of the camera to the object. This rough calibration was conducted to measure the FOV of the sensor. As result, the vertical angle was experienced at 44.6° and the horizontal angle at 41.7° (compared to a measurement conducted by Gut (2004) with same prototypes: horizontal 46.0° and vertical 42.9° angle). In Table 7 the FOV calibration results to both prototype 3D video range cameras (SwissRanger A and B) are presented when the sensor (model A and B) had specific distances to a background object. Some values could not be measured ("N/A") since the cameras FOV extended the reach of the target cross.

المنارة للاستشارات
www.manaraa.com

5.1.2 Non-Ambiguous Distance and Voxel Size

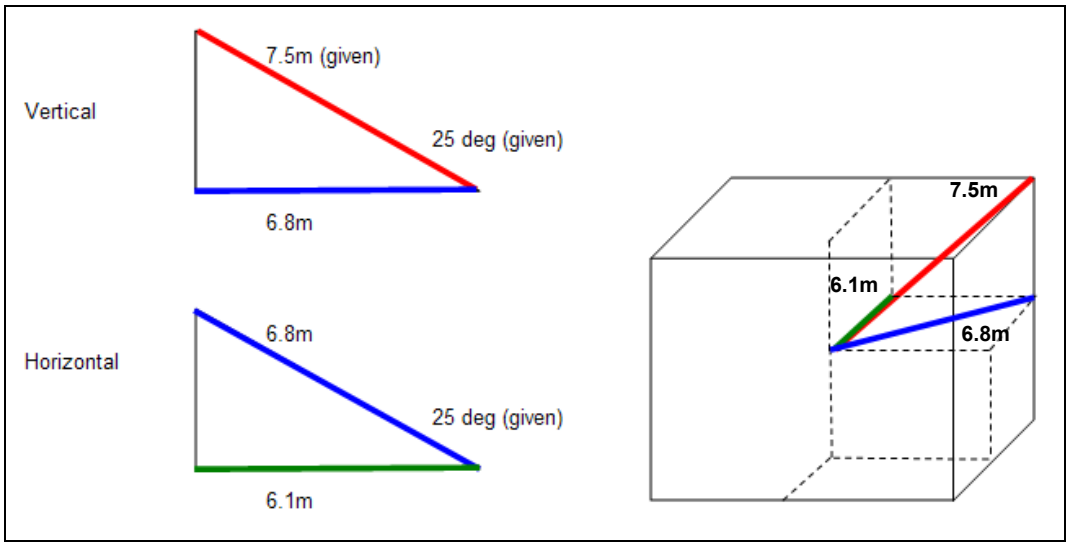


Figure 37: Sensor's maximum range limit to non-ambiguous range

The non-ambiguous range of the sensor is approximately 7.5m including range points to objects that appear in corners (as seen in Figure 37). As a result, this reduces the maximum distance to objects that are directly in front of the sensor to approximately maximal 6.1m (assuming a maximum angle of maximum 50° in angular view). To avoid saturation of pixels, objects also need to be further than 1.2m away from the sensor. As a note, future prototypes that may use multiple illumination frequencies may extend the un-ambiguous measurement range to multiples of 7.5m, however, the minimal and maximal distances are helpful boundaries to operate algorithm and sensor.

المنارة للاستشارات
www.manaraa.com

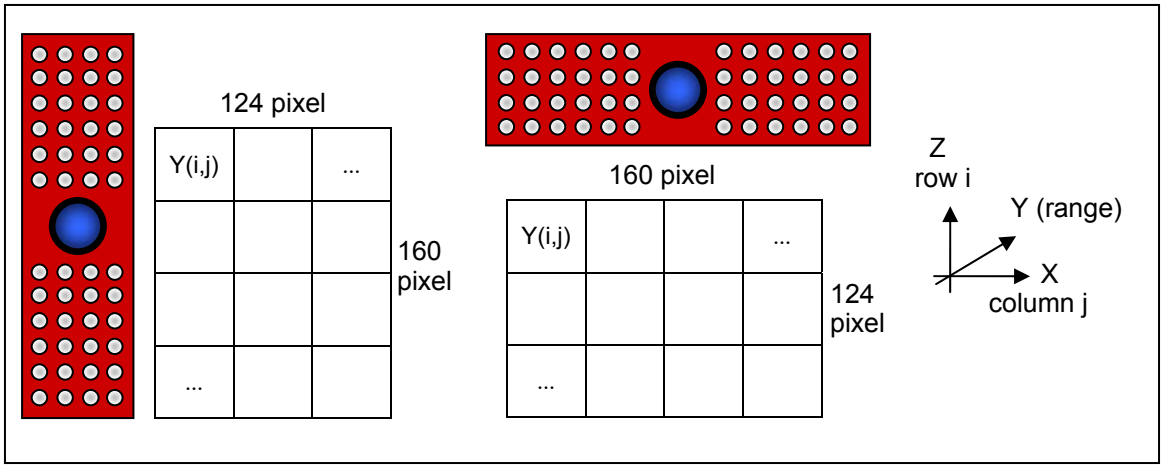


Figure 38: Vertical and horizontal position of SwissRanger 2 and its output matrix

Table 8: Analysis of field-of-view of SwissRanger 2b

SR-2b FOV	α_H [°]	d [m]	d_H [m]	Pixel width [cm]	SR-2b FOV	α_V [°]	d [m]	d_V [m]	Pixel width [cm]
Horizontal 160 pixel	41.721	7.50	5.716	3.57	Vertical 124 pixel	44.597	7.50	6.151	4.96
		7.00	5.335	3.33			7.00	5.741	4.63
		6.50	4.954	3.10			6.50	5.331	4.30
		6.00	4.573	2.86			6.00	4.921	3.97
		5.50	4.192	2.62			5.50	4.511	3.64
		5.00	3.811	2.38			5.00	4.101	3.31
		4.50	3.430	2.14			4.50	3.691	2.98
		4.00	3.049	1.91			4.00	3.281	2.65
		3.50	2.668	1.67			3.50	2.871	2.32
		3.00	2.286	1.43			3.00	2.461	1.98
		2.50	1.905	1.19			2.50	2.050	1.65
		2.00	1.524	0.95			2.00	1.640	1.32
		1.50	1.143	0.71			1.50	1.230	0.99
		1.00	0.762	0.48			1.00	0.820	0.66
		0.50	0.381	0.24			0.50	0.410	0.33

Another important feature for the occupancy grid algorithm was to determine the minimal usable voxel size in an occupancy grid. Due to the angular resolution and the collected number range points (horizontal 124 pixels and vertical 160 pixels) the pixel at a maximum distance 7.5m covers a physical area of width/height of 3.57cm/4.96cm (see Table 8). As a result, objects that are closer than 4.96cm at a distance of 7.5m would still be represented by only one range point, thus making the minimal reasonable grid size 5cm (assumed an equally sized grid width, length, and height).

المنارة للاستشارات
www.manaraa.com

5.1.3 Characteristics of Target Objects in Experiments

Types of construction resources include static and moving objects, and living and non-living objects, e.g. humans, vehicles, equipment, and materials. The preliminary experimental setup was based on the principle "from simple to complex" objects and scenes. Initial experiments had one geometrically simple target (e.g. box) that initially was static. Once the algorithm successfully performed the complexity level was increased to test the algorithm performance again, e.g. object was propelled. Through adding one or more static or moving targets to the sensor's FOV at the same time the complexity level of the experimental setup was elevated further.

Object specific parameters and the physical experimental setup are chosen to verify the working of the three-dimensional occupancy grid algorithm. Moving objects were limited to linear direction and had speeds of less than 3.5km/h at angles to the camera of 0, 30 and 90 degrees. Since no inertial motion unit was available to determine the object speed, marked control points on ground allowed recognizing the entry and exit of objects in the FOV of the 3D range camera via a video camcorder (Canon EZ80) that was installed under the lab roof. This post-range data processing method allowed a direct comparison between the original object speed from the camcorder with the speed determined by the occupancy grid based algorithm based on data from the 3D range camera. Better access to time controlled information access can solve inaccuracies and the limitation of this approach to determine the precise entry and exit and time needed of the object to travel through the FOV of the range camera. The maximum size of the face of objects (length multiplied with height) was $5m^2$ (skid steer loader: 2m x 2.5m) and the smallest object had a face area of $0.16m^2$ (cardboard box: 0.4m x 0.4m). Target objects further had distributed locations in the sensor's FOV.

المنارة للاستشارات
www.manaraa.com

5.1.4 Characteristics of Experimental Field-Of-View

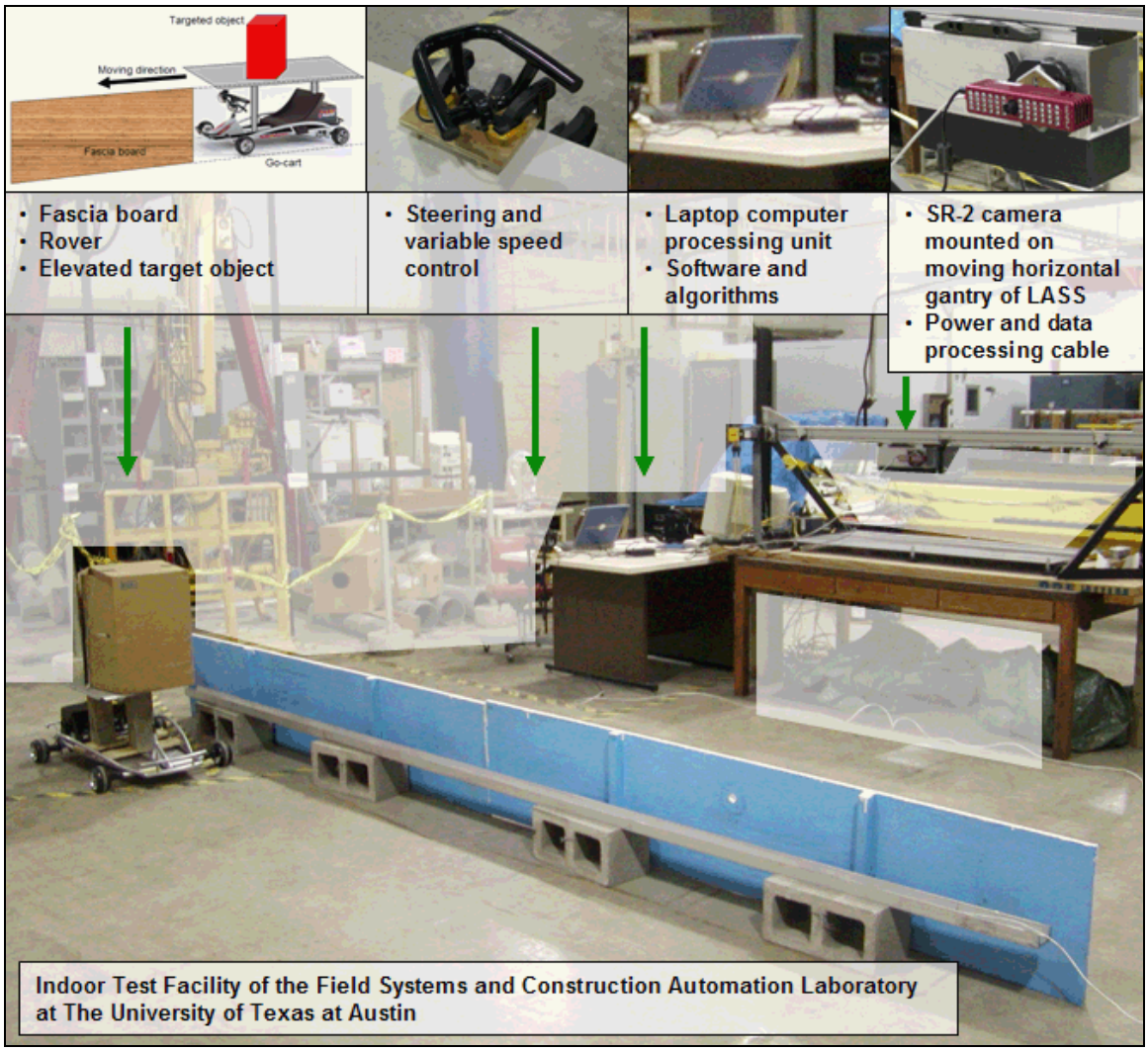


Figure 39: Components of indoor experiments

In indoor experimental research the range sensor was mounted on an existing horizontal gantry (length: 1.5m) of the Laser Aggregate Scanning System (LASS). Using the LASS system was advantageous since its pneumatic system allowed generating a constant speed at various levels (i.e. 0.05m/s, 0.1m/s, and max. 0.2m/s). As seen in Figure 39 a "remote" (cable-controlled) electric go-cart (max. speed 3.5km/h) was

المنارة للاستشارات
www.manaraa.com

developed with an elevated platform to propel objects at three different angles and speeds (see Figure 40). The initial experimental setup used a static fascia board to cover the cart from the FOV of the sensor. This allowed collecting only those range points that actually came from the moving object itself.

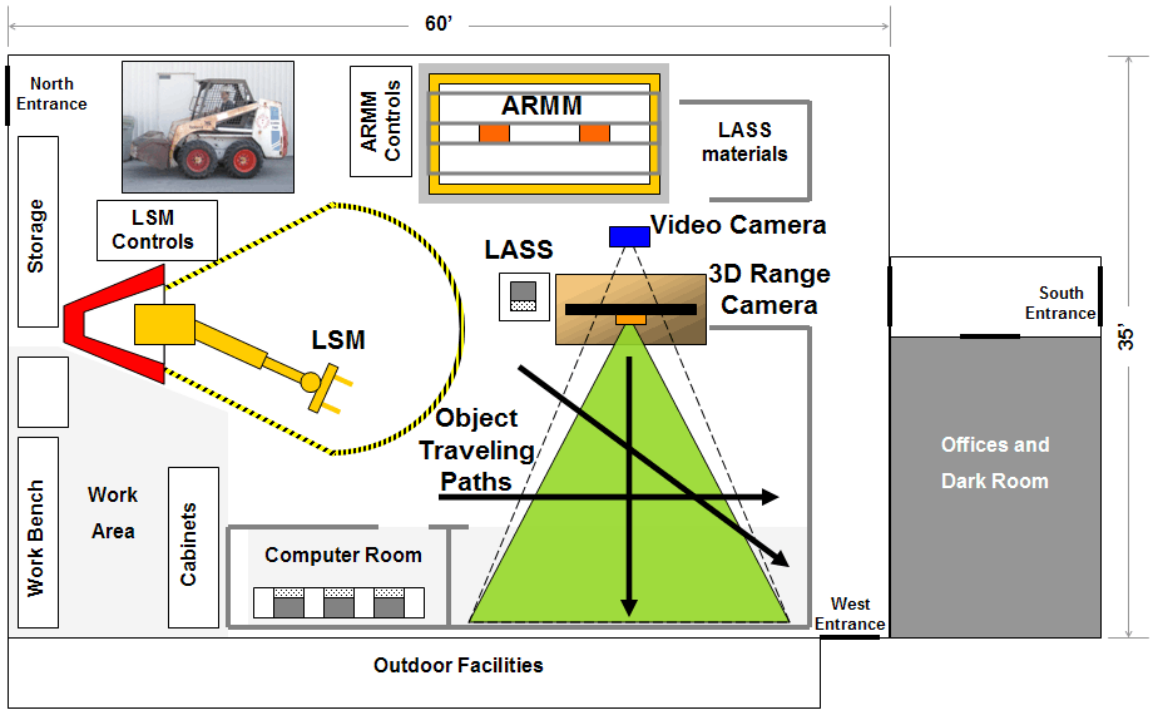


Figure 40: Overview of indoor experimental setup

In a next step, a Total Station was used to determine characteristic points in space of the sensor's FOV. Visible points on the ground and wall marked the FOV that allowed a video camcorder (installed underneath the lab roof) to track the time the object needed to enter and exit the sensor's FOV.

Since the LASS and the cart had a ramp up and down phase to accelerate and decelerate to reach the desired experimental speed (defined to reach a steady slow or fast), the start point of the object was located considerably before entering the FOV of the sensor.

المنارة للاستشارات
www.manaraa.com

Since the cart had no speedometer, the accelerator knob was held steady once the cart carrying the object entered the sensor's FOV. Knowing the time difference between entry and exit of the object in the sensor's FOV (time elapsed on the linear travel path) allowed calculating the real speed of the object. The real speed then was compared with the speed that resulted from the range processing algorithm.

Preliminary experiments had shown that the cart (no inertial steering or guidance control) needed some kind of guidance to not drift in unpredictable directions. Even though the front tires were fixed with screws to the frame and body of the cart so they could not move, often the cart slided off in different directions. To prevent derailing and to ensure that the original direction and position of an object can be compared to the processed range information, a fixed-on-the-ground U-shaped metal was installed that guided on side of the cart's wheels.

5.1.5 Summary of Indoor Experimental Setup

In the indoor experimental setup a 3D video range camera was used to measure dense range point clouds to the field-of-view (FOV) of the camera. The location of points characterizing the camera's FOV was determined using a Total Station. Static objects were positioned in the FOV and the location position measured with the Total Station. Moving objects were propelled on a linear travel path through the FOV. The video range camera's FOV, entry and exit points of variously shaped objects, and their direction and travel path are illustrated in Figure 41.

المنارة للاستشارات
www.manaraa.com

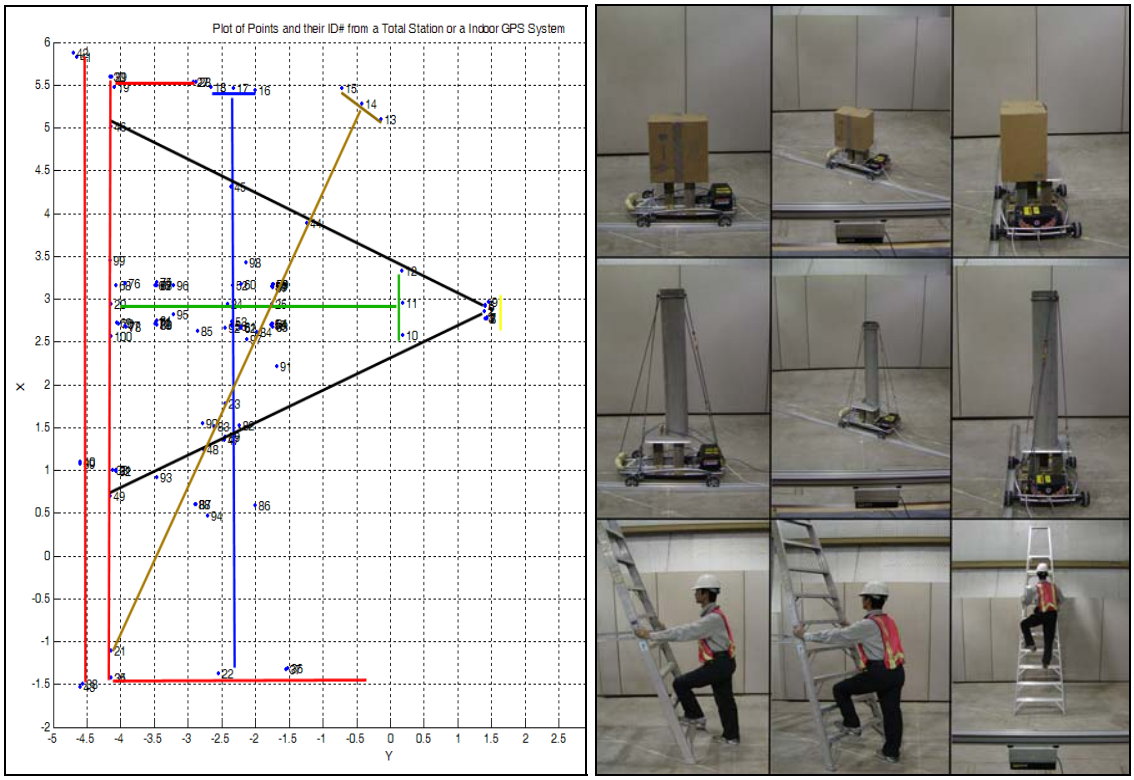


Figure 41: Geodesy to indoor experimental setup (sensor's FOV, boundaries, tracks) for different resources and angles

5.1.6 Summary of Outdoor Experimental Setup

The outdoor experimental setup is illustrated in Figure 42 and Figure 43. The 3D range acquisition used a static 3D video range camera mounted on a two-axis pan-tilt unit (PTU) of Directed Perception Inc. that was put on top of a tripod. A mobile computing unit (Dell Inspiron 5150, Pentium 4, 3.06Ghz, 512 MB, 40GB) hosted the range data acquisition software and a C program code that continuously reads pan and tilt angles from the PTU. The range data was acquired and stored on the hard drive of the computer processing unit and then processed indoors. The separation of data acquisition and processing was necessary to ensure the best possible quality of collecting and analyzing the range data. Ultimately both steps will be combined.

المنارة للاستشارات
www.manaraa.com

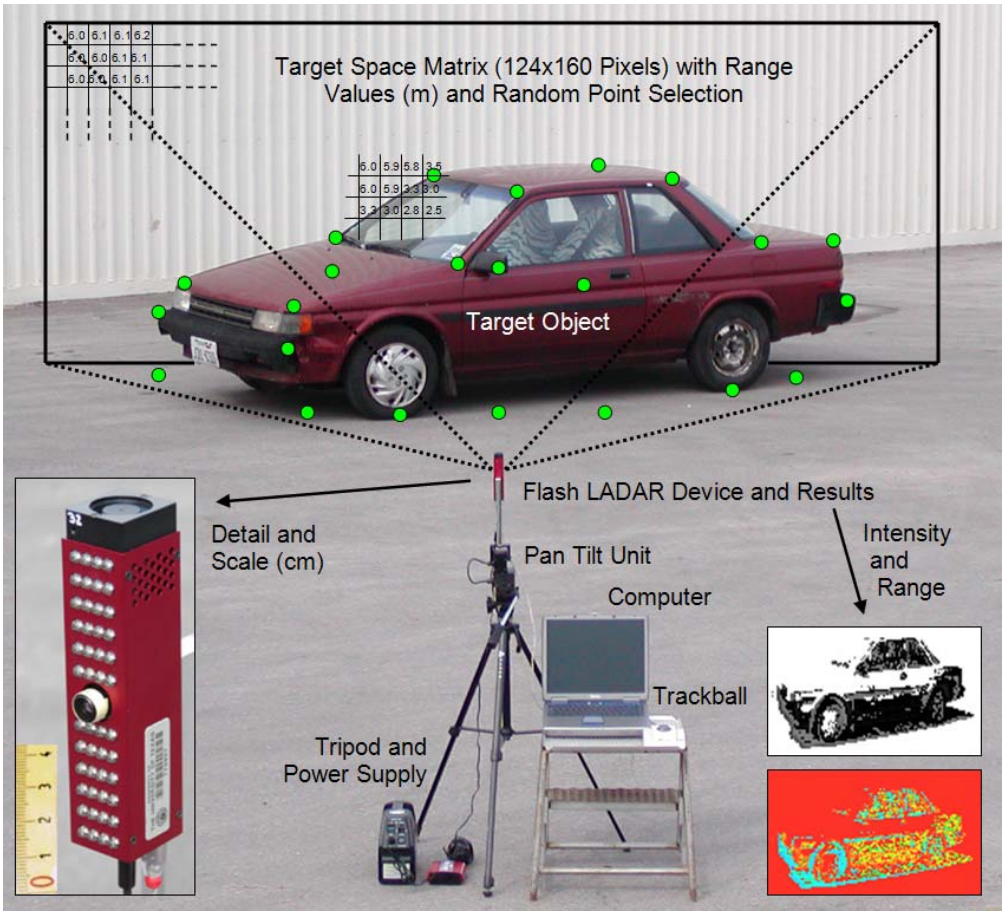


Figure 42: 3D video range camera setup in outdoor environment

Figure 43: Overview of outdoor experimental setup

المنارة للاستشارات
www.manaraa.com

5.2 EXPERIMENTAL RESEARCH METHODOLOGY

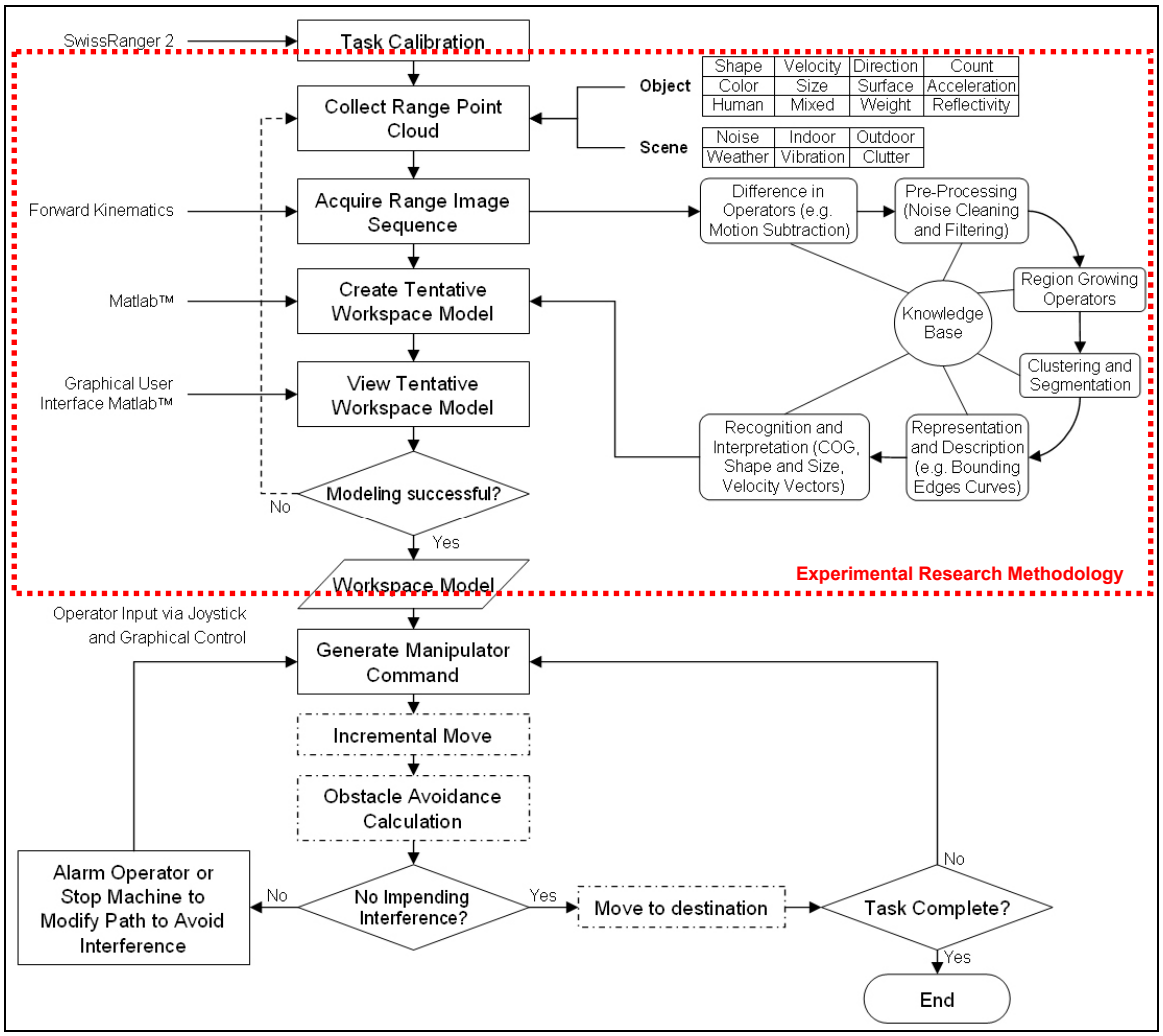


Figure 44: Flowchart of experimental plan in an obstacle-avoidance process for heavy machine operation (extended, after Gonzales and Woods, 1992, and Kim et al., 2004)

In an overview picture, this research is part of a bigger methodology that focuses on the development of obstacle detection and navigation algorithms for safe heavy construction equipment movement. Figure 44 illustrates the entire algorithm of the machine movement. All steps within the dotted frame are part of the algorithm developed in this research and concentrate on object detection and tracking.

المنارة للاستشارات
www.manaraa.com

A typical example for an obstacle avoidance system based on a 3D range camera the experimental plan foresees some of the following steps:

1. Calibrate sensor to achieve accurate range measurements to obstacles.
2. Create an indoor environment which is less applicable to ambient noise.
3. Determine location and orientation of start position of sensor. Build World Model including a priori information of permanent static objects.
4. Collect range point clouds in frame sequences (Cartesian format). Store on computer hard drive (data acquisition and processing separated to avoid errors).
5. Transform in spherical coordinate system and allocate range points in 3D occupancy grid. Apply fill factor and noise removal thresholds on occupancy grid.
6. Segment and group voxels to objects based on hierarchical agglomerative clustering technique. Apply threshold to define object as static or moving.
7. Determine to each object: Size of voxels; position of volume centroid (VC); object dimension; direction, travel length, and velocity of same VC from one frame to the next. Save analysis on disk.
8. Visualize detection and tracking of 3D model in frames (Local Model) and movies. Create workspace model if successful and supply data to World Model.
9. If experiment and data processing were unsuccessful, repeat experiment, adjust experimental setup, or refine processing algorithm.
10. Create a manual measurement protocol to each experiment and document observations and findings.
11. Continue with on-machine obstacle avoidance system, e.g. generate manipulator command, movement, alarm, and correctional movements.

المنارة للاستشارات
www.manaraa.com

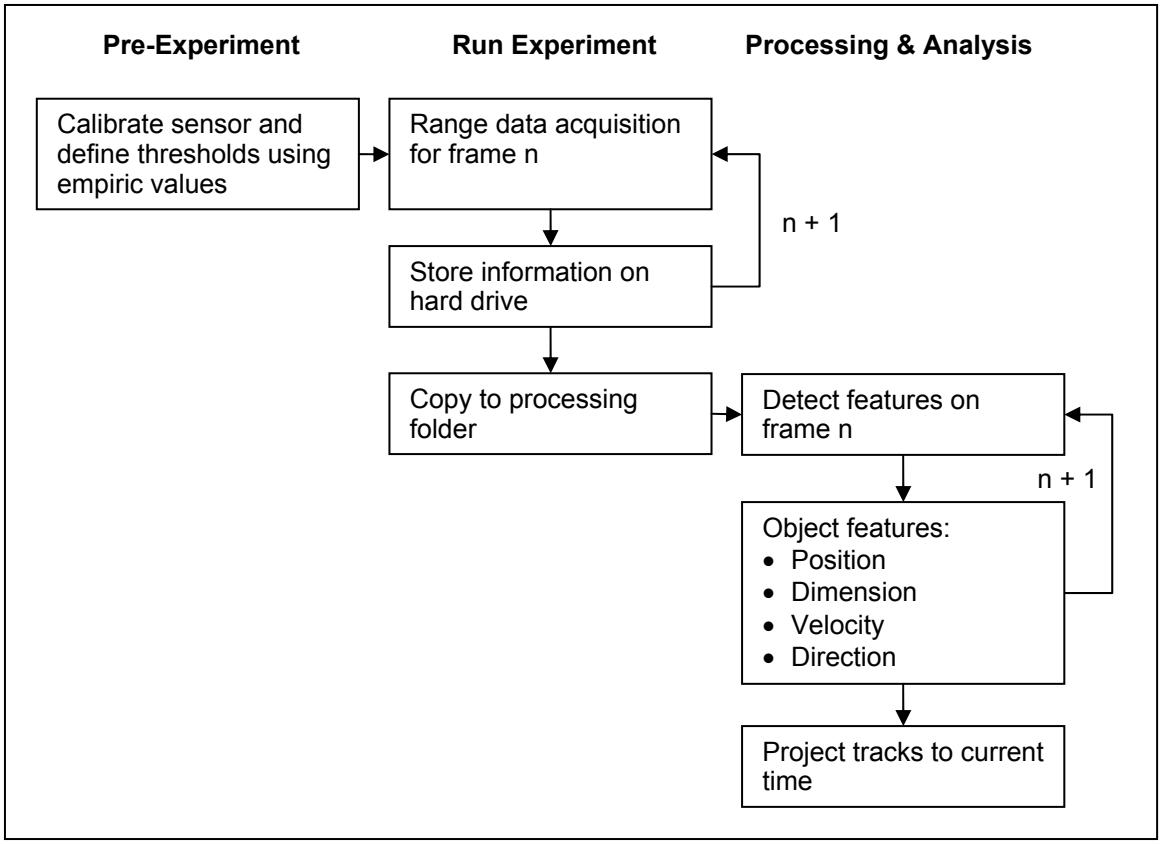


Figure 45: Phases of range data collection, processing, and analyses

The part of the obstacle avoidance system solved in this research concentrates on three main steps (see Figure 45): (1) Range data acquisition from a prototype 3D range video camera, (2) Range data processing using the developed occupancy grid algorithm to detect and track 3D object model information, and (3) Analyses of experimental results. These three individual steps that ultimately work seamlessly in a row were physically separated. This simplification allowed concentrating on developing the most accurate and best performance of each step. Several other reasons justify this step as well, e.g. a real-time closed loop tracking system can be part of the acquisition system. If chosen to be, it often requires special purpose (expensive) hardware that ultimately keeps focused on the target.

المنارة للاستشارات
www.manaraa.com

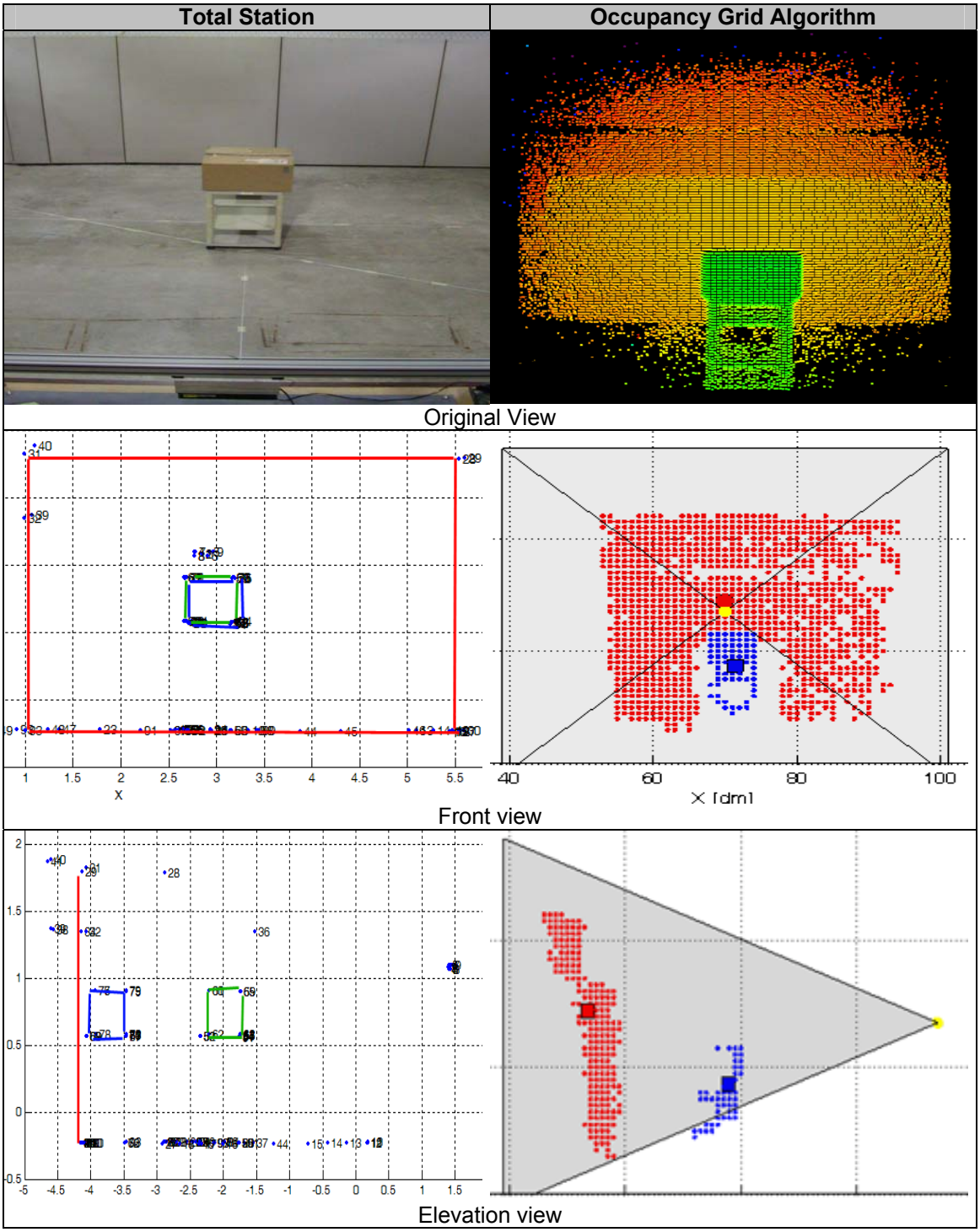
5.3 EXPERIMENTS TO DETERMINE POSITION AND DIMENSION ACCURACY

Different indoor experiments (presented are Test #20 and Test #28) were conducted to determine the optimal configuration of the prototype 3D range camera for achieving accurate object position and dimensions. The goal of these experiments was to measure and compare the true to the modeled position of significant object points, e.g. box corners and edges, and the dimension (width, depth, height = X,Y,Z) of objects.

In a first experiment, as seen in Figure 46, the scene consists of one object once at a medium distance and the same object once at a far distance (close to wall). In a second experiment several boxes at different locations were positioned in the FOV of the 3D video range camera. In the second experiment the 3D video range camera was located at various distances to the target objects (cardboard boxes).

As explained in previous chapters, a Total Station was used to measure (calibrate) the "true" location of relevant scene points including the edges to the cardboard boxes and the distance to the background wall. The output of the Total Station generated a point cloud in which each point is numbered in the order it was taken. All object and sensor locations (e.g. an edge of a box) were then directly compared to the same point in the occupancy grid based 3D model that was created using the range data output from the 3D video range camera. The differences in location were measured, recorded, and analyzed. The accuracy values of the distance to single points and the dimension of objects were determined. Results are displayed in Table 9 and Table 10. In both tables the deviations to results to single experiments is recorded and compared to each other. They show which points are compared against each other (Start and end point of object, their point identification number, and their relative distance in each axis) in the 3D model that was generated by the Total Station and the 3D occupancy grid. Differences in their measurement and their errors are given in [m] and [%].

المنارة للاستشارات
www.manaraa.com



المنارة للاستشارات
www.manaraa.com

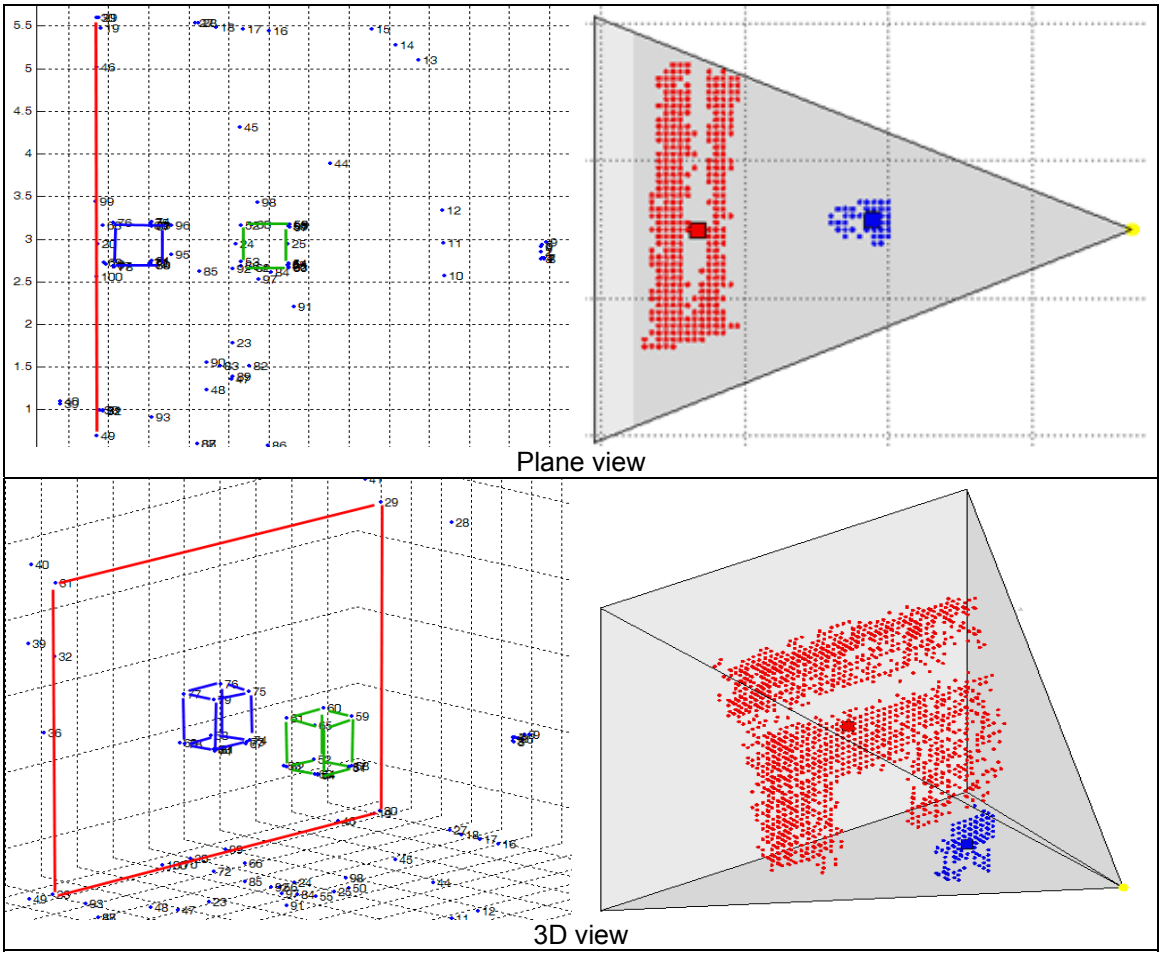


Figure 46: Position and dimension measurement – Experiment #20

In Figure 46 two boxes are displayed since the Total Station stores all sparse points in one file. For illustration purposes both boxes were plotted in the same figure, one appears closer to the sensor, and the other one, has very close distance to the front wall. The sensor can be seen in all Total Station images as a denser point cloud, since several reference measurements were taken to acquire the orientation information.

المنارة للاستشارات
www.manaraa.com

Table 9: Position measurement error of experiment #28

Position Verification Exp. #20

			Total Station [m]				Occupancy Grid [m]			Error [m]			Error [%]		
			Points	Distance			Distance*						compared to Total Station		
Experiment #20	Start point	End point	from-to	X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z
Test 20-01-01	Box	Box	59 , 65	0.50			0.70			0.20			40.0%		
	SR-2	Box	5 , 57		3.20			3.50			0.30			9.4%	
	Box	Box	58 , 59			0.33			0.50			0.18			53.8%
	SR-2	Lower Wall	5 , 100		5.56			5.60			0.04			0.8%	
	SR-2	Upper Wall	5 , 40		6.01			6.30			0.29			4.8%	
Test 20-01-02	Box	Box	75 , 79	0.50			0.60			0.10			20.2%		
	SR-2	Box	5 , 75		4.88			5.30			0.42			8.6%	
	Box	Box	74 , 75			0.32			0.50			0.18			55.3%
	SR-2	Lower Wall	5 , 100		5.56			5.60			0.04			0.8%	
	SR-2	Upper Wall	5 , 40		6.01			6.20			0.19			3.2%	
				delta x											
Test 20-01-01	SR-2	Front Box	1 , 63	0.19			0.20			0.01			7.5%		
	SR-2	Front Box	1 , 58	0.32			0.40			0.09			27.0%		
Test 20-01-02	SR-2	Wall Box	1, 80	0.16			0.10			-0.06			-38.3%		
	SR-2	Wall Box	1 , 74	0.35			0.40			0.05			15.3%		
Test 20-01-01	Box	Front Box	50 , 51			0.79			0.80			0.01			1.0%
	Box	Front Box	54 , 55			0.79			0.80			0.01			0.9%
Test 20-01-02	Box	Wall Box	66 , 67			0.79			0.80			0.01			1.1%
	Box	Wall Box	71 , 72			0.79			0.80			0.01			0.9%
Average				0.25	5.20	0.79	0.28	5.42	0.80	0.02	0.21	0.008	2.9%	4.6%	1.0%

* measured by "hand" from single occupancy grid figures

** cannot be used for position but for dimension calculation

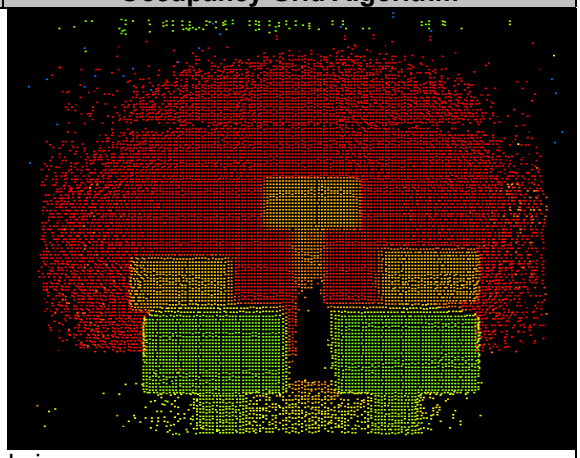
المنارة للاستشارات
www.manaraa.com

Table 10: Position measurement error of experiment #28

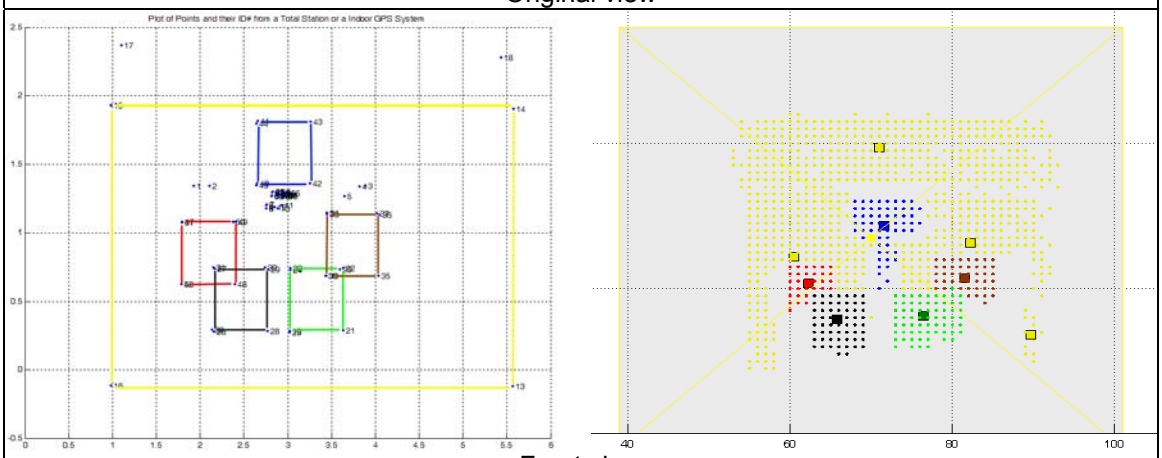
Position Verification Experiment #28

		Total Station [m]				Occupancy Grid [m]			Error [m]			Error [%]		
		Points	Distance			Distance						compared to Total Station		
	Start Point and End Point of frame #50 only	From-to	X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z
Test 28-01	SR-2 to bottom left corner of box front left	14, 28	-0.07	0.31	-0.09	-0.07	0.33	-0.13	0.00	0.02	-0.04	0.9%	6.9%	43.0%
	SR-2 to bottom left corner of box front right	14, 21	0.02	0.31	-0.09	0.03	0.33	-0.13	0.01	0.02	-0.04	74.4%	7.3%	43.8%
	SR-2 to bottom right corner of box back left	14, 50	-0.05	0.43	-0.06	-0.05	0.47	-0.08	0.00	0.04	-0.02	10.4%	9.5%	42.1%
	SR-2 to bottom left corner of box back middle	14, 42	-0.02	0.45	0.02	-0.02	0.47	0.01	0.00	0.02	-0.01	0.0%	4.9%	-42.2%
	SR-2 to bottom left corner of box back right	14, 35	0.06	0.44	-0.05	0.08	0.47	-0.08	0.02	0.03	-0.03	35.6%	6.3%	59.4%
Test 28-02	SR-2 to bottom left corner of box front left	14, 28	-0.07	0.31	-0.09	-0.07	0.34	-0.14	0.00	0.03	-0.05	0.9%	10.1%	54.0%
	SR-2 to bottom left corner of box front right	14, 21	0.02	0.31	-0.09	0.03	0.34	-0.13	0.01	0.03	-0.04	74.4%	10.6%	43.8%
	SR-2 to bottom right corner of box back left	14, 50	-0.05	0.43	-0.06	-0.05	0.46	-0.08	0.00	0.03	-0.02	10.4%	7.2%	42.1%
	SR-2 to bottom left corner of box back middle	14, 42	-0.02	0.45	0.02	-0.02	0.48	0	0.00	0.03	-0.02	0.0%	7.1%	-100.0%
	SR-2 to bottom left corner of box back right	14, 35	0.06	0.44	-0.05	0.08	0.47	-0.08	0.02	0.03	-0.03	35.6%	6.3%	59.4%
Test 28-03	SR-2 to bottom left corner of box front left	72, 28	-0.08	0.25	-0.10	-0.02	0.29	-0.08	0.06	0.04	0.02	-74.1%	17.7%	-19.9%
	SR-2 to bottom left corner of box front right	72, 21	0.01	0.25	-0.10	0.01	0.29	-0.08	0.00	0.04	0.02	7.5%	18.3%	-19.5%
	SR-2 to bottom right corner of box back left	72, 50	-0.05	0.37	-0.07	-0.06	0.39	-0.08	-0.01	0.02	-0.01	12.8%	6.3%	22.5%
	SR-2 to bottom left corner of box back middle	72, 42	-0.03	0.39	0.01	-0.03	0.4	0.08	0.00	0.01	0.07	7.5%	3.7%	863.9%
	SR-2 to bottom left corner of box back right	72, 35	0.05	0.38	-0.06	0.07	0.39	0.01	0.02	0.01	0.07	37.0%	2.7%	-116.9%
Test 28-04	SR-2 to bottom left corner of box front left	55, 28	-0.08	0.21	-0.10	-0.02	0.26	-0.09	0.06	0.05	0.01	-75.6%	21.2%	-9.2%
	SR-2 to bottom left corner of box front right	55, 21	0.00	0.21	-0.10	0.01	0.27	-0.09	0.01	0.06	0.01	108.3%	26.5%	-8.7%
	SR-2 to bottom right corner of box back left	55, 50	-0.06	0.34	-0.06	-0.06	0.36	-0.01	0.00	0.02	0.05	4.0%	7.4%	-84.5%
	SR-2 to bottom left corner of box back middle	55, 42	-0.03	0.35	-0.01	-0.03	0.38	0.07	0.00	0.03	0.08	-7.4%	7.3%	-869.2%
	SR-2 to bottom left corner of box back right	55, 35	0.05	0.35	-0.06	0.07	0.36	0	0.02	0.01	0.06	50.2%	3.4%	-100.0%
Average									0.16	0.10	-0.05	15.64%	9.54%	-4.81%

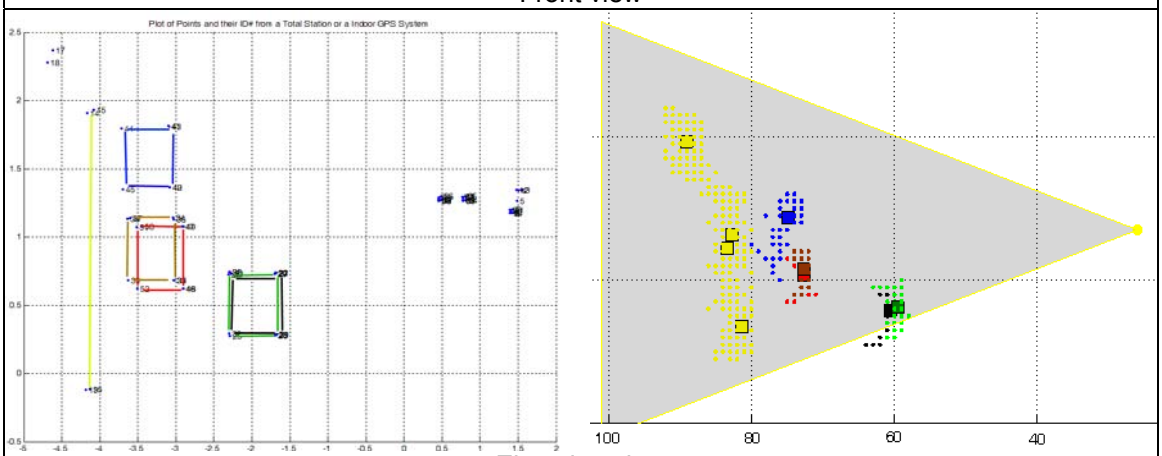
المنارة للاستشارات
www.manaraa.com

Total Station	Occupancy Grid Algorithm
	

Original view



Front view



Elevation view

المنارة للاستشارات
www.manaraa.com

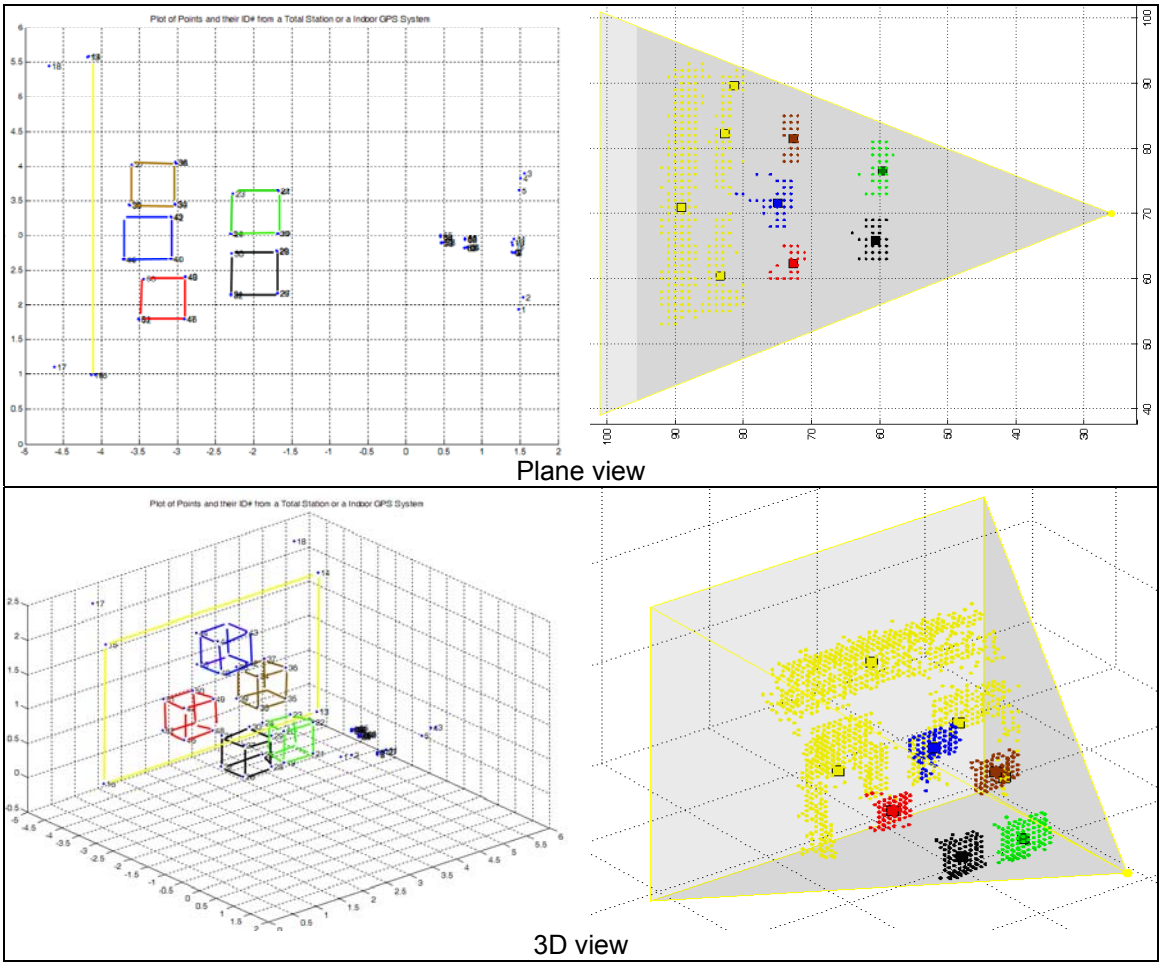


Figure 47: Position and dimension measurement – Experiment #28

In total, experiment #20 had 4 distinct contributions to determine the position accuracy. Experiment #28 had 20 distinct contributions of such kind. Results can be seen in Table 11. Maximum positioning error in either axial direction is 10cm or about 10%.

Table 11: Summary of position accuracy

	Measurements	Error [m]			Error [%]		
		X	Y	Z	X	Y	Z
Experiment #20	4	0.02	0.22	0.01	2.9	4.6	1
Experiment #28	20	0.16	0.10	-0.05	15.6	9.5	-4.8
Average weighted	24	0.10	0.09	-0.03	10.1	6.5	-2.9

المنارة للاستشارات
www.manaraa.com

Table 12: Dimensional error from experiment #22 with a box

Experiment #22		Occupancy Grid [10cm]			Reality [10cm]			Error [10cm]		
Object: Box	Frame # picked	X	Y	Z	X	Y	Z	X	Y	Z
Test 22-01-01	35	6	5	7						
	40	7	3	8						
	45	6	4	7						
	Average	6.3	4.0	7.3	5.0	4.0	6.0	1.3	0.0	1.3
Test 22-01-03	25	7	3	6						
	30	6	3	7						
	33	5	6	6						
	Average	6.0	4.0	6.3	5.0	4.0	6.0	1.0	0.0	0.3
Test 22-01-05	35	7	5	7						
	40	6	4	7						
	45	6	5	6						
	Average	6.3	4.7	6.7	5.0	4.0	6.0	1.3	0.7	0.7
Test 22-01-07	25	7	5	7						
	30	7	7	8						
	33	5	6	7						
	Average	6.3	6.0	7.3	5.0	4.0	6.0	1.3	2.0	1.3
Test 22-01-09	40	6	8	5						
	45	6	7	6						
	50	5	4	6						
	Average	5.7	6.3	5.7	5.0	4.0	6.0	0.7	2.3	-0.3
Test 22-01-11	15	6	4	7						
	20	6	6	6						
	25	5	4	8						
	Average	5.7	4.7	7.0	5.0	4.0	6.0	0.7	0.7	1.0
Average		6.1	4.9	6.7	5.0	4.0	6.0	1.1	0.9	0.7
Error	"+" means object in occupancy grid is bigger than in reality							21.1%	23.6%	12.0%

Table 13: Dimensional error from experiment #22 with pipe

Experiment #22		Occupancy Grid [10cm]			Reality [10cm]			Error [10cm]		
Object: Pipe	Frame # picked	X	Y	Z	X	Y	Z	X	Y	Z
Test 22-02-01	50	2	4	14						
	55	2	4	14						
	60	2	5	15						
	Average	2.0	4.3	14.3	2.0	2.0	13.0	0.0	2.3	1.3
Test 22-02-03	25	2	5	15						
	30	2	5	15						
	35	2	4	16						
	Average	2.0	4.7	15.3	2.0	2.0	13.0	0.0	2.7	2.3
Test 22-02-05	45	3	5	15						
	50	2	3	15						
	55	2	3	15						
	Average	2.3	3.7	15.0	2.0	2.0	13.0	0.3	1.7	2.0
Test 22-02-07	35	2	4	16						
	40	2	3	15						
	44	2	4	15						
	Average	2.0	3.7	15.3	2.0	2.0	13.0	0.0	1.7	2.3
Test 22-02-09	25	2	2	12						
	30	2	2	14						
	35	2	3	12						
	Average	2.0	2.3	12.7	2.0	2.0	13.0	0.0	0.3	-0.3
Test 22-02-11	20	2	3	11						
	25	2	2	12						
	30	2	2	14						
	Average	2.0	2.3	12.3	2.0	2.0	13.0	0.0	0.3	-0.7
Average		2.1	3.5	14.2	2.0	2.0	13.0	0.1	1.5	1.2
Error	"+" means object in occupancy grid is bigger than in reality							2.8%	75.0%	9.0%

المنارة للاستشارات
www.manaraa.com

Table 14: Dimensional error from experiment #28 with 5 boxes

	Experiment #28		Occupancy Grid [10cm]			Reality [10cm]			Error [10cm]		
		Frame # picked	X	Y	Z	X	Y	Z	X	Y	Z
BOX Front Left	Test 28-01	30	7	4	7						
		50	7	5	7						
		70	7	4	7						
		Average	7.0	4.3	7.0	7.0	7.0	5.0	0.0	-2.7	2.0
	Test 28-02	30	7	7	7						
		50	7	6	6						
		70	7	3	7						
		Average	7.0	5.3	6.7	7.0	7.0	5.0	0.0	-1.7	1.7
	Test 28-03	30		7	7						
		50		9	7						
		70		9	7						
		Average		8.3	7.0		7.0	5.0		1.3	2.0
	Test 28-04	30		9	8						
		50		9	8						
		70		10	8						
		Average		9.3	8.0		7.0	5.0		2.3	3.0
BOX Front Right	Test 28-01	30	9	4	8						
		50	9	4	7						
		70	9	3	7						
		Average	9.0	3.7	7.3	7.0	7.0	5.0	2.0	-3.3	2.3
	Test 28-02	30	9	4	7						
		50	9	3	6						
		70	9	6	7						
		Average	9.0	4.3	6.7	7.0	7.0	5.0	2.0	-2.7	1.7
	Test 28-03	30	10	9	7						
		50	10	9	7						
		70	10	9	7						
		Average	10.0	9.0	7.0	7.0	7.0	5.0	3.0	2.0	2.0
	Test 28-04	30	10	9	8						
		50	10	9	8						
		70	10	10	8						
		Average	10.0	9.3	8.0	7.0	7.0	5.0	3.0	2.3	3.0
BOX Back Left	Test 28-01	30	6	4	6						
		50	6	5	6						
		70	6	4	6						
		Average	6.0	4.3	6.0	7.0	7.0	5.0	-1.0	-2.7	1.0
	Test 28-02	30	6	5	6						
		50	7	5	6						
		70	6	5	6						
		Average	6.3	5.0	6.0	7.0	7.0	5.0	-0.7	-2.0	1.0
	Test 28-03	30		5	7						
		50		5	6						
		70		4	6						
		Average		4.7	6.3		7.0	5.0		-2.3	1.3
	Test 28-04	30		5	6						
		50		5	6						
		70		3	6						
		Average		4.3	6.0		7.0	5.0		-2.7	1.0
BOX Back Middle	Test 28-01	30	8	3	5						
		50	8	4	5						
		70	8	3	5						
		Average	8.0	3.3	5.0	7.0	7.0	5.0	1.0	-3.7	0.0
	Test 28-02	30	8	6	5						
		50	8	4	5						
		70	8	4	5						
		Average	8.0	4.7	5.0	7.0	7.0	5.0	1.0	-2.3	0.0
	Test 28-03	30	9	7							
		50	9	5							
		70	9	5							
		Average	9.0	5.7		7.0	7.0		2.0	-1.3	
	Test 28-04	30	9	8							
		50	9	9							
		70	9	8							
		Average	9.0	8.3		7.0	7.0		2.0	1.3	
BOX Back Right	Test 28-01	30	8	4	5						
		50	8	3	5						
		70	8	3	6						
		Average	8.0	3.3	5.3	7.0	7.0	5.0	1.0	-3.7	0.3
	Test 28-02	30	9	4	6						
		50	8	3	6						
		70	9	4	6						
		Average	8.7	3.7	6.0	7.0	7.0	5.0	1.7	-3.3	1.0
	Test 28-03	30	9	5	5						
		50	9	6	5						
		70	9	6	6						
		Average	9.0	5.7	5.3	7.0	7.0	5.0	2.0	-1.3	0.3
	Test 28-04	30	9	5	7						
		50	9	5	7						
		70	9	5	6						
		Average	9.0	5.0	6.7	7.0	7.0	5.0	2.0	-2.0	1.7
Average			8.3	5.6	6.4	7.0	7.0	5.0	1.3	-1.4	1.4
Error	"+" means object in occupancy grid is bigger than in reality								18.8%	-20.2%	28.1%

Missing point data: Some points were occluded by another object

المنارة للاستشارات
www.manaraa.com

The dimensional accuracy of object was measured by comparing the Total Station measurement ("true" width, depth, height of an object) to the corresponding values generated by the occupancy grid. In experiment #22 and experiment #28 a total of 32 different measurements were compared. The averaged error rate is displayed in Table 12, Table 13, and Table 14 . In a method explained earlier, these tables summarize the results to individual experiments. In each experiment, three randomly selected frames were chosen and analyzed by the dimensional values of occupancy grid vs. Total Station measurement (Reality). The differences are noted as errors. As a result, Table 15 summarizes the errors in width (X), depth (Y), and height (Z). The maximum error is in the Z-axis of 12cm (or 21.3% taller than the original object). The object once modeled in the occupancy grid algorithm appears to be too tall. This finding is similar to the X-axis, where as in Y axis objects are modeled smaller. As explained before, the main reason can be line-of-sight since the part of an object in the shadow can not be seen in 3D (only 2½ D).

Table 15: Summary of dimensional accuracy

	Measurements	Error [m]			Error [%]		
		X	Y	Z	X	Y	Z
Experiment #22	12	0.06	0.12	0.09	15.9	40.7	9.9
Experiment #28	20	0.13	-0.14	0.14	18.8	-20.2	28.1
Average weighted	32	0.10	-0.04	0.12	17.7	2.6	21.3

An observation that is important to note is the fact that the prototype camera includes a lot of fluctuations during range measurements. In the occupancy grid that can be seen in experiment #28 (illustrated in Figure 47) the two background walls both appear to have a depth value of an average of 3-4 voxels in the Y-axis. Due to line-of-sight of the range sensing method, however, the depth value should not be recognizable or very minimal. Subsequently a more accurate distance measurement cannot be given of the 3D location of the walls at a distance of 6.1m from the sensor.

المنارة للاستشارات
www.manaraa.com

5.4 EXPERIMENTS TO DETERMINE DIRECTION AND VELOCITY

The indoor experimental setup shown in Figure 41 was used to determine how accurately objects that were mounted on the cart traveled along a determined linear path (cart guided on a rail). It also determined the difference of the experienced velocity of the algorithm compared to the observed real velocity. To this point a total of 24 experiments were analyzed that used a box, a pipe, and a human to demonstrate the successful working principle of the algorithm which is, to detect and track moving objects. Detecting means to recognize (not identify) an object and tracking means determining the objects' travel path in space. Each of these objects was used in eight experiments, propelled at three different angles to the range sensor, at two different velocities (slow and fast).

As explained previously, the cart carrying the box and pipe, and the human did not have an inertial velocity measurement that could be compared to the measured value after the sensor's range data was processed. Thus, a digital video camcorder shot the field-of-view of the range sensor simultaneously to each experimental run to measure and post-process the time elapsed between entry and exit of the field-of-view. A Total Station delivered the actual travel distance value of the object within the FOV, which in summary, allowed calculating the original speed. Results and analyses are presented next.

In Table 16 the list of experiments to determine the errors in directional measurements is presented. Four different experiments were conducted: Human climbing a ladder, human walking on ground, cart with box traveling on ground at slow and fast speed, and cart with pipe traveling at slow and fast speed. The experiment with all four objects were conducted at three different angles with 0° as a horizontal on-ground movement or in the direction of the negative X-axis to the sensor's FOV (left to right), 90° as in positive Y-axis (on-ground movement away from the sensor), 30° as in between

المنارة للاستشارات
www.manaraa.com

(on-ground movement to the far left), and 72.5° (in vertical movement at the angle of ladder).

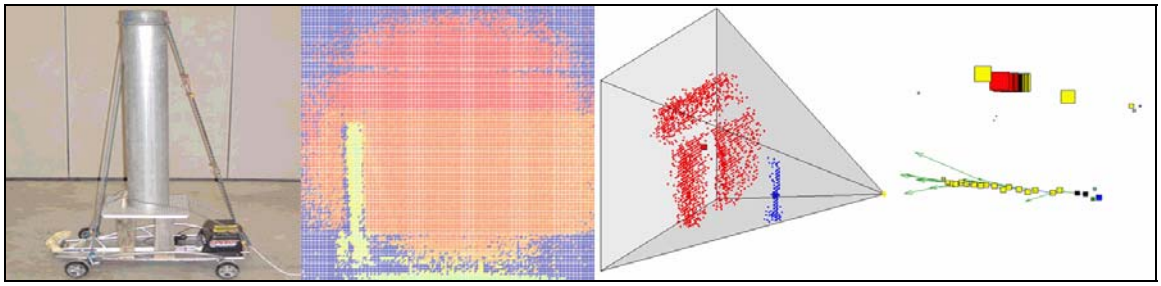


Figure 48: Indoor experiment – Moving pipe

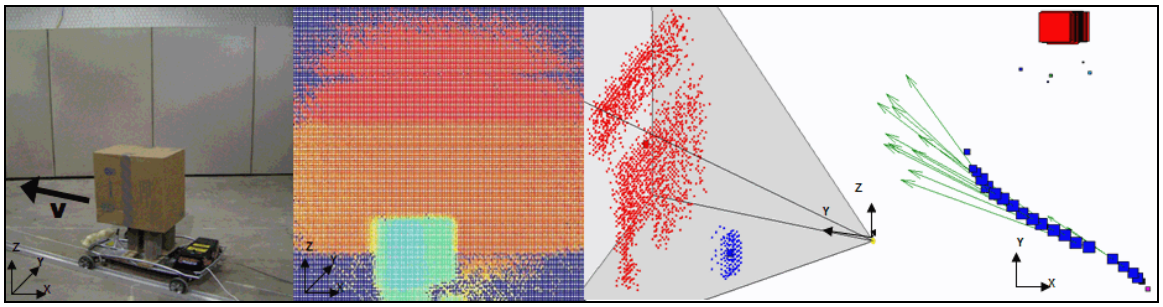


Figure 49: Indoor experiment - Moving box

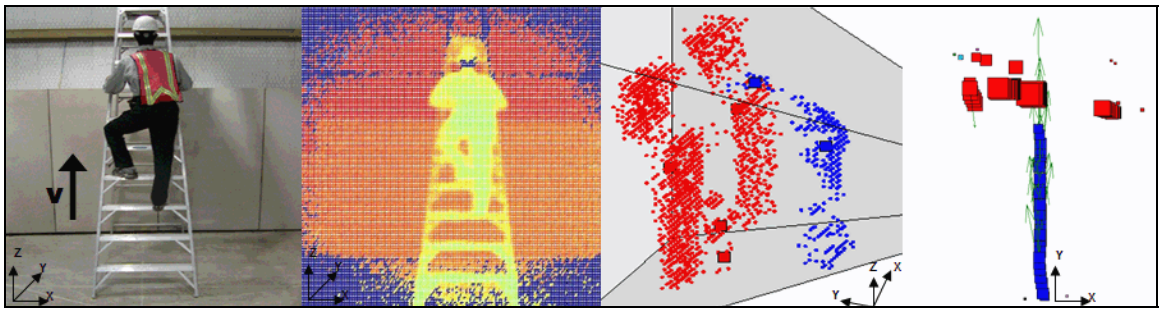


Figure 50: Indoor experiment – Human climbing a ladder

Three objects were used in indoor experiments, such as a pipe & cart in Figure 48, box & cart in Figure 49, and human & ladder in Figure 50. In these figures the first image shows the original experimental setup with object and direction. To the right are

المنارة للاستشارات
www.manaraa.com

illustrated: One range frame image of the sensor's FOV, the same frame in its processed form indicating the volume centroid to objects with a small square, and to the very right is presented the plot of all consecutive center of gravities to found objects in this particular measurement. These figures demonstrate that objects were detected as well as tracked.

Table 16: Differences in directional measurements

Experiment #	Object	Angle			Exp. used for direction		
		Experienced	Original	Difference	X	Y	Z
21-01-01	Human/Ladder	-9.8	0.0	-9.8	yes		
21-01-01	Human/Ladder	68.4	72.5	4.1			yes
21-01-02	Human/Ladder	-35.8	30.0	-5.8	yes	yes	
21-01-02	Human/Ladder	69.3	72.5	3.2	yes		yes
21-01-03	Human/Ladder	-88.4	90.0	1.6		yes	
21-01-03	Human/Ladder	66.6	72.5	5.9			yes
21-01-04	Human/Ladder	-6.9	0.0	-6.9	yes		
21-01-04	Human/Ladder	62.1	72.5	10.4			yes
21-01-05	Human/Ladder	-25.1	30.0	4.9	yes	yes	
21-01-05	Human/Ladder	71.6	72.5	0.9			yes
21-01-06	Human/Ladder	-89.2	90.0	0.8		yes	
21-01-06	Human/Ladder	64.5	72.5	8.0			yes
21-02-01	Human	-10.0	0.0	-10.0	yes		
21-02-02	Human	-25.9	30.0	4.1	yes	yes	
21-02-03	Human	-88.3	90.0	1.7		yes	
21-02-04	Human	-4.1	0.0	-4.1	yes		
21-02-05	Human	-29.2	30.0	0.8	yes	yes	
21-02-06	Human	-89.6	90.0	0.4		yes	
22-01-01	Box slow	-5.4	0.0	-5.4	yes		
22-01-03	Box fast	-8.1	0.0	-8.1	yes		
22-01-05	Box slow	-35.0	30.0	-5.0	yes	yes	
22-01-07	Box fast	-33.7	30.0	-3.7	yes	yes	
22-01-09	Box slow	-88.5	90.0	1.5		yes	
22-01-11	Box fast	-87.7	90.0	2.3		yes	
22-02-01	Pipe slow	-5.4	0.0	-5.4	yes		
22-02-03	Pipe fast	-8.1	0.0	-8.1	yes		
22-02-05	Pipe slow	-30.0	30.0	0.0	yes	yes	
22-02-07	Pipe fast	-28.8	30.0	1.2	yes	yes	
22-02-09	Pipe slow	-89.5	90.0	0.5		yes	
22-02-11	Pipe fast	-89.3	90.0	0.7		yes	
Average error in X [°]					0.35		
Average error in Y [°]						1.71	
Average error in Z [°]							4.81

The known direction of the object movement (0°, 30°, 90°, and 72.5°) needed to be compared to the directional value measured by the algorithm. The experienced values

المنارة للاستشارات
www.manaraa.com

were derived from the algorithm output as it can be seen and as already explained in Figure 33. The designed (true) values of the directional movement are then directly compared to the experienced averaged directional angle values. Only those experiments that had a directional component in the expected object movement direction were used to determine directional errors for that axis (as seen in Table 16). In this table, each experiment (identification number is in the left column) contains one object. The object itself (e.g. pipe, box, human) is described in the next column. The next columns list the object's motion path in reality, in the experiment, and the difference in directional movement. Only those experiments which had a component in the axis of the movement were used to determine the error. For example, in experiment #22-02-09, a pipe that was moving at slow speed on a designed travel direction of 90 degrees achieved only abs(-89.5)=-89.5 degree (thus was 0.5 off the original designed path), was used only in the y-direction of the error assessment. The overall error of experiments was the average of all individual experiments.

As a result of the experiments, objects that traveled in negative x-direction on average were 0.35 degrees off the original travel path and thus slowly gained went off in the Y-axis. Objects were designed to travel directly away from sensor slowly tended to move to the right on average by 1.71 degrees. Objects that had a vertical velocity achieved a lower than expected angle of about on average 4.81 degrees. 75 percent of all objects were within less than ±5 degrees off the original path. As another result or recommendation, the experimental setup can be significantly improved to reduce systematic error sources (e.g. inertial velocity measurement, rigid guidance of objects, better sensor calibration or prototype, etc.).

A similar strategy was used to determine the accuracy of object velocity values comparing the original objects speeds to the speed experienced after range data

المنارة للاستشارات
www.manaraa.com

processing. The original speed was measured upon two values: (1) using a video camcorder to count the time elapsed between the entry and exit of the object in the sensor's field-of-view, and (2) using a Total Station measurement to obtain the travel distance of the object within the sensor's field-of-view. The experienced speed was calculated upon three known values: (1) Appearance and disappearance of the VC of the target object allowed determining the entry and exit frame during the frame grabbing sequence by subtracting the frame identification number, (2) using the fixed frame grabbing frequency of 15.2 Hz during the experiments to determine elapsed time, and (3) travel distance on ground in the sensor's field-of-view. Due to a fixed voxel size of 0.1m the algorithm actually converted the time stamp that was acquired in the data acquisition process to each frame and related it to the volume centroid of each object. This repetitive procedure allowed to compare one frame to next frame immediately and to measure the velocity of the same volume centroid until the object disappeared. Taking the average led to the experienced velocity value referred here.

Table 17: Differences in velocity measurements

Speed [m/s]	X	Y	Z
Reality	1.350	1.140	0.400
Algorithm	1.400	1.202	0.404
Difference	0.050	0.062	0.004
Error [%]	3.7	5.4	1.1

In Table 17 the results of velocity differences to each axis is presented. Objects in x-direction are 3.7% faster, in Y-direction 5.4% faster and in Z-direction 1.1% faster than expected. Overall, objects appearing in the real-time 3D model appeared to be faster. This research did not intend to find (statistical) evidence whether these values are significant and/or can be reduced by correcting, calibrating, or changing the developed algorithm and its parameters. In the discussion of the error sources (following chapter) some improvements to the environmental setup as well as to the algorithm are suggested.

المنارة للاستشارات
www.manaraa.com

In summary, all objects traveled closely to their original path and speed and the developed three-dimensional occupancy grid algorithm achieved satisfactory results to demonstrate in this initial research effort that detection and tracking of resources such as for example humans in the travel path of a heavy equipment machine is possible.

5.5 OUTDOOR EXPERIMENTS WITH STATIC 3D VIDEO RANGE CAMERA

Several outdoor experiments had the goal to demonstrate the successful working principle of the algorithm. The following figures to 4 experiments show typical detection and tracking results that were found.

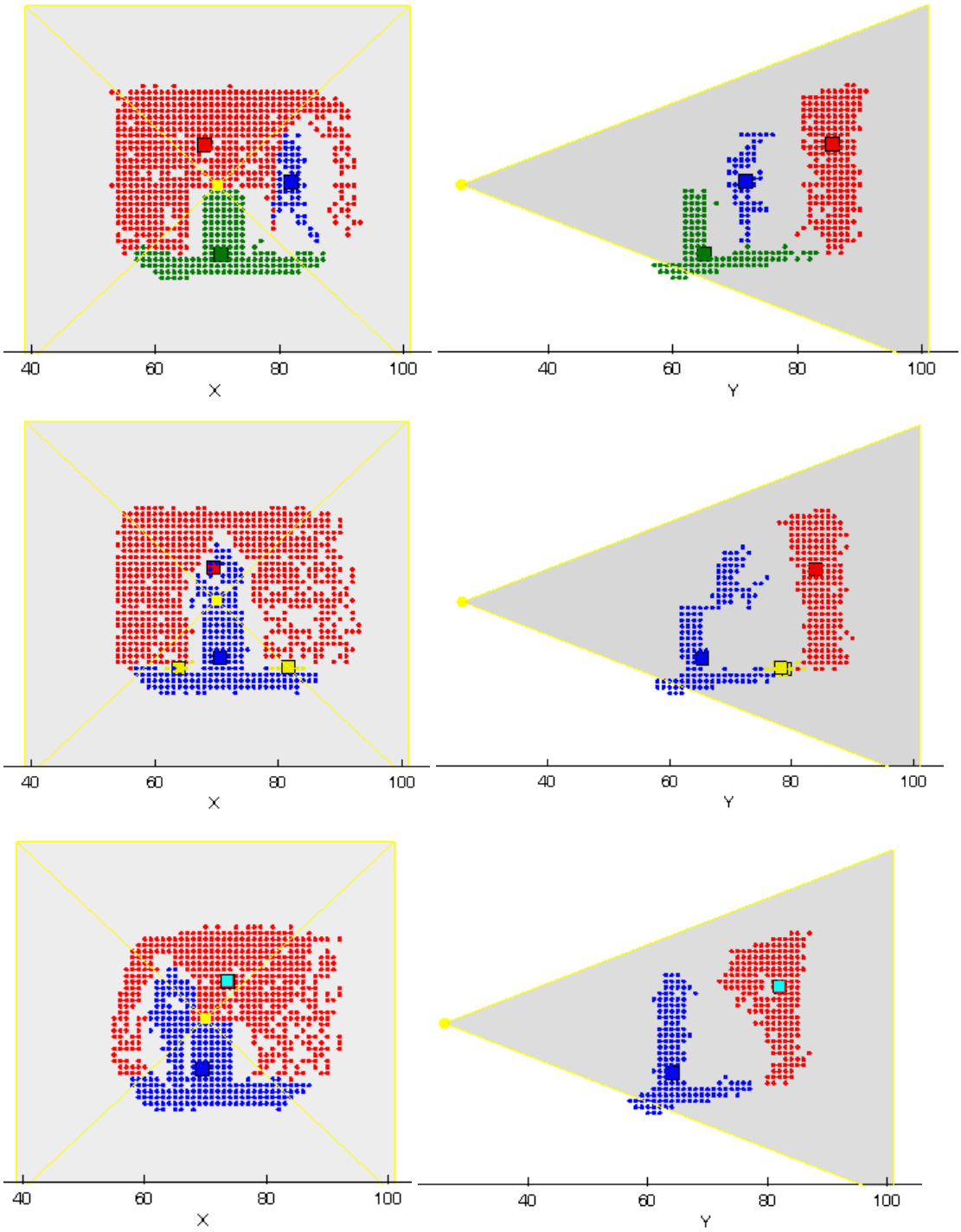
As expected from the prototype sensor, the biggest impact in the outdoor environment was the background noise. The noise level made it impossible to work in direct sunlight. All experiments explained next were conducted in cloudy weather or night conditions. This once again shows the feasibility to operate at darkness.

In Figure 51 the front view and its corresponding elevation view of 4 processed range frames are illustrated. The first processed range frame that is displayed in this figure shows a person entering the FOV of the 3D video range camera. The oil barrel and the ground are in one cluster since they have physical contact. The background wall is separated from all existing clusters since the angle the video range camera generates only a few filled voxels at larger distances. This principle separates these clusters due to the minimum distance threshold that is predefined before the algorithm processes the range information.

In the next frame sequence the person has reached a position behind the barrel that is very close or touching it. The cluster of person, barrel, and ground are merged due to physical contact.

The following two frame sequences show the person taking a 90 degree directional change towards the sensor. Clusters again are merged.

المنارة للاستشارات
www.manaraa.com



المنارة للاستشارات
www.manaraa.com

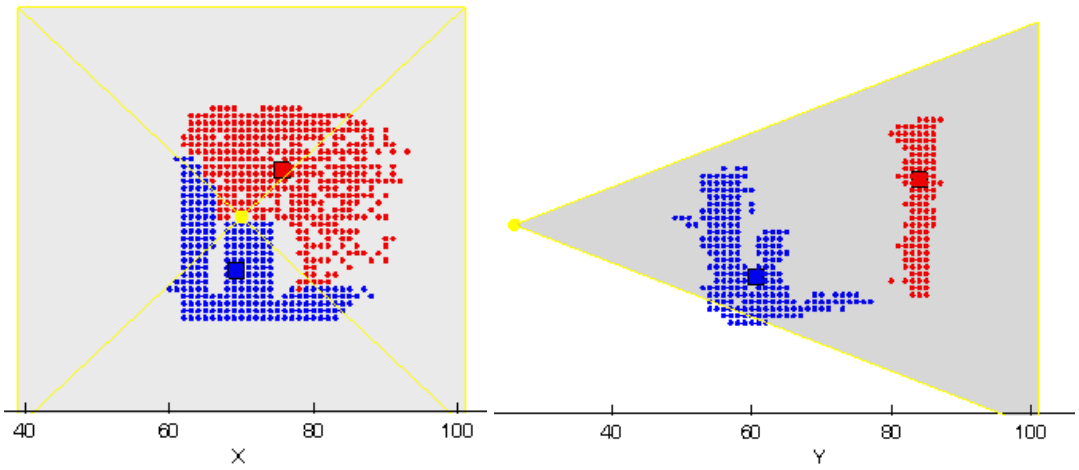


Figure 51: Outdoor experiment #24-04-03: Human walking around barrel

As a result of this experiment, detection and tracking of objects used in this research (size, speed, color etc.) was successfully. Future work is in the improvement of the algorithm, e.g. it needs to separate objects that are standing on ground level. One method to solve this problem is to calibrate the ground level before the range camera takes range frames. The current limitation of the prototype status of the camera did not allow to prove these issues.

In Figure 52 two humans walked adjacent two each other. While the front worker and his wheel barrow are merged within one cluster, the range data for the worker in the back was processed in its own cluster. Range data to an oil barrel between the travel paths of both persons were clustered to one object. The background wall appears in its own cluster. Due to the minimal physical distance between front person and smaller metal scaffolding structures in the front, both clusters are merged to one object.

المنارة للاستشارات
www.manaraa.com

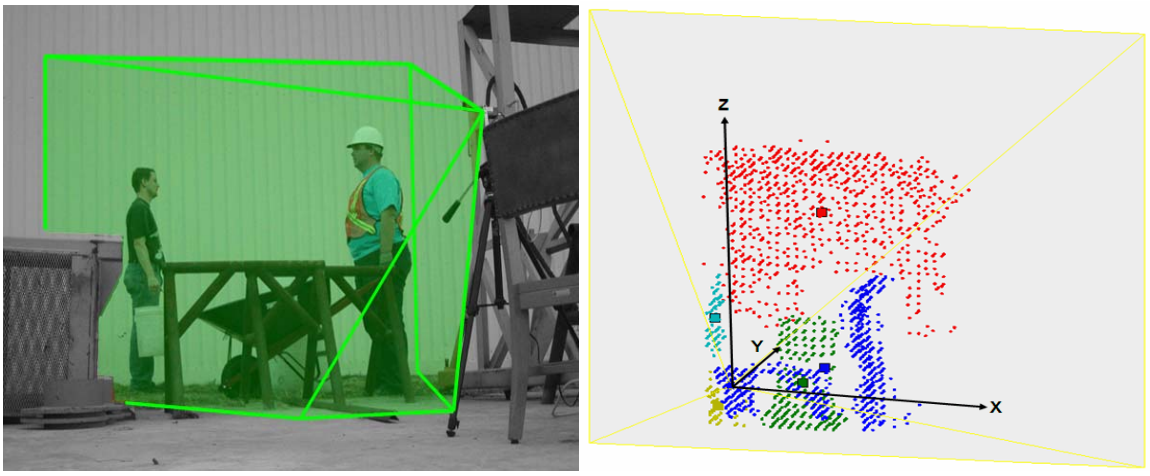


Figure 52: Outdoor experiment #24-03-02: Two humans and object clutter

In final experiments, heavy equipment machinery was used. The experimental setup that can be seen in Figure 43 was used to model the travel path of a skid steer loader.

In Figure 53 a skid steer loader travels at very slow speed through the FOV of the 3D video range camera. In the left image the original scene can be seen. Thereafter is one original range frame and the same processed range frame. The image on the far right shows the plane view of all processed range frames superimposed. Since the bobcat occluded the background and due to ambient light many small clusters were generated with this prototype sensor (this particular sensor version did not offer active background light suppression).

المنارة للاستشارات
www.manaraa.com

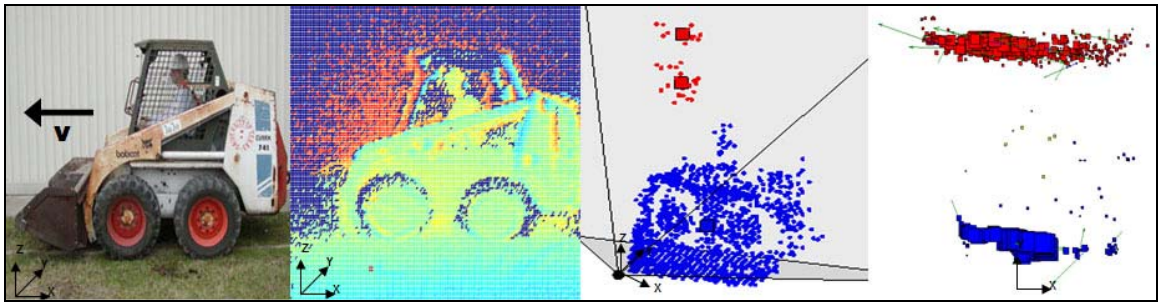


Figure 53: Outdoor experiment #18 with a moving skid steer loader

Since the FOV of the prototype was limited in distance and angular FOV, the machine almost completely filled the FOV of the 3D video range camera. The algorithm and its 3D modeling feature, however, successfully modeled the object. Black colored objects, such as tires did not return any valuable range information that can be used for range image processing. The algorithm was not able to model such scene features which were missing subsequent range data.

5.6 EXPERIMENTS WITH MOVING 3D VIDEO RANGE CAMERA

5.6.1 Indoor Experiment

Several indoor experiments were conducted to verify that the algorithm works when the 3D video range camera is in moving status and the objects are either static or moving. Due to the limitation of the experimental setup, the horizontal gantry provided a distance of 40cm to precisely control the move of the sensor, only a very limited number of experiments were conducted. The sensor was put at speed of 5cm/s, 10cm/s, and 20cm/s. In Figure 54 the 3D, front, plane, and elevation view of the occupancy grid can be seen. The start and end field-of-view of the sensor and all processed occupancy grid frames that were generated in between are displayed in the figure. While moving the sensor in this particular experiment at a speed of 5cm/s in the actual scene five objects were present: fascia board, three cardboard boxes of different sizes and location, and the

المنارة للاستشارات
www.manaraa.com

background wall. Due to the close proximity of the box to the sensor the box cluster was most times merged with the cluster of the fascia board. As a result, the detection of static objects in an indoor environment was successful, but the experimental setup can be improved allowing for longer sensor movements.

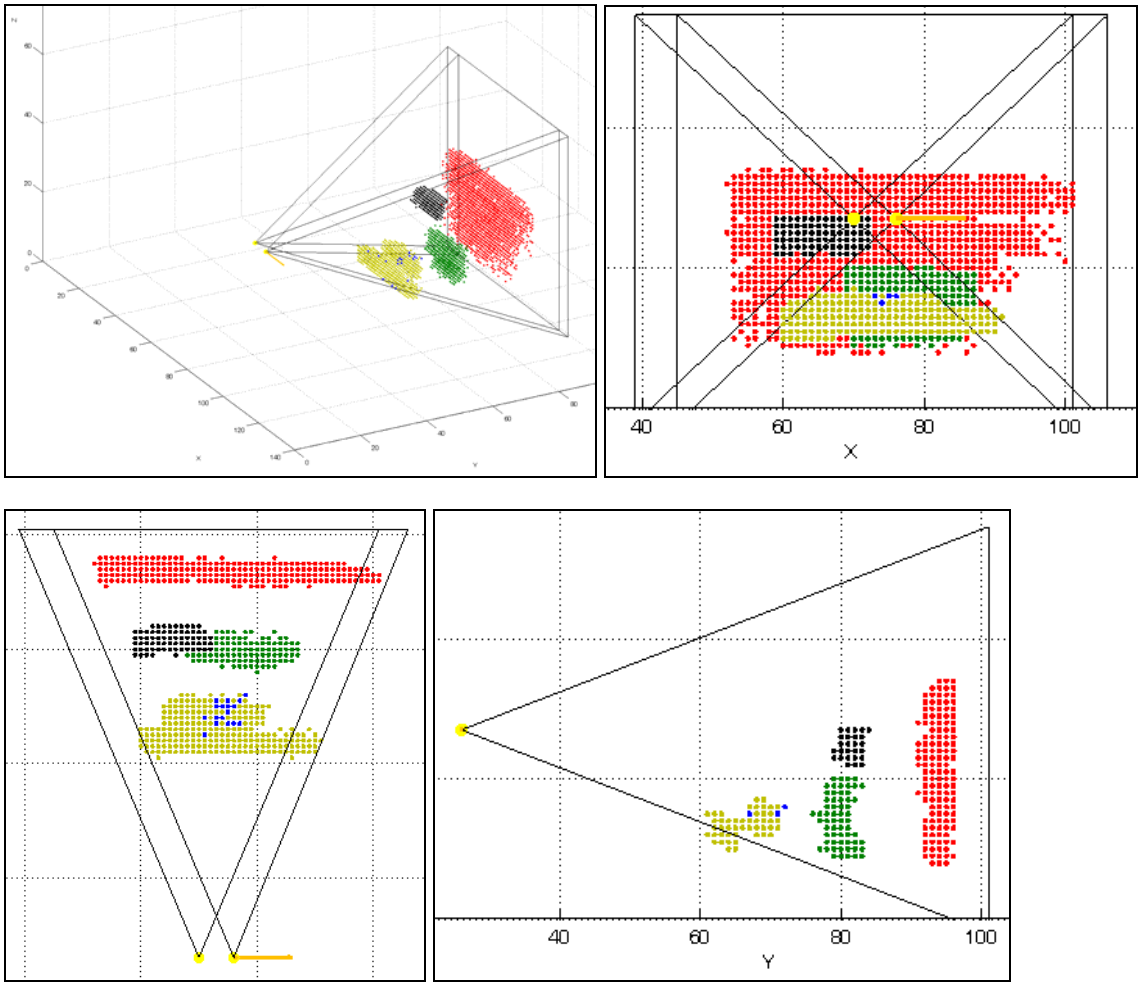


Figure 54: Start and end position of moving 3D video range camera capturing static objects

المنارة للاستشارات
www.manaraa.com

5.6.2 Outdoor Experiment

The 3D video camera was mounted on the skid steer loader and moved towards the laboratory building wall as seen in Figure 55. Perpendicular to the machine movement a person walked through the sensor's FOV.

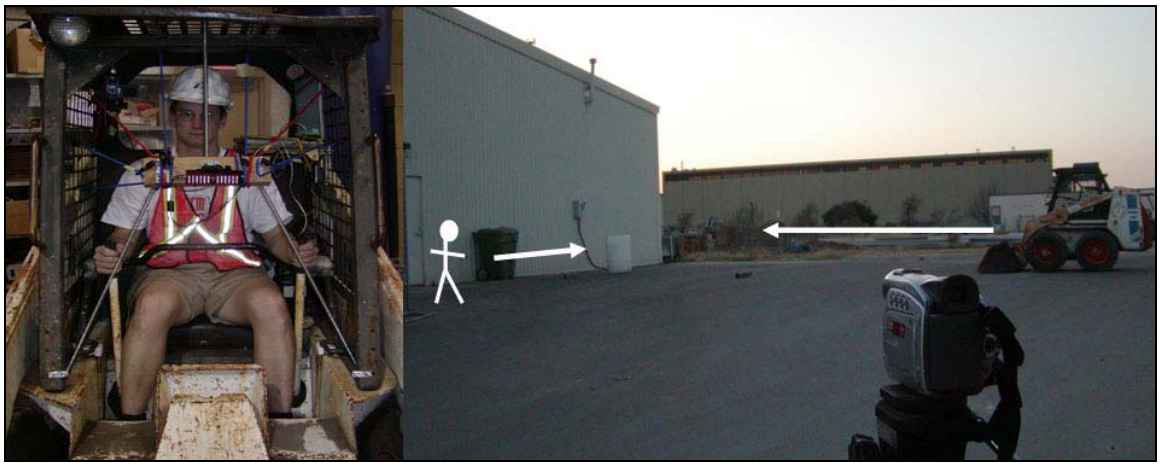


Figure 55: Sensor mounting on bobcat and outdoor experimental setup

The goal to detect the person in the occupancy grid was successful, however, to represent it in its own cluster (separated from the cluster that includes the ground level) was unsuccessful. In Figure 56 the person is highlighted with a circle.

المنارة للاستشارات
www.manaraa.com

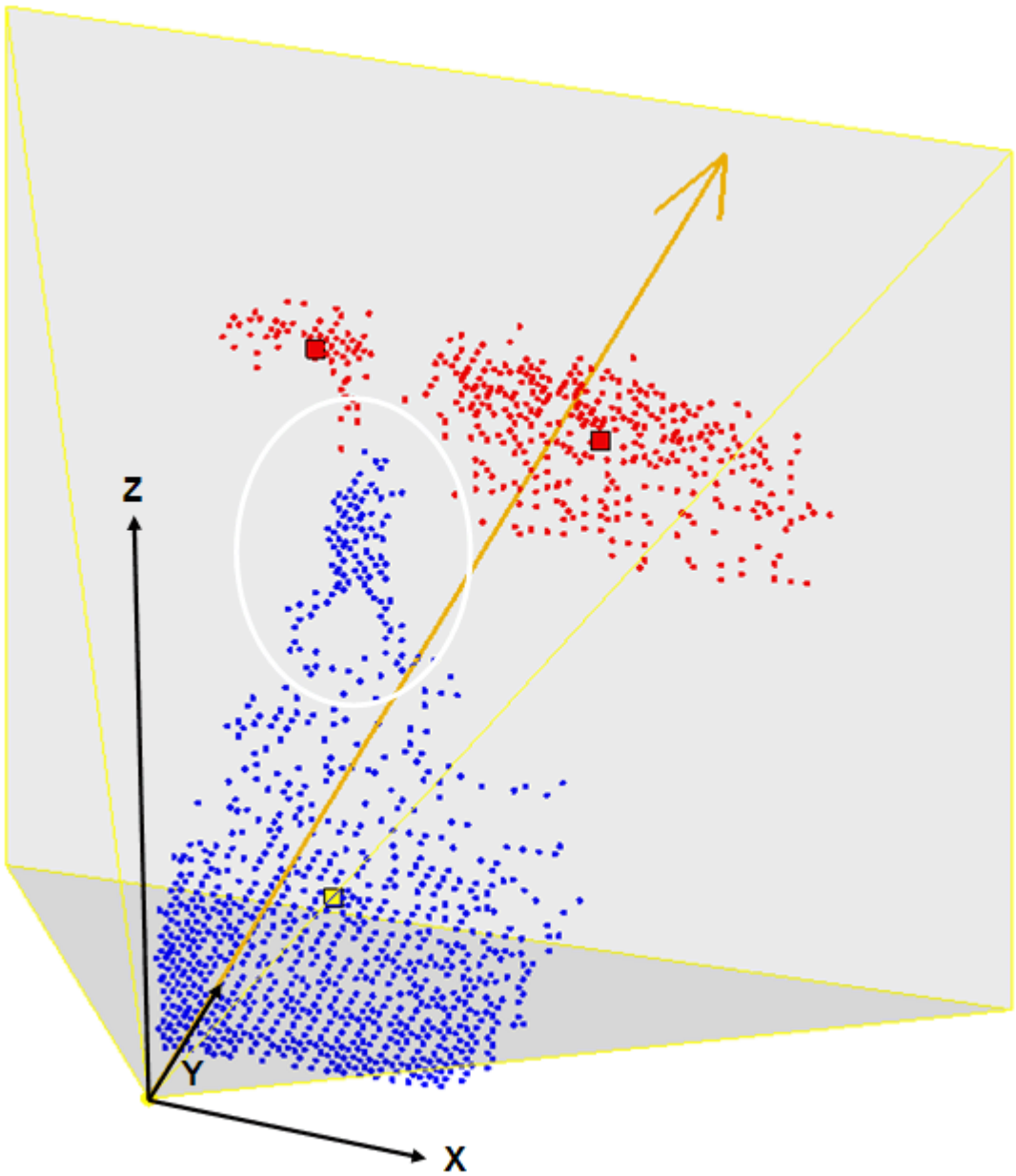


Figure 56: Processed range frame from moving sensor at a velocity of 2m/s

5.7 Results and Analyses of Experiments

The experiments and results demonstrated that objects appearing in the field-of-view of a range sensor can be detected and tracked. The developed technology has

المنارة للاستشارات
www.manaraa.com

potential to be applied in safety applications for heavy construction equipment operation to assist the equipment operator to prevent accidents, e.g. hitting co-workers close by to the machinery. Latest sensor developments have surpassed the non-ambiguous sensing range of 7.5m and extended it to up to 150m (see Table 3).

This algorithm was targeted towards rapid modeling of three-dimensional environments and objects for safety in construction, however, each application has its own specific requirements of how to take best advantage of range data. The resolution or number of voxels remaining in a scene, for instance, is a key for successful segmentation and clustering. Finely-sampled data can support creation of surfaces and rendered to create easily understood photo-realistic or enhanced images, e.g. range data for comparison of as-planned vs. as-built data. Coarsely-sampled data can degrade an image to surface normals and intensity, or simply clusters that bound their voxels and still can be tracked.

5.7.1 Time to process data

The information of one frame was processed in between 8.5 and 0.4 seconds with a mean of 5.4 seconds using a Matlab™ environment.

Programming the algorithm in Matlab™, known not to offer the fastest data processing for large data sets, can be improved by C++ coding at considerably lower computational cost. Running the algorithm on a Linux based environment is estimated to reduce the processing speed of single frames to tens of microseconds per frame.

5.7.2 Discussion of Errors

In the section to preliminary findings already many error sources were discussed. This section focuses on errors that have their origin the developed experimental

المنارة للاستشارات
www.manaraa.com

environment and which error sources have influence of limiting high accuracy range measurements.

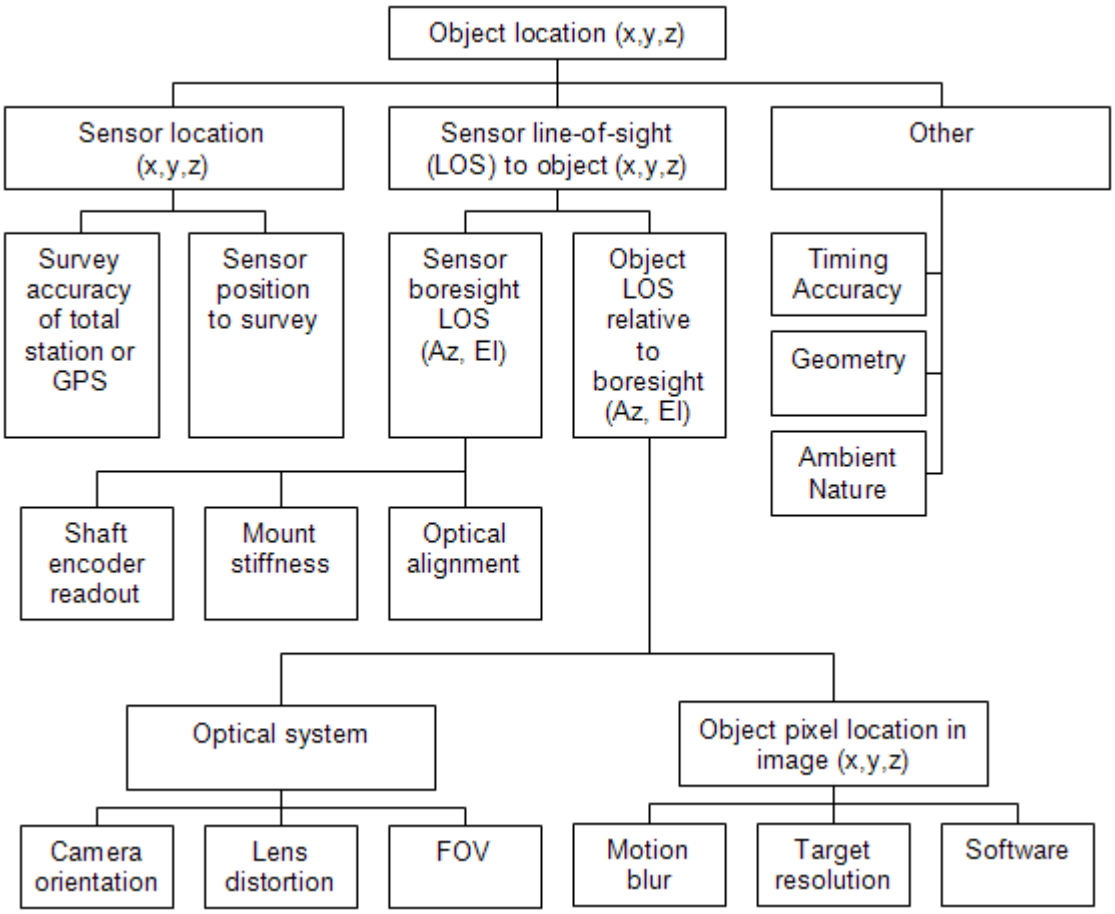


Figure 57: Sources of errors defining the accuracy of detection and tracking

Figure 57 displays sources of errors in an experimental setup based on optical range measurements (Roth, 2006). As stated in the assumption previously, only some are collected and discussed in this research. The main error relates to the accuracy of objects in the 3D model:

- Is the object positioned in the 3D model at the same location as it is in real space?
- Are the object dimensions accurate to the real dimensions?

المنارة للاستشارات
www.manaraa.com

- If an object is moving:
- Is the direction of the movement similar to the one in reality?
- Is the velocity the same as in reality?

Answers to these questions need to include a discussion what error sources influence the accuracy:

- *Absolute and relative errors:*

Accuracy of measurement setup depending on absolute and relative errors: Per definition absolute errors have common magnitude and direction. Relative errors have random magnitude and direction (Sanders-Reed, 2006).

Absolute error is necessary when computing absolute position data (e.g. for comparison with other systems). The absolute error in general depends on all error sources. The relative error is used when measuring the distance between two items, either two different items in the same scene (like two boxes from each other), or the same item from frame to frame (like in a range frame data acquisition). The relative error depends only on those quantities which change over time. For fixed cameras, errors in camera location and boresight line-of-sight (LOS) provide a constant bias to the absolute position, but do not affect the relative position. Once the camera is moving or tilting at different directions or angles, however, the error receives an additional absolute error value.

- *Errors in experiments:*

Stochastic errors are random events that produce error measurements, e.g. different background light source can influence the power of light return on the sensor. The developed algorithm was able to filter out stochastic errors.

المنارة للاستشارات
www.manaraa.com

Systematic errors are errors that can be defined as the "mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus a true value of the measurand". Standard uncertainty: is the "uncertainty of the result of a measurement expressed as a standard deviation". The linearity is defined as "the maximum deviation of any points from a straight line drawn a a "best fit" through the point cloud". Systematic errors can be solved via algorithmic corrections.

- *Errors in developed algorithm:*

According to the research objectives and given this laboratory setup the question was how accurately the object(s) position can be determined. In the measurement that contains object and camera the basic geometrical measurements such as x, y, and z, the angles, the field-of-view and number of pixels to and from the camera are measured with a Total Station. It is assumed that the Total Station values are true values (although the Total Station has a measurement itself). The sensor FOV and pixel size, and its range values from several frames were compared to the Total Station values. Once the measurement error σ has been found the correction (or calibration) values can be determined and stored in a table or formula. This finally allows to compute the object position (x,y,z) and error (σ_{target}).

المنارة للاستشارات
www.manaraa.com

Chapter 6: Potential Application Areas

The following overview presents important needs and target applications in object identification and tracking, transportation, safety, asset management, and construction. It is expected that there are many more areas which may benefit from this approach. Envisaged areas in construction of important applications are not limited to, but mainly are focused in security, target detection and tracking, engineering work, and work zone safety:

6.1 Target Identification and Tracking

Machine navigation on construction sites is often difficult as the environment is characterized by unknown location of equipment, materials, structures and moving objects, like humans or other machine operations. A significant improvement can be made if such unstructured or structured scenes can be made more understandable for the path planning of machine navigation. This requires to (a) analyze a scene rapidly to know the physical location of objects, e.g. a sand pile or parts of formwork, (b) identify them, (c) track moving targets and static objects in the motion path, and (c) store the information in a workspace model to (d) guide construction equipment (semi-) autonomously through Fobstacles.

6.2 Assisted Site Construction

Mechanical and construction work can be made faster and cheaper by applying new practices and techniques to the field. Major improvements are considered to be in (a) safety, e.g. trench cutting and earth movement work, (b) picking and placing with cranes materials such as pipes and spools, (c) rapid design modeling, e.g. pipe fitting of point data in 3D model, (d) surveillance and security, (e) analysis of site layout, i.e. to measure

المنارة للاستشارات
www.manaraa.com

and verify the quantity of materials or the direction and depth of trenches for pipe-laying, (f) 3D and 4D construction applications, e.g. to simulate construction, planning, documentation, and inspection analyses for payment purposes, (g) hazardous environments where humans have restricted access and 3D information is needed to proceed with the work, (h) coordinate issues to better design of retrofit projects and to help lower the time, effort and cost of process plant retrofit engineering and construction projects by helping engineers take advantage of the accuracy and completeness of point clouds (Kim et al., 2004). Such data integration translates into less construction rework due to interferences and fit-up problems and helps eliminate costly field fabrication.

6.3 WORK ZONE SAFETY

The construction industry continues to be one of the largest drivers of the U.S. economy, but it also offers one of the deadliest, most dangerous and unprotected work environments for workforce, equipment, and materials (CPWR, 2002). In 2005, the industry employed close to seven million construction workers, or 5.2 percent of the U.S. workforce, and the Bureau of Labor Statistics reported that it experienced a total of 1,224 fatalities. Construction accounted for almost 25% of all workforce fatalities and has a fatality rate (11.9% per 100,000 workers) three times higher than the average of all other industries (BLS, 2005).

In 2004, 36% of all construction fatalities occurred in “falls”, 23% were related to “transportation incidents”, 22% had “contact to objects and equipment”, 14% were due to “exposure to harmful substances or environments”, and 5% occurred due to other factors. The primary cause of death in 13% of all construction fatalities was related to infrastructure equipment, such as excavating, loading, and road grading machinery (BLS, 2005). In one past study, most equipment-human or equipment-object accidents were found to have resulted from missing safety features on operating heavy equipment

المنارة للاستشارات
www.manaraa.com

(OSHA, 1990). Installing new devices on machinery can make operations safer and extend the operator's limited vision and awareness (Kim et al., 2004).

The Construction Industry Institute (CII) estimated in 1993 that the direct and indirect cost of construction injuries in the U.S. was approximately $17 billion per year, an amount which has climbed since then (CII, 1993). Furthermore, CII's study demonstrated that initial investments in safety prevention like use of best practices, safety devices, and training pays off through improved project performance (CII, 2004). Indirect costs such as costs related to rescuing and attending an injured employee, productivity and quality lost while getting a replacement worker on site and up to speed, and time spent on filling out accident reports, are not covered by insurance and can account for up to four to six times the direct cost. (ENR, 2003 and OSHA 2005a and 2005b).

Good jobsite safety can keep up morale and workmanship. A study by the U.S. Department of Labor reports increasing demand for skilled construction workers, but retirements of industry veterans and less motivation among young Americans to join construction make it difficult for the industry to hire and retain skilled and safety-trained personnel (ENR, 2000). Statistical employment data indicates that the fast growth in the number of foreign-born workers may be related to the increase of more than twice the recent historical average in recorded fatalities in 2005 (BLS, 2005). As a result, extensive training, best management practices, and protective equipment and technology will be needed to improve safety and to help comply with existing Occupational Safety and Health Administration (OSHA) standards or rules of the National Institute for Occupational Safety and Health (NIOSH) (OSHA, 1990 and NIOSH, 2005).

6.4 ASSET MANAGEMENT

Another important field is asset management, i.e. in transportation. Determining the status quo of existing federal, state or local transportation and construction

المنارة للاستشارات
www.manaraa.com

infrastructure and resources becomes more important in aging infrastructure. Almost 2% of the construction budget is assigned to manual quality control and tracking of work package completion, including operations involving earth moving and bulk materials handling (Cheok et al, 1999). Thus operating and managing transportation systems are turning to become more important. Automated and rapidly related information will benefit in cost and time of (a) as-built and before and after drawings for 3D/4D CAD modeling of scenes to allow user a rapid access for accurate spatial information planning of future tasks, such as modifications of roadway, bridge and building design and rehabilitation, (b) condition assessment to determine the status of state of materials or objects, e.g. like asphalt or concrete cracks in highway or runway lanes, (c) quality control to repair and maintain at the critical point and to save budgets, e.g. rebar inspections, correct position of anchor bolts, and inspections of rutting, (d) interior modeling and architectural preservation for inventory of existing historical (infrastructure) structures, and (e) resource allocation for more strategic, performance based approaches to manage assets. Using accurate sensing systems can improve and preserve maintenance of federal and state wide transportation departments such as the facility management of transportation related buildings, roads or other infrastructural resources.

6.5 INTELLIGENT TRANSPORTATION SYSTEMS (ITS)

In addition to static tasks, other target areas for range sensors are intelligent transportation systems which require sensing systems able to communicate between vehicles and autonomously or semi-autonomously navigate vehicles without hitting any obstacles or able to count vehicles to regulate the traffic flow. Sensor applications in intelligent transportation systems become increasingly important, e.g. to navigate traffic faster and without accidents. Consequently, potential applications of range sensor are (a)

المنارة للاستشارات
www.manaraa.com

obstacle avoidance including humans, vehicles, and objects by identifying range and velocity vectors, e.g. forensic analysis of range data captured in black boxes during vehicle accidents, (b) semi-autonomous and autonomous navigation, e.g. on highways, to keep minimum distance between vehicles, as well in autonomous navigation of complex process like overtaking where many different real-time range measurements are needed, (c) day and night traffic monitoring, control and counting by adding make, model and dimension, (d) automotive and aviation infrastructure environments, e.g. detect structural conditions, (e) survey in heavy traffic zones without positioning surveyors in the roadway and without closing the road, (f) build the simulation model for storm structures, (g) 3D models for accident investigation and documentation.

In transportation related construction enhanced quantity tracking, project control systems, and safety play an important role. The field of view of equipment operators is often limited by environmental factors such as surrounding activities, noise, illumination, precipitation, and dust interference. Thus, without the operator's ability of a clear sight of their surroundings, the control and guidance of automated or semi-automated vehicles is a formidable task. In certain cases, highway maintenance and project schedules routinely require less protected night work. Operators should exercise extreme care as safety features like hardly visible signals, vests or information of backside mounted TV cameras on equipment, if available at all, almost becomes meaningless under such conditions.

Automotive restraint systems have reached a high level of occupant protection. Belts and airbags are important safety systems that are able to react once the impact takes place. To further increase safety, range sensing and imaging may help to acquire and process information before and much earlier dangerous situations occur. Preventive active action can prevent from accidents happening through object detection, classification.

المنارة للاستشارات
www.manaraa.com

6.6 SUMMARY OF APPLICATION AREAS

The importance of 3D vision is rising rapidly and the number of applications is increasing steadily. Table 18 lists some more applications for potential research topics.

Table 18: List of potential application areas for 3D range sensing

Industrial	Transportation and Automotive
• Safety • Work zone and machine protection • Safeguarding machine operation • Security and surveillance • Object recognition • Object measurement • 3D map generation • Automated production • Quality inspection, assurance, and control • Fixed and autonomous robots • Mobile systems • Picking and placing • Optimization in lift control (count waiting people)	• Navigation (Simultaneous localization and mapping - SLAM) • Pre-crash, obstacle and collision avoidance • Scene reconstruction for object localization • Pedestrian safety • Stop & go traffic • Parking aid • Emergency break • Smart airbag and occupant observation • Large scale geometrical modeling of space for modeling and simulation • Classification of cars in ongoing road traffic • Ice detection on aircraft or roads
Multimedia	**Medical**
• Computer games • PC pointing devices • Virtual reality and user interfaces • Video conferences • Film and television • Product presentation	• Biometrics, medical imaging and health care • Custom garment and shoe manufacturing • Geometric human modeling for gesture and activity recognition • Intensity and range data based tracking and detection of body parts
Military and Government	**AEC/FM and Oil Industry**
• Automatic target detection and missile guidance • Coherent LIDAR and LADAR for airport turbulences and windshear detection (Lockhead Martin, www.ctilidar.com) • Aerospace docking and 3D modeling	• As-planned vs. as-built modeling • Productivity and schedule reduction • Machine operation safety • Automated materials tracking and identification • Simulation and subwater 3D modeling

المنارة للاستشارات
www.manaraa.com

Chapter 7: Conclusions

This chapter reviews the achievement of the research objectives. The conclusions that grew out of this research are presented next, and the contributions of the research to infrastructure construction, maintenance, and transportation are discussed. Finally, some possibilities for future research are outlined, as well as some application areas are mentioned.

7.1 REVIEW OF OBJECTIVES

The main objectives of this research were to develop a fast, accurate, and reliable method of acquiring spatial information in a real-time, to immediately detect and track objects in the field-of-view of a static and moving three-dimensional video range camera, and to demonstrate the applicability of the method to the infrastructure construction, maintenance, and transportation discipline. According to the research hypothesis, the following major sub-objectives have been solved and achieved:

1) To find, understand, experiment with, and evaluate advantages and disadvantages of the working principle of emerging 3D video range sensing technology.

2) To develop a real-time, on-site experimental system for acquisition of spatial information using dense range data at fast range frame update rates exceeding with its characteristics the existing 3D laser scanning systems for detecting and tracking objects.

3) To program a computational algorithm based on three-dimensional occupancy grids that is capable of handling dense range point clouds to extract and trail static and moving objects from a static or moving sensor position.

4) To demonstrate and evaluate in indoor experiments and field tests the overall efficiency of the spatial-information acquisition system (hardware and software) through implementation in real-life scenario laboratory and construction-like environments.

المنارة للاستشارات
www.manaraa.com

7.2 CONTRIBUTIONS AND CONCLUSIONS

A key feature of the proposed system for on-site acquisition and handling of spatial information is its capability to update data on the positions, dimensions, directions and velocities of objects and people, which makes it a useful tool for thorough analysis of site layouts, identification of safety zones, utility tracking, general construction-management decision-making, traffic control on large sites, and enrichment of as-built drawings. As a result of this research, the developed algorithm allows automatic detection and tracking of clustered range volume pixel (voxels) and thus avoids manual tracking where an operator needs to specify objects or tracks. The developed tracking features chosen to be very helpful in three-dimensional range sensing using line-of-sight (LOS) to detect and track objects for obstacle avoidance in heavy equipment operation in a construction environment are based on the following defined characteristics: Position of objects, i.e. the volume centroid (VC), dimension of objects (x, y, and z), velocity of objects (speed), and direction of objects (angle).

This research shows that the approach of using emerging 3D range sensing technology and the developed data processing algorithm based on 3D occupancy grids was feasible to convert the acquired dense range data in real-time into 3D models that were then used to detect and track objects. Experimental results have shown promise as a source of improvement for modeling infrastructure operations and maintenance that would ultimately enhance safety and productivity in situations where heavy equipment is in use (e.g. potentially detect and track humans in the field-of-view of a 3D video range camera or travel path of heavy equipment in construction). Specifically this research has shown that:

1) The spatial-data acquisition method for generating models of objects from a dense set of data acquired by a three-dimensional video range camera is computationally

المنارة للاستشارات
www.manaraa.com

efficient, in real-time (better than 1Hz), and of sufficient versatility for application to the representation of construction sites.

2) The proposed spatial-information acquisition system can be employed in a variety of applications, e.g., in as-built 3D modeling for infrastructure operation and maintenance (with a small voxel size) or as object detection and tracking for enhancing safety in heavy equipment navigation control (bigger voxel size sufficient).

3) The review of technologies and use of technology as well as the developed algorithm has been demonstrated as an efficient tool, one that could play a significant role in accident prevention on an actual construction site.

7.3 RECOMMENDATIONS

This research has investigated the fundamental principles of optical range sensing and delivered a valid framework for real-time 3D modeling for safety in construction. Future research can lead into the combination of different range sensing methods that have been introduced in the beginning of this dissertation. Future short-term oriented research can address some of the following needs and applications:

1) Many applications require three-dimensional representation of the environment. To avoid time-consuming processing steps and in order to ensure the high quality of the specific task, the need in technology improvements includes: High distance accuracy per pixel, successful suppression of background light, high dynamic 3D imaging in terms of distance and light variations within the scenery, high speed demodulation, high depth and lateral resolution or in all three dimensions, large and dynamic measurements range, operate in harsh environments, e.g. bright sunlight, small pixel dimensions, e.g. Video Graphics Array (VGA), short and long range measurements at the same time, insensitivity to measurement conditions such as background light level, minimal power consumption, safe operation (comply with all eye-safety regulations),

المنارة للاستشارات
www.manaraa.com

ease of interface and programming, robustness, accuracy, small size, and low purchase price. Researchers in general state that future research on the hardware side can lead into the control of a light source that is modulated temporarily and spatially ("shuttered light pattern projector") and a custom time-of-flight camera at a specific triangulation angle.

2) Concentration on building smaller sensors that offer smaller systems at smaller costs. Faster processing leads to precise timing and higher resolution. A high optical fill factor improves the intensity image, reduces noise and allows measuring longer ranges.

3) A goal can be to make 3D data as easy to use as 2D data, track editing to fix broken tracks, include intensity data in the data processing analysis, and identify objects and name it (may allow to handle hazardous material differently).

4) A combination of technologies to overcome their individual limitations. In the particular application of safety for heavy equipment operation in construction, laser range finders using the sparse point cloud approach, for example, could provide 3D models permanent objects, such as traffic zones. A video 3D range camera can detect semi-stationary objects (moved once in a while). RFID tags can help in identifying objects, e.g. objects that potentially need special enlarged distances, such as power lines may require a larger distances than a concrete pipe. GIS can host all project related information and assist the project management, e.g. to chose the best construction method to complete a work task, such as the maximum allowable size of hoisting equipment to reach a confined space.

5) Create a rigid testbed (experimental setup) that improves the experimental environment and reduces systematic error sources. A rigid experimental environment can allow research to proceed in calibrating existing and emerging range sensing technology. Standardization of the technology, the data output, and data communication is another

المنارة للاستشارات
www.manaraa.com

important research feature to allow as many applications as possible to benefit from this approach.

6) Integration of range sensing with other modeling and simulation methods, e.g. a priori information from CAD, to build rapid 4D models to improve construction visualization before work tasks are executed, conventionally or automatically.

7) As this research has demonstrated, the development of an algorithm based on thresholding has limitations. The developed algorithm may only work in specific situations. A future research effort can be on the development of automated/adaptive thresholds.

8) Advances in clustering algorithms itself will allow a better (more accurate and detailed, with less random errors) representation of a scene at faster speeds. Various methods exist and research needs to investigate the potential of these approaches, such as prototype based clustering algorithms (fuzzy, mixture models, self-organizing maps), density-based clustering algorithms (grid based, subspace, kernel-based), graph-based clustering algorithms (sparse, minimum spanning tree, optimal partitioning of sparse similarities using METIS, chameleon with hierarchical clustering with dynamic modeling, shared nearest neighbor similarity, Jarvis-Patrick, SNN density), and scalable clustering algorithms (Tan et al., 2006).

9) Once the hardware to deliver raw range data and the algorithm to process the raw range data have been optimized, field application and test can be started. Future research can determine guidelines of how to best use the technology and software.

المنارة للاستشارات
www.manaraa.com

Appendix A – Technical Details to 3D Range Camera and Interface

Figure 58: Components of 3D camera

The 3D range camera consists of five components as seen in Figure 58:

- 3D range camera (48 LEDs, lens, frame grabber (sensor) and processing unit)
- Digital data cable assembly (USB 2.0)
- Power supply (100—240Volt, 50-60Hz, 1.5A, output 12Volts)
- Computer processing unit (Dell Inspiron I1510, Mobile Intel® Pentium® 4 3.06Ghz, 1.59Ghz, 512 MB RAM, 60MB hard drive, NVIDIA GeForve FX Go 5200)
- Data acquisition and camera control software interface

The manufacturer CSEM describes the SwissRanger 2 RevB with the following electrical characteristics (CSEM, 2005): The camera is based on 0.8 micron CMOS/CCD

المنارة للاستشارات
www.manaraa.com

technology A 160 by 124 pixels with the size of 39.2µm by 54.8µm based on "2-tap" lock-in pixel structure measurement distances after the TOF using the phase shift principle. The sensor is controlled like a 1-tap structure. The lens manufacturer is Kogaku Universe and the lens has an 8mm focal length with F/# 1.4. The filter is made by MK Photonics at CWL of 880nm and BW of 70nm. The filter diameter is 0.5inch. The illumination source, the LEDs, is produced by Vishay Semiconductor and operates at a wavelength of 870nm with a spectral bandwidth of 40nm. The beam divergence is ±22° with an average LED current of 70mA and the average optical power of one LED at 70mA of 15mW. The mean optical output power by continuous wave modulation is 720mW. After the start of the range camera it acquires tap0, tap90, tap180, and tap270 before it calculates and reports the readout to the computer (see Figure 59).

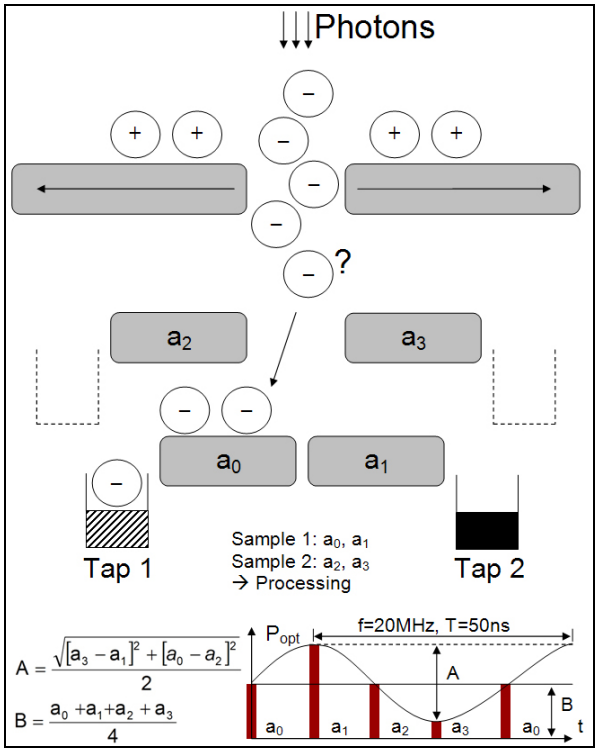


Figure 59: 4-tap structure on sensor (after Gut, 2004)

المنارة للاستشارات
www.manaraa.com

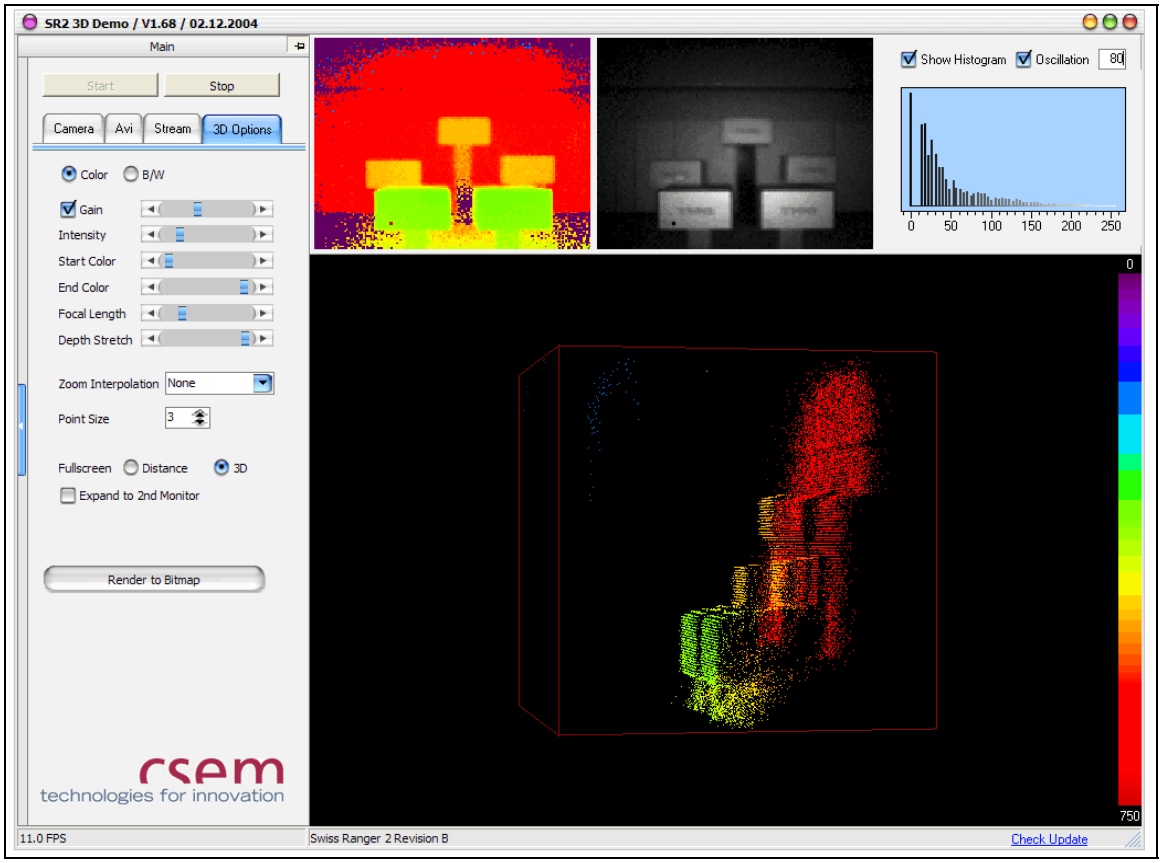
The APS readout allows every pixel to have its own amplifier stage to convert the signal charge into voltage. Every pixel can be addressed individually and in a nondestructive way. All pixels can be reset with the same voltage at the same time.

The button "Start" on the upper left corner of the software interface (see Figure 60a and b) allows the start of the 3D camera. Amplitude Threshold LSB and MSB [0..255], Intensity Saturation LSB and MSB [0..255], as well as Integration Time [0..255] can be set in the left column. For the experiments they were left on default value (50/0/0/36/50). Several register values can be set. The software interface displays several other information values: (1) Depth display window from the front perspective (upper left image), (2) intensity/amplitude window from the front perspective (upper middle image), (3) histogram as an optional output window (upper right image), (4) 3D view window of range information with oscillation option (large center image), (5) range frame acquisition rate (lower left corner), and (6) sensor name and option for new software updates (lower center and right corner). The distance window displays range information as a color coded pixel. Close and far colors can be selected. The intensity window displays the average of a pixel's four taps. The histogram window shows the distribution of amplitude over the entire pixel field. The 3D window is an OpenGL cube which can be configured.

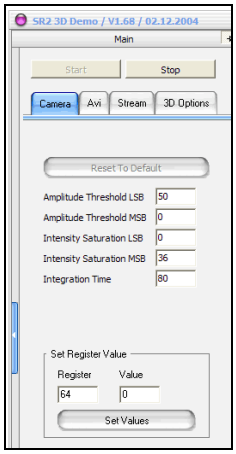
In Figure 60c movie files (AVI format) can be recorded.

In Figure 60d the stream function (frame-grabbing) of collecting several range frames is displayed. Enter a value of how many range frames should be scanned allows to measure moving targets over a period of time. The histogram option has been selected as well as a different 3D view has been chosen.

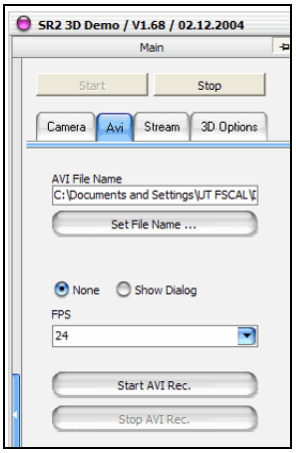
المنارة للاستشارات
www.manaraa.com



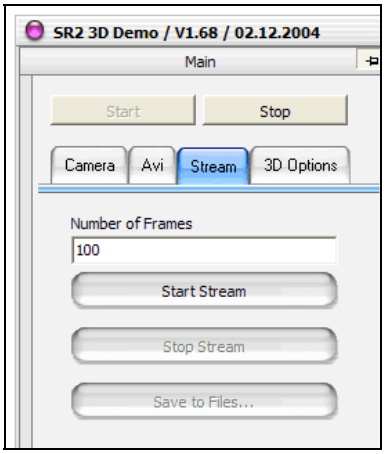
(a)



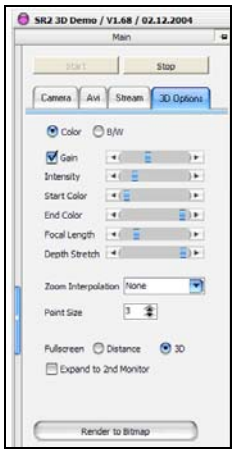
(b)



(c)



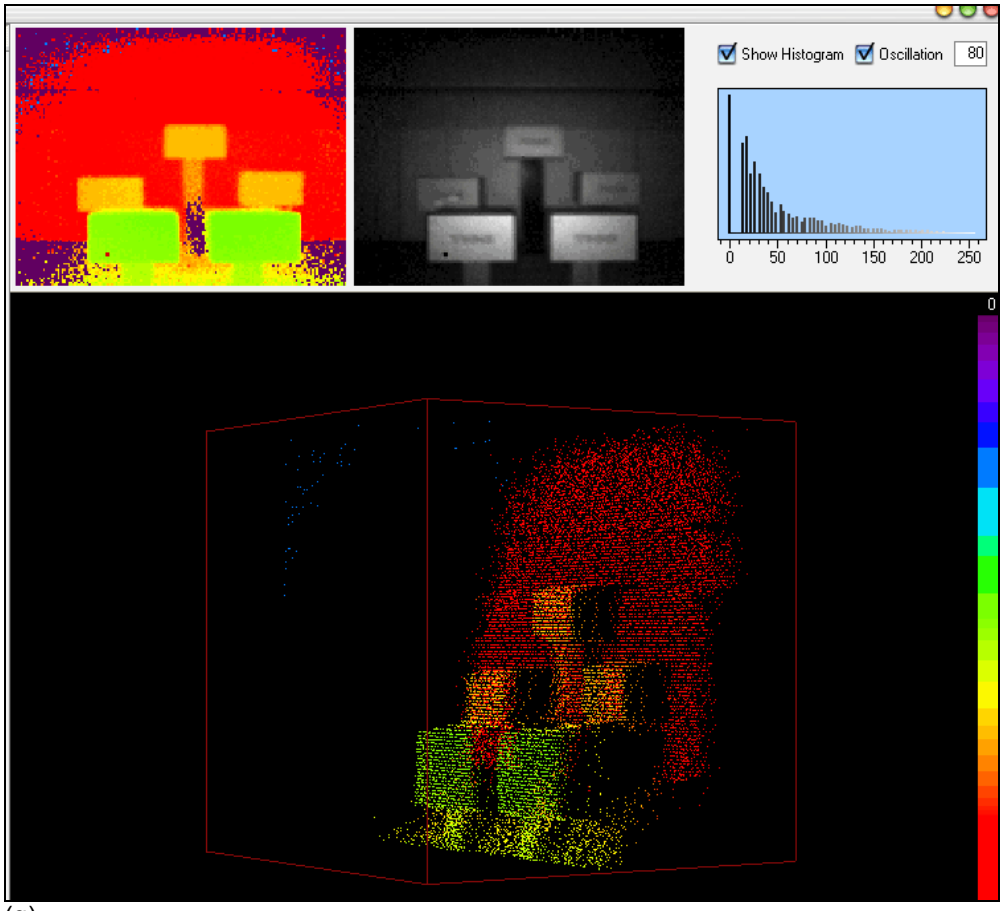
(d)



(e)

(f)

المنارة للاستشارات
www.manaraa.com



(g)

Figure 60: (a) Screenshot of camera software interface with main tab options, (b) Start tab, (c) Movie tab (d) Range grabbing tab, (e) 3D display options tab, (f) Frame rate, sensor model, update display, and (g) Depth, intensity/amplitude, and histogram window.

In Figure 60e, f, and g options allow to modify the image properties like color signature or black/white, gain, intensity, start color, end color, focal length, depth stretch, zoom interpolation, point size, full screen or 3D, expand to second monitor, render to bitmap. By modifying the registry values the readout from the camera to the computer can be directly controlled. Real-time application to submit the range readout directly to the image processing tools stored on a computer.

المنارة للاستشارات
www.manaraa.com

References

Arman F., and Aggarwal J.K. (1993). "Model-based Object Recognition in Dense-Range Images – A review". *CM Computing Surveys*. 25(1), 5-43.

Besl, B. (1988). "Active Optical Range Imaging Sensors", *Machine Vision and Applications*, 1, 127-152.

Besl, P.J., and McKay, N.D. (1992). "A method for registration of 3D shapes". *IEEE PAMI*, 14(2), 239-256.

Blostein, S.D. and Huang, T.S. (1987). "Algorithms for Motion Estimation Based on 3-D Point Correspondences in Motion Understanding". ed. J. K. Aggarwal and W. N. Martin, Kluwer Pub.

Bouma B.E. and Tearney G.J. (2001). Handbook of Optical Coherence Tomography, Marcel Dekker Inc. New York.

Bureau of Labor Statistics (2005). U.S. Department of Labor. Washington D.C., http://www.bls.gov/iif/oshcfoi1.htm (Accessed December 8, 2005).

Büttgen, B., Oggier, T., Lehmann, M., Kaufmann, R. (2005). "CCD/CMOS Lock-in Pixel for Range Imaging: Challenges, Limitations and State-of-the-Art". *1st Range Imaging Research Day*, ETH-Zürich, 21-32.

Center to Protect Workers' Rights (2002). *The Construction Chart Book.* Third edition, Silver Spring, MD.

Chen, Y. and Medioni (1991). "Object Modeling by Registration of Multiple Range Images". *Proceedings International Conference Robotics and Automation*, 2724-2729.

Cheok, G.S. and Stone, W.C. (1999). "Non-intrusive scanning technology for construction assessment". *Proceedings of the 16th International Symposium on Automation and Robotics in Construction*, 645-650.

Construction Industry Institute (1993). "Zero Injury Economics." Special Publication 32-2, The University of Texas at Austin, TX.

CSEM AG (2005). Operation Manual SwissRanger 2. Different versions. http://www.swissranger.com. Last accessed June 12, 2006.

المنارة للاستشارات
www.manaraa.com

Daily, M.J., Harris, J.G., and Reiser, K. (1987). “Detecting Obstacles in Range Imagery”. *Proceedings DARPA Image Understanding Workshop*.

DeWeert, M. and Gilbert, G. (2006). “When Sonar Isn’t Enough: Marine Optics for Port Security”. SPIE Short Course SC586 Course Notes, Orlando, Florida, April 17, 2006.

Dietiker, Miro (2003). “3D Video – Bilderfassung und -Verarbeitung”. Diploma Thesis. Hochschule für Technik Rapperwil, Switzerland.

Duchaineau, M., Wolinsky, M., Sigeti, D.E., Miller, M.C., Aldrich, C., and Minneev-Weinstain, M.B. (1999). “ROAMing terrain: Real-time optimally adapting meshes.” *Proceedings of IEEE Visualization 97*, 81-88.

ENR (2000). “Special Issue – Education for Careers in Construction.” *Engineering News Record*, 245(17), 10-12.

ENR (2003). “Construction Facts.” *Engineering News Record*, 251(20a), 108.

Freedman and Kaufmann (2006). http://eosweb.larc.nasa.gov/EDDOCS. Originally from: Universe by Freedman and Kaufmann, Accessed May 21, 2006.

Gonzales, R.C., and R.E. Woods (1992). *Digital Image Processing*. Addison-Wesley Publishing Company.

Gordon, G., M. Billinghurst, M. Bell, J. Woodfill, B. Kowalik, A. Erendi, and J. Tilander (2002). “The Use of Dense Stereo Range Data in Augmented Reality”. *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR02)*.

Guralnik, V. and Karypis G. (2001). "A Scalable Algorithm for Clustering Protein Sequences". *Workshop on Data Mining in Bioinformatics*, 73-80.

Gut, Oliver (2004). Untersuchungen des 3D-Sensors SwissRangerTM. Diploma-Thesis, Swiss Federal Institute (ETH), Zürich, Switzerland.

Haas, C.T.. (2004). “Innovative Construction Workforce”. Proceedings of The Sloan 2004 Annual Conference in Atlanta.

Haas, C.T. and Kim, Y-S. (2002). “Automation in Infrastructure Construction”, *Construction Innovation*, 2(3)3, 191-209.

Haala, N. and Brenner, C. (1999). “Extraction of buildings and trees in urban environments”. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3), 130-137.

المنارة للاستشارات
www.manaraa.com

Hähnel, D., Burgard, W., Thrun, S. (2001). “learning 3D Models of Indoor and Outdoor Environments With a Mobile Robot”. The 4th European Workshop on Advanced Mocile Robots (EUROBOT 01), 91-98.

Harlick, R.M., Joo, H., Lee, C., Zhunang, X., Vaidya, V.G., and Kim, M.B. (1989). “Pose estimation from corresponding point data”. *IEEE Transactions on Systems, Man and Cybernetics*, 19(6), 1426-1446.

Harvard-Smithsonian Center for Astrophysics (2006). http://cfa-www.harvard.edu/ , May 21, 2006).

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

Holmgren, J. and Persson, A. (2004). “Identifying species of individual trees using airborne laser scanner”. *Remote Sensing Environment*, 90(4), 415-423.

Hoppe, H. (1998). “Smooth view-dependent level-of-detail control and its application to terrain rendering”. *Proceedings of IEEE Visualization*, 35-42.

Hoover, A, D. Goldgof, K. Bowyer (2003). “Egomotion estimation of a range camera using the space envelope”. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 33(4), 717-721.

Hosticka B., Seitz, P. and Simoni A. (2005). “Optical TOF Sensors for Solid-State 3D Vision, Encyclopedia of Sensors:. Editors: Grimes, C. and Dickes E.C., American Scientific Publishers.

Hsu, S., Samarasekera, S., and Kumar, R. (2003). “Automatic registration and visualization of occluded targets using ladar data”. *Proceedings SPIE 5086*, 209-220.

Iqbal, Q. and Aggarwal J.K. (1999). “Using Structure in Content-based Image Retrieval”. Proceedings of the IASTED International Conference Signal and Image Processing (SIP), 129-133.

Jain, A.K. and Dubes, R.C. (1988). Algorithms for Clustering Data. Prentice Hall.

Jain, A.K., Murty, M.N., and Flynn, P.J. (1999). “Data Clustering: A Review”. ACM Computing Surveys, 31(3).

Jensen, B., Weingarten, J., Kolski, S., and Siegwart, R. (2005). “Laser Range Imaging using Mobile Robots: From Pose Estimation to 3D Models”. 1st Range Imaging Research Day, ETH-Zürich, 129-144.

المنارة للاستشارات
www.manaraa.com

Johnson, A.E. and M. Herbert M. A (1997). “System for Semi-automatic Modeling of Complex Environments”. *Proceedings of International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, 213-220.

Johnson, S.C. (1967). "Hierarchical Clustering Schemes" Psychometrika, 2, 241-254.

Jong, Arie de (2006). Infrared Threat Detection Systems. SPIE Short Course SC628 Course Notes, Orlando, Florida, April 21, 2006.

Kahlmann, T. and Ingensand, H. (2005). “Range Imaging Sensor Properties and Calibration”. *1st Range Imaging Research Day*, ETH-Zürich, 71-80.

Karbacher S., X. Laboureux, N. Schön, and G. Häusler (2001). “Processing Range Data for Reverse Engineering and Virtual Reality”. *Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling, IEEE,* 314-321.

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”. Conference Proceedings. Earth & Space, 9th ASCE Aerospace Division International Conference.

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.

Kubacki, J. and Pfeiffer, K. (2005). “Using Range Imaging Sensor with Color Imaging Sensors in Cognitive Robot Companions: A new and Simple Calibration Technique Based on Particle Swarm Optimization”. *1st Range Imaging Research Day*, ETH-Zürich, 43-57.

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.

Lange, R. (2000). 3D Time-Of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology. Dissertation, University of Siegen.

Lange, R. and Seitz, P. (2001). “Solid-state, Time-Of-Flight Range Camera”. IEEE Journal of Quantum Electronics, 37, 390-397.

Lindstrom, P,, Koller, D., Ribarsky, W., Hodges, L.F., Faust, N., and Turner, G. (1996). “Real-time continuous level of detail rendering of height fields”. *Proceedings of SIGGRAPH*, 109-118.

Lindstrom, P. and Pasucci, V. (2001). “Visualization of Large Terrain Made Easy”. *Proceedings of IEEE Visualization 2001*, 363-370.

المنارة للاستشارات
www.manaraa.com

Luebke, D., Reddy, M., Cohen, J.D., Varshney, A., Watson, B., and Huebner, R. (2003). *Level of detail for 3D graphics*. Morgan Kaufmann Publishers, San Francisco, CA.

Lytle, A.M., Saidi ,K.S., Stone, W.C., and Scott, N.A. (2003). "Towards an Intelligent Job Site: Status of the NIST Automated Steel Construction Test Bed". *Proceedings of 20th International Symposium on Automation and Robotics in Construction*.

Lytle, A.M. Katz, I., and Saidi, K.S. (2005). "Performance Evaluation of a High-Frame Rate 3D Range Sensor for Construction Applications", *Proceedings of 22th International Symposium on Automation and Robotics in Construction*.

McLaughlin, J., Sreenivasan, S.V., Haas, C.T., and Liapi, K.A. (2004). "Rapid Human-Assisted Creation of Bounding Models for Obstacle Avoidance in Construction". *Journal of Computer-Aided Civil and Infrastructure Engineering*, 19, 3-15.

Mine Safety and Health Administration (2003). http://www.msha.gov/equipmentsafety/equipmentaccidents.asp. Accessed May 21, 2006.

Moravec, H.P. (1996). "Robot Spatial Perception by Stereoscopic Vision and 3D Evidence Grids". Technical Report Canergie Mellon University, CMU-RI-TR-96-34.

Moravec, H.P. and Elfes, A. (1985). "High-Resolution Maps from Wide-Angle Sonar". Proceedings of the IEEE International Conference on Robotics and Automation.

Murray, D., and Little, J.J. (1998). "Using Real-Time Stereo Vision for Mobile Robot Navigation". Proceedings of the IEEE Workshop on Perception for Mobile Agents.

NIOSH (2005). http://www.cdc.gov/niosh/topics/safety.html. Accessed December 8, 2005.

Occupational Safety and Health Standards (1990). "Excavation Final Rule". Safety and Health Administration, Federal Register, U.S. Department of Labor, 54, 209.

OSHA (2005a). Keeping Your Workplace Safe - Q&A's for Small Business Employers. http://www.osha.gov/Publications/98-40brochure.html, Accessed November 18, 2005

OSHA (2005b). OSHA Construction eTool. http://www.osha.gov/SLTC/etools/construction/index.html. Accessed November 18, 2005.

المنارة للاستشارات
www.manaraa.com

Oggier, T. (2003). "An All-Solid State Optical Range Camera for 3D Real-Time Imaging with Sub-Centimeter Depth Resolution (SwissRangerTM)". *Proceedings of SPIE*, Vol. 5249, No. 65.

Oggier T., Kaufmann, R., Lehmann, M., Buettgen, B., Neukom, S., Richter, M., Schweizer, M., Metzler, P., Lustenberger, F., Blanc, N. (2005). "Novel Pixel Architecture with Inherent Background Suppression for 3D Time-of-Flight Imaging", *SPIE Electronic Imaging*, San Jose.

PMDtec GmbH (2005). Presentation Slides for *1st Range Imaging Research Day*, ETH-Zürich.

Pulli, K. (1999). "Multiview registration for large data sets". *Proceedings of the 2nd International Conference on 3D Digital Imaging and Modeling*, 160-168.

Puppo, E. (1998). "Variable Resolution Triangulations". *Computation Geometry*, 11(3-4), 219-238.

Robotics Trends. http://www.roboticstrends.com/displayarticle389.html. Accessed September 14, 2004.

Rohner, M., Siercks, K., and Hinderling, J. (2005). "Single Photon Detection for High Precision Ranging – A Trade-off Study". *1st Range Imaging Research Day*, ETH-Zürich, 33-42.

Roth, M.W., Scheck, A.E., Chui, W.W., Murphy, K.E. (2005). "Precision geo-location at longe range with multiple-look Lidar". *Proceedings SPIE 5791*.

Roth, Michael (2006). "3D Visualization Techniques for Laser Radar". The Johns Hopkins University Applied Physics Laboratory. SPIE Short Course SC717 Course Notes, Orlando, Florida, April 21, 2006.

Rottger, S., Heidrich, W., Slussallek, P., and Seidel, H.-P. (1998). "Real-time Generation of continuous levels of detail for height fields". *Proceedings of the International Conference in Central Europe on Computer Graphics and Visualization*, 315-322.

Sanders-Reed, J. (2006). "Image Based Motion Analysis". SPIE Short Course SC536 Course Notes, Orlando, Florida, April 20, 2006.

Sato, K. and Aggarwal, J.K. (2004)"Tracking and Recognizing Two-Person Interactions in Outdoor Image Sequences," Computer Vision and Image Understanding, 96, 100–128.

Schwarte, R., Haeusler G. Malz, R.W. (2000). "Computer Vision and Applications: A guide for Students and Practitioners - Three-dimensional Imaging Techniques. Academic Press.

المنارة للاستشارات
www.manaraa.com

Seitz, Peter (2005). "Unfied analysis of the peprformance and physical limitations of optical range-imaging techniques". *1st Range Imaging Research Day*, ETH-Zürich, 9-17.

Shaw, Joseph (2006). Introduction to Optical and Infrared Sensor Systems". Montana State University-Bozemann. SPIE Short Course SC789 Course Notes, Orlando, Florida, April 21, 2006.

Shum, H., K. Ikeuchi, and R. Reddy (1994). "Virtual Reality Modeling from a Sequence of Range Images". The Robotics Institute, Carnegie Mellon University.

Singh, S. and P. Keller (1991). "Obstacle Detection for High Speed Navigation" *Proceedings of IEEE International Conference on Robotics and Automation.*

Sithole, G. and Vosselman, G. (2004). "Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds". *ISPRS Journal of Photogrammtery and Remote Sensing*, 59(1-2), 85-101.

Stamos, I. and Allen, P.K. (2000). "3D Model Construction Using Range and Image Data". Submitted to CVPR 2000. Department of Computer Science, Columbia University, New York, NY 10027.

Steitz, A. and Pannekamp, J. (2005). "Systematic Investigation of Properties of PMD-Sensors". *1st Range Imaging Research Day*, ETH-Zürich, 59-69.

Stone, W.C., Juberts, M., Dagalskis, N. Stone, N.,Gorman, J. (2004). Performance Analysis of Next Generation LADSAR for Manufacturing, Construction, and Mobility. NISTIR 7117, United States Department of Commerce – National Institute of Standards and Technology, NIST.

Södermann, U. and Elstrom, M.D. (2001). "Stereo-based registration of range and projective imagery for data fusion and visualization". *Optical Engineering*, 40, 352-361.

Surmann, H., Nichter, A., and Hertzberg, J. (2003). "An Autonomous Mobile Robot with 3D Laser Range Finder for 3D Exploration and Digitalization of Indoor Environments". *Journal of Robotics and Autonomous Systems.*

Tan, P., Steinbach, M., and Kumar, V. (2005). Introduction to Data Mining. Addison Wesley.

Takarski, V.I. (1961). Wave propagation in a turbulent medium. McGraw-Hill, New York.

Thrun, S., Burgard, W., Fox, D. (2000). "A Real-time Algorithm for Mobile Robot Mapping with Applications to Multi-robot and 3D Mapping." Proceedings ICAR.

المنارة للاستشارات
www.manaraa.com

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

Teizer, J., Liapi, K.A., Caldas, C.H., and Haas, C.T. (2005b), "Experiments in Real-time Spatial Data Acquisition for Obstacle Detection." Proceedings of the Construction Research Congress, 547-551.

Teizer, J., Haas, C.T., Caldas, C.H., Bosche, F. (2006). "Real-Time Three-Dimensional Object Detection and Tracking in Transportation". *Proceedings of the 9th International Conference on Applications of Advanced Technology in Transportation*.

Wang, C., Thorpe, C., and Thrun, S. (2003). "Online Simultaneous Localization and Mapping with Detection and Tracking of Moving Objects: Theory and Results from a Ground Vehicle in Crowded Urban Areas". IEEE International Conference on Robotics and Automation.

ZuWhan, K (2001). "Multi-view 3-D Object Description with Uncertain Reasoning and Machine Learning". PhD Dissertation. Department of Computer Science of the University of Southern California.

Zywitza, F., Massen, J., Brunn, M., Lang, C., Göring, T. (2005). "One-to-Three-Dimensional Rnaging for Future Automotaive Safety Systems". *1st Range Imaging Research Day*, ETH-Zürich, 109-116.

Weingarten, J., Gruener, G., and Siegwart, R. (2004). "A State-of-the-Art 3D Sensor for Robot Navigation." *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3, 2155- 2160.

المنارة للاستشارات
www.manaraa.com

Vita

Jochen Teizer was born in Ettlingen, Germany, on December 09, 1975, and is the son of Marlitt and Franz Teizer. He graduated as Diplom-Ingenieur in Civil Engineering from the University of Karlsruhe (TH) in Germany in 2002 with specialty areas in construction technology and management, and facilities management. Jochen Teizer attended the University of Texas at Austin, U.S.A., from September 2002 until August 2006 to pursue his doctorate in civil, architectural and environmental engineering with a concentration in construction engineering and management.

Permanent address: Jochen Teizer, Schoellbronner Str. 32, 76275 Ettlingen, Germany

This dissertation was typed by Jochen Teizer.

المنارة للاستشارات
www.manaraa.com