

**THE AUGMENTATION OF URBAN SEARCH AND RESCUE DOGS WITH SENSING,  
CONTROL AND ACTUATION—EXTENDING THE METAPHOR, “DOG AS ROBOT”**

By

Jimmy Quang Minh Ngoc Tran  
Bachelor of Science  
in the program of Computer Science,  
Ryerson University 2007

Master of Science  
in the Program of Computer Science,  
Ryerson University 2009

A dissertation  
presented to Ryerson University

in partial fulfillment of the  
requirements for the degree of

Doctor of Philosophy  
in the Program of  
Computer Science

Toronto, Ontario, Canada, 2019  
© Jimmy Quang Minh Ngoc Tran, 2019

## **AUTHOR'S DECLARATION**

I hereby declare that I am the sole author of this dissertation. This is a true copy of the dissertation, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this dissertation to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this dissertation by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my dissertation may be made electronically available to the public.

## ABSTRACT

### **THE AUGMENTATION OF URBAN SEARCH AND RESCUE DOGS WITH SENSING, CONTROL AND ACTUATION—EXTENDING THE METAPHOR, “DOG AS ROBOT”**

Jimmy Quang Minh Ngoc Tran

Doctor of Philosophy, Computer Science, Ryerson University, 2019

When disaster strikes in urban areas, the devastating results are collapsed structures that may contain voids, and trapped people within. To a large extent, the speed with which these victims can be found and extricated determines the likelihood of their survival. Specially trained and equipped emergency first responders are tasked with trying to save their lives by locating and extricating trapped victims from these dangerous environments. Telepresence systems can help first responders search for casualties from a safe location. Most automated search systems intended for use in urban disasters, come in the form of remotely operated robots. This work takes a different approach to telepresence and robotics. This work is an extension of previous work that exploits the intelligence and characteristics of trained search dogs combined with compatible technology and used as components in new kinds of telepresence systems for urban search and rescue (USAR) operations.

The Canine Remote Deployment System (CRDS) is a tool that emergency responders can use to deliver critical supplies to trapped victims in rubble using dogs. The first contribution of this work is the development of the bark detection system for automatically triggering deployment of packages near trapped victims from the CRDS—guaranteeing accurate package deployment even when remote communication with the dog is impossible.

A well-known ground robot problem is the difficulty in designing a mobility mechanism to traverse rubble. Another contribution of this thesis is the Canine Assisted Robot Deployment (CARD) framework and the design of a robot capable of being carried by a search dog. This work extends the responder’s telepresence in rescue operations by bringing robots much deeper into the disaster site than current methods.

Visual odometry is used in location tracking in GPS-denied environments and can be used in rescue operations. This research explores the limitation of RGB-D cameras for visual odometry for this application. An algorithm called pseudo-Random Interest Points Extractor was developed

to track images over visually feature-sparse areas with the potential use of visually reconstructing canine search paths to victims. This work concentrates on using visual odometry from data collected from a search dog-mounted RGB-D camera. The task of model stabilization is difficult due to the nature of dog's constant and unpredictable movements, as the data contains many motion blurred images. The development of an algorithm called Intelligent Frame Selector is shown to improve visual odometry for systems carried by search dogs by intelligently filtering data and selecting only usable frames. The algorithm can be applied to any general visual odometry pipeline beneficially as the technique reduces cumulative error problems by using less data.

## ACKNOWLEDGEMENTS

I am grateful to my mother, who has provided me with unwavering love and support in my life. I am also grateful for my late father, who instilled in me my sense of curiosity and appreciation for education.

I would like to express my sincere gratitude to my supervisor and mentor Dr. Alexander Ferworn. He has given me the opportunity to be a part of the Network-Centric Applied Research Team (NCART), to follow and share his vision in our research work. Under his supervision, I learned how to conduct research and he helped foster my own research path. This work would not be possible without his guidance, support, and encouragement.

I would like to thank the other committee members, Dr. Vincent Chan, and Dr. Goldie Nejat for reviewing this dissertation. I especially want to thank Dr. Alireza Sadeghian, and Dr. Isaac Woungang for being on my committee, reviewing this dissertation, and also for having taught me in my undergraduate classes.

I want to convey my appreciation to my fellow NCART teammates, Alex Ufkes, and Martin Gerdzhev for their collaboration and help in conducting experiments. It was especially fun bouncing ideas and working through problems with Alex.

I would like to acknowledge and give special thanks to my friend, Devin Ostrom, who taught me much about fabrication, mechanical design, and who provided words of encouragements when I needed them most.

To canine handlers Kevin Barnum, Dan Bailey, Tom Haus, and Athena Haus, I greatly appreciate you and your dogs Dare and Fritag for helping us with our experiments. I was deeply saddened by the news of their passing and miss them very much.

My Ph.D. program was funded in part by Graduate Fellowships endowed by Ryerson University, the Ministry of Training, Colleges and Universities of Ontario, and the Natural Sciences and Engineering Research Council of Canada (NSERC). I remain grateful for this support.

## TABLE OF CONTENTS

<b>Abstract.....</b>	<b>iii</b>
<b>List of Figures.....</b>	<b>ix</b>
<b>List of Tables .....</b>	<b>xiii</b>
<b>Abbreviations .....</b>	<b>xiv</b>
<b>CHAPTER 1 Introduction.....</b>	<b>1</b>
1.1 Introduction.....	1
1.2 Problem Definition.....	4
1.3 Objectives .....	8
1.4 Structure of this Dissertation .....	9
<b>CHAPTER 2 Remotely Affecting Disaster Incidents.....</b>	<b>12</b>
2.1 Introduction.....	12
2.2 Related Work .....	12
2.3 Continuing work on CRDS.....	18
2.4 Bark Detection .....	22
2.5 Bark Release Experiments .....	25
2.5.1 Bark Release Results.....	28
2.6 Conclusions.....	29
<b>CHAPTER 3 Extending a Rescuer’s interaction Within a Disaster Environment.....</b>	<b>31</b>
3.1 Introduction.....	31
3.2 Canine Assisted Robot Deployment .....	32
3.2.1 Unmanned Ground Vehicles in Urban Search and Rescue Operations .....	32
3.2.1.1 Performance Metric of Rescue Robot for USAR .....	34
3.2.1.2 Marsupial Robot Operations .....	36
3.3 Canine Assisted Robot Deployment Framework.....	37
3.4 Canine-Delivered Marsupial Robots.....	39

3.4.1	<i>Drop an Explore (DEX)–A Design for Canine-Delivered Marsupial Robot</i>	39
3.4.2	<i>DEX Field Experiments</i>	44
3.4.2.1	DEX Deployment Mechanism Experiment	44
3.4.2.2	Tunnel Experiment	45
3.4.2.3	Rubble Experiment	48
3.4.3	<i>DEX Experiment Results</i>	50
3.4.3.1	DEX Deployment Mechanism Experiment Results	50
3.4.3.2	Tunnel Experiment Results	51
3.4.3.3	Rubble Experiment Results	53
3.4.4	<i>Conclusion and Future Work</i>	55
<b>CHAPTER 4 3D Disaster Scene Reconstruction with RGB-D Cameras</b>		<b>56</b>
4.1	Introduction	56
4.2	Related Work	57
4.3	Low-Cost 3D Scene Reconstruction for Response Robots in Real-time Technical Approach	60
4.3.1	<i>RGB-D Visual Odometry</i>	60
4.3.2	<i>pseudo-Random Interest Points Extractor Algorithm</i>	63
4.3.3	<i>3D Reconstruction Experiments</i>	67
4.3.3.1	Computing Hardware	67
4.3.3.2	Test Robot	67
4.3.3.3	Test Environment	68
4.3.3.4	Test Method	70
4.3.4	<i>Results</i>	70
4.3.5	<i>Discussion and Conclusion</i>	75
4.4	Initial Experiments on 3D Modeling of Complex Disaster Environments Using Unmanned Aerial	75
4.4.1	<i>Experiments</i>	76
4.4.1.1	UAV	76
4.4.1.2	Sensor	77
4.4.1.3	Computing Hardware	78

4.4.1.4	3D Point Cloud Stitching Algorithm .....	79
4.4.1.5	Tasks .....	81
4.4.2	<i>Results</i> .....	82
4.4.3	<i>Discussion</i> .....	87
<b>CHAPTER 5</b>	<b>Visual Odometry from Searching Canines with RGB-D Sensors .....</b>	<b>88</b>
5.1	Introduction.....	88
5.2	Related Work .....	89
5.3	Contributions.....	91
5.4	Technical Approach .....	91
5.4.1	<i>Data Collection</i> .....	92
5.4.2	<i>Data Analysis</i> .....	93
5.4.3	<i>Data Processing</i> .....	94
5.5	Experiments .....	97
5.6	Results.....	99
5.7	Discussion and Conclusion .....	101
<b>CHAPTER 6</b>	<b>Summary and Conclusions .....</b>	<b>105</b>
6.1	General Conclusion.....	105
6.2	Contributions.....	106
6.3	Limitations of this work.....	107
6.4	Future Work .....	108
6.5	Concluding Remarks.....	110
<b>References</b> .....		<b>111</b>

## LIST OF FIGURES

Figure 1.1 – Population density of the United States from 1790 to 2017 in residents per square mile of land area [7].....	2
Figure 1.2 – Pancake collapse diagram [13].....	4
Figure 1.3 – Lean-to collapse diagram [13].....	4
Figure 1.4 – ‘V’ collapse diagram [13].....	5
Figure 1.5 – A-Frame collapse diagram [13].....	5
Figure 2.1 – USAR canine wearing CAT dog goggles at exercise in Fergus, Ontario 2006 .....	13
Figure 2.2 – USAR Dog Dare wearing CAT 2.0 prototype.....	14
Figure 2.3 – CAT 3.0 worn by CA-TF1 USAR Dog Freitag .....	14
Figure 2.4 – Canine equipped with CRDS and underdog.....	16
Figure 2.5 – USAR Dog left underdog behind .....	17
Figure 2.6 – CRDS detail. From left to right: Remote, CRDS, harness, underdog .....	17
Figure 2.7 – Latest version of CRDS in Retail Package.....	19
Figure 2.8 – Latest version of CRDS with tactical vest, demoed on model dog .....	20
Figure 2.9 – Block diagram of the electronics components of the CRDS main unit.....	20
Figure 2.10 – Block diagram of the electronics components of CRDS remote unit .....	21
Figure 2.11 – Bark Release Algorithm .....	23
Figure 2.12 – Block diagram of the electronics components of Bark Release System .....	24
Figure 2.13 – Audio amplifier circuit schematic .....	25
Figure 2.14 – Bark Barrel Test .....	26
Figure 2.15 – Location on rubble pile where the quarry hid .....	27
Figure 2.16 – Dog finds victim.....	27
Figure 2.17 – Circle indicates “hide” location, Square show pooling of smoke and, presumably, scent from the quarry .....	30

Figure 3.1 – Kenaf rescue robot with features including crawler enclosing its entire body and four sub-crawler arms that are each controlled independently [40, 41] .....	33
Figure 3.2 – Packbot having undergone track shedding after NIST Response Robot endurance testing.....	34
Figure 3.3 – Matilda robot with scratched and abraded armoured front visor. Damage caused by attempting to traverse rubble seen in background .....	34
Figure 3.4 – Purpose built rubble pile for USAR training at “Disaster City” .....	35
Figure 3.5 – An example of a robot test “step field” in one configuration of a NIST mobility test .....	36
Figure 3.6 – DEX 1.0, the first prototype .....	39
Figure 3.7 – DEX with retractable hook.....	41
Figure 3.8 – DEX attached with hooks to CRDS .....	41
Figure 3.9 – DEX in underdog attached to CRDS.....	42
Figure 3.10 – Block Diagram of DEX 1.0 components.....	43
Figure 3.11 – Rubble Pile at Disaster City .....	45
Figure 3.12 – Map of tunnel under rubble .....	46
Figure 3.13 – Centre-fire Shooting Target.....	47
Figure 3.14 – Dot-mil Confidence Target.....	47
Figure 3.15 – IEA Resolution Chart 1956 .....	48
Figure 3.16 –Top view of OPP's Rubble Pile .....	49
Figure 3.17 – Side view of OPP's Rubble pile.....	49
Figure 3.18 – DEX being thrown from Freitag.....	50
Figure 3.19 – DEX wrapped by underdog.....	51
Figure 3.20 – DEX released by bark barrel .....	51
Figure 3.21 – Image from DEX .....	52
Figure 3.22 – Dex’s deployment map.....	53

Figure 3.23 – Quarry's location-hidden from view by orange cover .....	54
Figure 3.24 – DEX's deployment on rubble (left), image taken from DEX (right).....	54
Figure 4.1 – Traditional 2D Perspective from a robot's camera.....	58
Figure 4.2 – 3D viewpoint of Figure 4.1 .....	58
Figure 4.3 – Microsoft Kinect gaming controller .....	60
Figure 4.4 – Block diagram of visual odometry pipeline .....	63
Figure 4.5 – pRIPE algorithm block diagram.....	65
Figure 4.6 – pRIPE Feature Location Example .....	66
Figure 4.7 – Pedsco RMI-9WT Remote Mobile Investigator equipped with Kinect sensor .....	68
Figure 4.8 – The effect of ambient sunlight on Kinect depth data. black regions show where the Kinect failed to measure the depth.....	69
Figure 4.9 – Environment 1 (lab).....	69
Figure 4.10 – Left: Environment 2 (narrow hallway), Right: Environment 3 (wide hallway).....	70
Figure 4.11 – Top and side view of model created by SURF + descriptor.....	71
matching in Environment 1 (lab) .....	71
Figure 4.12 – Top and side view of model created by Shi-Tomasi corners tracked with optic flow in Environment 1 (lab).....	72
Figure 4.13 – Top and side view of model created by pRIPE in Environment 2 (narrow hallway) .....	72
Figure 4.14 – Top and side view of model created by SURF + descriptor matching in Environment 2 (narrow hallway).....	73
Figure 4.15 – Top and side view of model created by pRIPE in Environment 3 (wide hallway) 73	
Figure 4.16 – Reference Rubble Pile at UCRT HQ, Bolton, Ontario.....	76
Figure 4.17 – MikroKopter Hexakopter 2 .....	77
Figure 4.18 – UAV with hardware payload.....	79
Figure 4.19 – Overhead view of the rubble pile. Circled are two areas surveyed by the UAV ...	82

Figure 4.20 – Sunny vs. shady depth data .....	82
Figure 4.21 – 3D Rubble swath, side and top view .....	84
Figure 4.22 – 3D Rubble 'U', top and side views.....	85
Figure 4.23 – Void, front and side 3D views.....	85
Figure 4.24 – Void entrance, 2D image.....	86
Figure 4.25 – Top and side views of hidden void.....	86
Figure 4.26 – Aerial image of hidden void.....	87
Figure 5.1 – RGB-D Sensor mounted on search dog Dare.....	92
Figure 5.2 – Depiction of sensor motion on a moving dog .....	93
Figure 5.3 – Example of feature matching between a clear and blurry image .....	94
Figure 5.4 – Consecutive series of blurry frames (red) between clear frames (green) on the inflection points. ....	94
Figure 5.5 – Sample histogram of a window of matches.....	96
Figure 5.6 – IFS Algorithm Block Diagram .....	96
Figure 5.7 – Plain interior hallway at George Vari Engineering Building at Ryerson University	98
Figure 5.8 – From to top bottom, model generated from Cart, Slow1, Medium1 and Fast1 datasets.....	99
Figure 5.9 – Models generated without IFS (top-Slow1, bottom-Fast1 dataset).....	100
Figure 5.10 – Models generated by human assisted process (top-Slow1, bottom-Fast1) .....	100
Figure 5.11 – Boston Dynamics’ BigDog robot walks on cinder blocks as reported in [104]...	102
Figure 5.12 – Boston Dynamics’ Petman robot with and without suit as reported in [105] .....	103
Figure 5.13 – Picture of Google Glass made by Google [106].....	103

## LIST OF TABLES

Table 2.1 – Distance from underdog to victim .....	28
Table 3.1 – Dimensions of DEX 1.0.....	40
Table 4.1 – Computing Hardware used .....	67
Table 4.2 – Model Completeness.....	73
Table 4.3 – Feature Extraction and tracking Speed Comparison.....	74
Table 5.1 – Hallway trials.....	98
Table 5.2 – Process time comparison .....	101

## ABBREVIATIONS

<b>3-D</b>	<b>3-dimensional</b>
<b>ASTM</b>	<b>American Society for Testing and Materials</b>
<b>BARS</b>	<b>Bark Release System</b>
<b>BI</b>	<b>Biological Intelligence</b>
<b>CARD</b>	<b>Canine Assisted Robot Deployment</b>
<b>CAT</b>	<b>Canine Augmentation Technology</b>
<b>CAT-5</b>	<b>Category-5</b>
<b>CDMR</b>	<b>Canine-Delivered Marsupial Robots</b>
<b>COTS</b>	<b>Commercial Off-The-Shelf</b>
<b>CRDS</b>	<b>Canine Remote Deployment System</b>
<b>DEX</b>	<b>Drop an Explore</b>
<b>DHS</b>	<b>Department of Homeland Security</b>
<b>DOF</b>	<b>Degrees of Freedom</b>
<b>EMS</b>	<b>Emergency Medical Services</b>
<b>FEMA</b>	<b>Federal Emergency Management Agency</b>
<b>FPS</b>	<b>Frames Per Second</b>
<b>GPS</b>	<b>Global Positioning System</b>
<b>GPU</b>	<b>Graphics Processing Unit</b>
<b>HUD</b>	<b>Head-Up Display</b>
<b>ICP</b>	<b>Iterative Closest Point</b>
<b>IEEE</b>	<b>Institute of Electrical and Electronics Engineers</b>
<b>IFS</b>	<b>Intelligent Frame Selector</b>
<b>IR</b>	<b>Infrared</b>
<b>LiPo</b>	<b>Lithium Polymer</b>
<b>NIST</b>	<b>National Institute of Standards and Technology</b>
<b>OPP</b>	<b>Ontario Provincial Police</b>
<b>PERT</b>	<b>Provincial Emergency Response Team</b>
<b>pRIPE</b>	<b>pseudo-Random Interest Points Extractor</b>
<b>PSF</b>	<b>Point Spread Function</b>
<b>RANSAC</b>	<b>Random Sample and Consensus</b>
<b>RF</b>	<b>Radio Frequency</b>
<b>RGB</b>	<b>Red Green Blue</b>
<b>RGB-D</b>	<b>Red Green Blue - Depth</b>
<b>SA</b>	<b>Situation Awareness</b>
<b>SBC</b>	<b>Single board computer</b>
<b>SD</b>	<b>Secure Digital</b>
<b>SLAM</b>	<b>Simultaneous Localization and Mapping</b>
<b>SSD</b>	<b>Solid State Drive</b>

<b>SURF</b>	<b>Speeded-Up Robust Features</b>
<b>SVD</b>	<b>Singular Value Decomposition</b>
<b>UAV</b>	<b>Unmanned Aerial Vehicle</b>
<b>UCRT</b>	<b>USAR CBRN Response Team</b>
<b>USAR</b>	<b>Urban Search and Rescue</b>
<b>USD</b>	<b>United States Dollar</b>
<b>UVG</b>	<b>Unmanned Ground Vehicles</b>
<b>WiFi</b>	<b>Wireless Fidelity</b>
<b>WMN</b>	<b>Wireless Mesh Network</b>

## **CHAPTER 1 Introduction**

### **1.1 Introduction**

“Telepresence” is a set of technologies that allow a person to feel they are present at a remote location and provide them with the ability to affect that environment [1]. This technology is used in a plethora of applications. In the office, the technology has been applied to applications as simple as video conferencing [2], to complex applications like telepresence robotics capable of driving around an office and interacting with people [3]. Even more complex applications include remote robotic surgery where a surgeon can operate on a patient without having to be in the same location by using sophisticated robotics. In these applications, the technology is used in a well known and structured environment where communications are generally assured, and the environment controlled. Other telepresence applications allow humans to explore hazardous environments such as volcanos, the deep sea, and outer space. Sometimes the environment itself is not inherently dangerous but other factors make it dangerous—like the presence of explosives, radiation or dangerous chemical agents. Each of these applications presents a set of unique challenges. An important application area involving inherently dangerous, unstructured environments that this dissertation focuses on is that of Urban Search and Rescue (USAR).

Disasters can be thought of as events which are much larger in scope and magnitude than would be expected in a routine emergency [4-6]. Typically, emergency management relies on planned services to respond to events predicted through a process of risk analysis. For example, a fire service might be planned and equipped to respond to residential and commercial fires based on population and building density. When someone calls the fire department because of their house is on fire, they and their insurance provider expect a fire truck to arrive within a specified time and to put out the fire. A disaster is different. In a disaster, it matters little which emergency service receives calls, all of them will be overwhelmed and not able to respond effectively without outside assistance.

Throughout human history, there have been many disasters. Disasters are especially devastating when they occur in urban areas with large populations. As the human population continues to increase worldwide, people congregate to live in metropolises which have more

people living in smaller spaces than have ever existed before [7]. A clear indication of this can be seen in Figure 1.1.

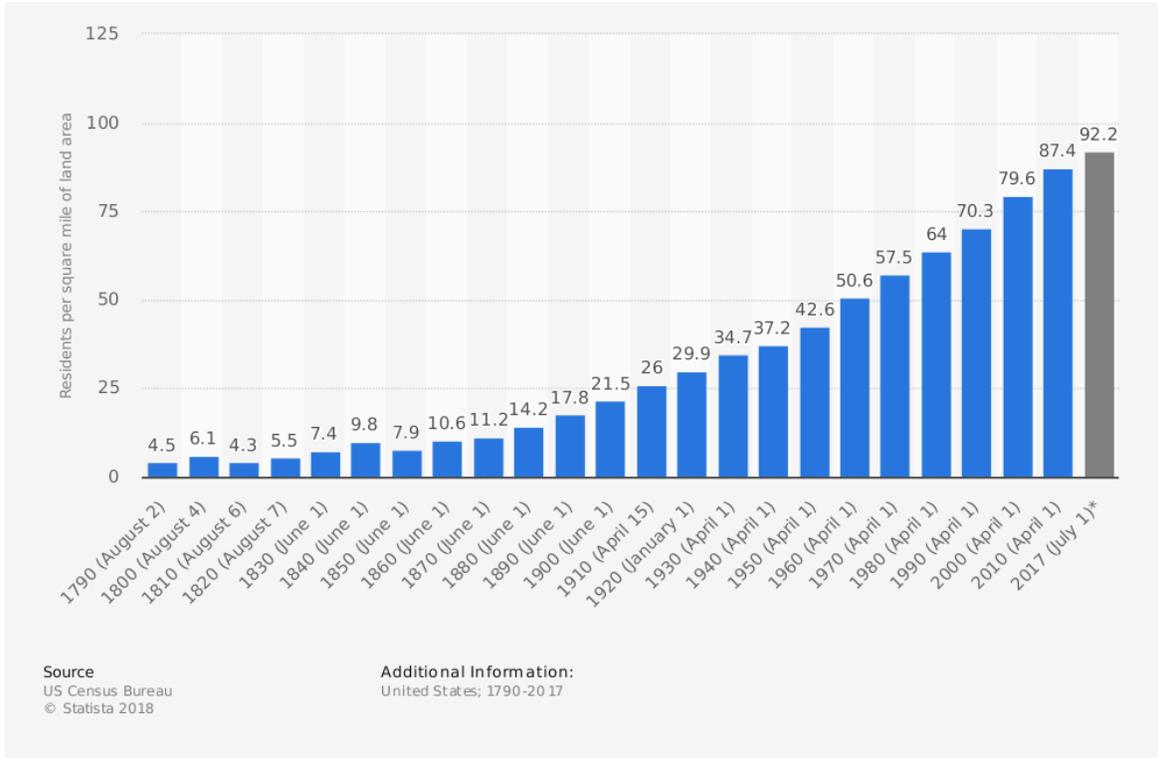


Figure 1.1 – Population density of the United States from 1790 to 2017 in residents per square mile of land area [7]

When a disaster strikes, it will have an affect on more people than would have been the case earlier in human history. Effects can be wide-ranging including direct mortality, economic loss, and long-term disadvantage. The causes of disasters can be naturally occurring such as, earthquakes, hurricanes, tornados, tsunamis, volcanic eruptions, or they can be man-made such as bombings, industrial accidents, and faulty engineering. A few notable examples of disaster incidents include the World Trade Center terrorist attacks leading to the structural collapse of several buildings (2001, 2996 people killed) [8], the Sichuan China earthquake (2008, 87000 people killed) [9], the Haiti earthquake (2010, 230000 people killed) [10], the Japan earthquake, tsunami and nuclear reactor meltdown (2011, 0 people killed directly, 1600 people killed from evacuation stress with an unknown number to be effected by low-level radiation in the atmosphere) [11].

Even in a country like Canada, where natural disasters are rare, the Algo Centre Mall collapse occurred in Elliot Lake, Ontario in 2012. While only two people were killed directly, the loss of the mall had an economic impact on the community of Elliot Lake which is felt even today. To mitigate the costs of disasters, it is then very important to respond to them effectively.

Public Safety Canada defines USAR as a general term for a group of specialized rescue skills that are integrated into a team with resources that include search, medical and structural assessment capacity [12]. Most USAR organizations are called “task forces” and are comprised of usually a regional or national organization that is staffed with first responders with special skills and equipped with specialized equipment that would not normally be available to local responders. While most USAR task forces are permanently established, their personnel are mostly volunteers who are regular fire fighters, police, Emergency Medical Services (EMS) or heavy equipment operators in their regular careers. Task Forces are deployed when the regional or national authority makes a formal declaration of a disaster situation that would require the response of the task force.

A common occurrence in urban disasters is the collapse of buildings or structures. While it is hard to imagine anyone surviving inside a collapsed structure, it is possible and relatively common that people are entombed in spaces within the resulting rubble called “voids”. USAR operations involve the dangerous task of locating, extracting, and medically stabilizing the survivors of such a collapse. Most of the challenge in USAR work is the result of unpredictable and unstable environment. As the structures formed within rubble as a result of an initial collapse are inherently unstable, any form of activity on or under the rubble could cause secondary collapses to occur. If a rescue is not performed precisely and carefully, secondary collapses can cause further deaths and injuries—including the first responders doing the work. However, speedy rescues are crucial as a victim’s survival rate is dependent on how fast they can be found and rescued.

It is hypothesized that search and rescue worker can be assisted in rescuing people more effectively and safely with the aid of an appropriate telepresence system or systems. The proposed system is a set of algorithms and technologies that allows a first responder the ability to remotely interact with an incident search site, interact with the casualties, acquire relevant metrics and visual data about the site, and possibly even determine a path to casualties under certain circumstances.

## 1.2 Problem Definition

To understand the technical challenges associated with applying the promise of telepresence systems within disaster environments, one must understand the environment that rescuers must deal with. When a building collapses, it typically does so in a pattern. These patterns are known as internal, external or total collapse [13].

There are four internal collapse patterns:

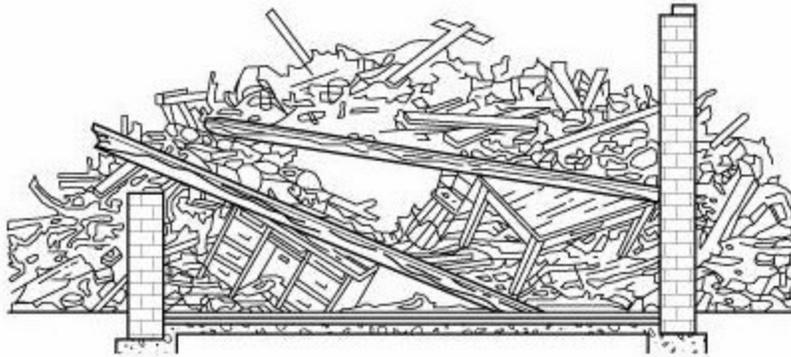


Figure 1.2 – Pancake collapse diagram [13]

**Pancake collapses** happen when load bearing walls fail, and the floors suddenly drop on top of each other, creating a stack up configuration, like pancakes. Voids are often created by large machinery, applications or furniture.

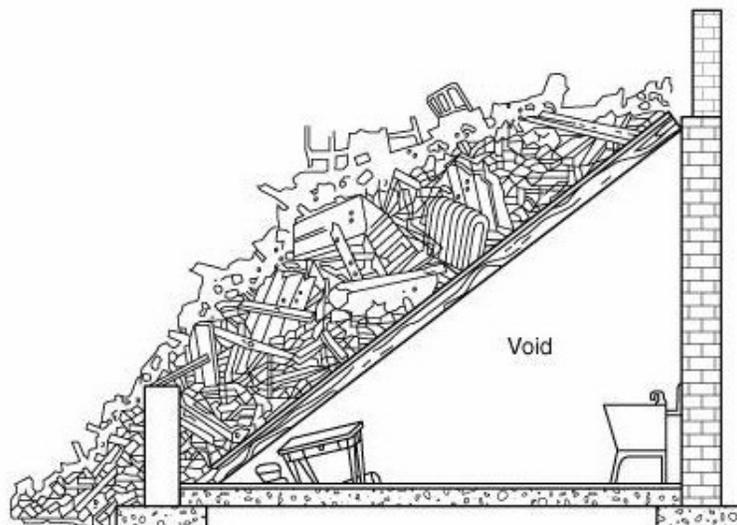


Figure 1.3 – Lean-to collapse diagram [13]

**Lean-to collapses** are the results of a supporting wall, beam or column at one end fails. This results in large voids, which has a better chance of finding survivors.

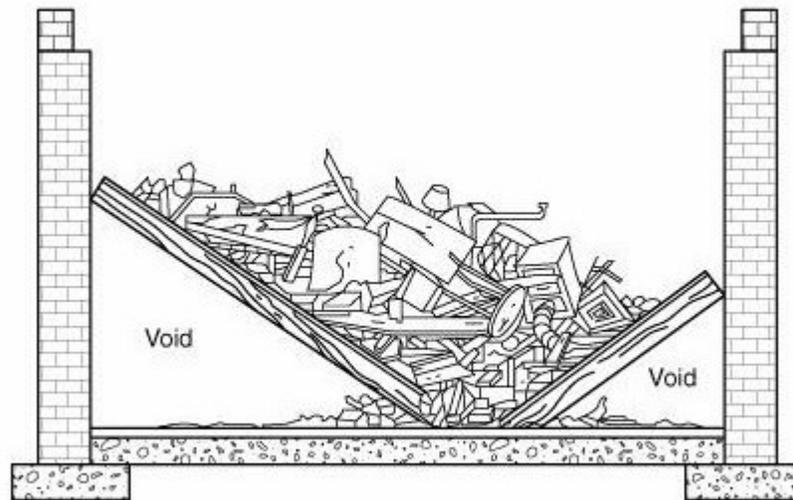


Figure 1.4 – 'V' collapse diagram [13]

'V' collapses occur when a heavy load at any given point cause the floor to break (usually in the middle) and fall on to the second floor forming two triangular voids.

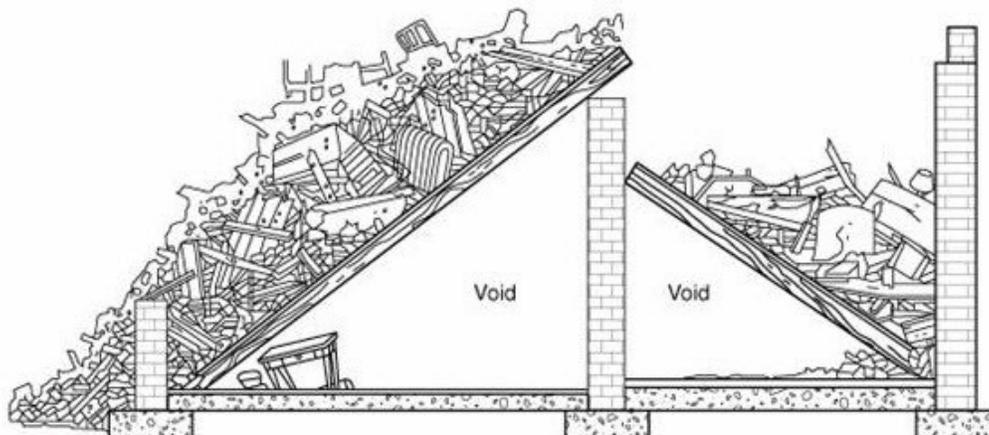


Figure 1.5 – A-Frame collapse diagram [13]

**A-Frame or tent collapses** are the opposite of V-collapses where the interior wall ends up partially supporting the ceiling, also forming triangular voids.

**Exterior collapses** occur when the outside walls fail. **Total collapses** are especially severe because all floors, walls, support structures fail and flatten down onto the ground floor or basement.

Victims of collapses have been found in all of these scenarios. The aftermath of a structural collapse is considered unsafe and no human exploration or search effort can begin until a structural engineer has examined the resulting rubble and determined a safe working area. Safe working areas are created by placing “shoring” and other structural supports at unstable points. Shorings are temporary supporting structures used to fortify weak or unstable points and take time to plan and construct. The search effort can also be impeded by a lack of entry points into underlying structures. Since doors and pathways no longer exist in their original form, gaining access requires finding or making access points by cutting through walls—often made of concrete. This is called breaching. The entire process is slow and arduous.

However, before rescuers decide to deploy all these resources in a particular area, they have to determine which area is best to begin with. Of course, it is desirable to make areas safe where there is a strong indication that victims are likely to be found.

Technologically, ground rescue robots are an obvious choice to assist in areas that are inherently unstable and dangerous. They can be equipped with a variety of sensors that can improve the situational awareness of the rescuers. The robot designs vary from wheeled to tracked. Some employ variable geometry that allow them to change shape, and others mimic biological systems like snakes that allow them to conform to the environment they find themselves in. However, while there are many rescue robots available, they are not used very often during USAR operations. The robots are still unable to travel very far on actual rubble.

The Intelligent Systems Division of the National Institute of Standards and Technology (NIST) of the U.S. Department of Commerce, sponsored by the Federal Emergency Management Agency (FEMA) and Department of Homeland Security (DHS) investigates how to measure the performance of rescue robots. Their goal is to determine how to evaluate robots in operation in an USAR environment [14] through the use of common metrics that determines the capabilities of a rescue robot. The purpose of these standardize metric is for emergency task forces to be able to purchase robots based on known capabilities. NIST has developed performance standards for many categories of characteristics of robots [15] through the American Society for Testing and Materials (ASTM) standards process, E54 Task Group [15-17].

Robots are evaluated on various performance characteristics including mobility, sensing capabilities and overall system performance (durability, communication, power) as well as secondary physical characteristics like “Cache Packaging”—indicating the physical size of the

robot and related equipment which must be transported by a task force to an operation. While there are niche areas where robot mobility is quite impressive, their overall performance in traversing open rubble is still inadequate.

As an alternative to ground robotics, Unmanned Aerial Vehicles (UAVs) can fly over rubble. UAVs in the form of a multicopter can easily fly over a rubble site, cover vast areas in short periods of time, and collect a multitude of data. However, UAVs cannot fly in the confined spaces of the interior of disaster sites. Various challenges face UAVs. Ground effects occur when multicopters fly close to the ground, ceiling or walls, and the airflow generated from the rotors interact with those surfaces and change the vehicles' thrust characteristics. This cause instability which quickly leads to crashes. Another problem is the wireless communication between the ground station and multicopter are interfered with by concrete with reinforced steel rebar construction (rebar) of most buildings. Lastly, most multicopters do not perform well in a Global Positioning System (GPS)-denied environment. While there have been major improvements in how some multicopter systems employing computer vision can succeed in object avoidance and positionings unfortunately, rubble environments still thwart efforts to determine a precise position as is required for confined space flight.

Currently, the most effective means for locating survivors in rubble are trained search dogs. Canine teams are comprised of a trained search dog with human handler. They are excellent at finding survivors. Dogs can quickly search an area and provide an indication of the presence of live people and can even determine their numbers. This is due to their acute sense of smell and agility. Canines primarily use their olfactory sensory system to track and identify items of interest to them as opposed to their visual sensory system [18]. Their sense of smell can be used to great advantage, as a dog's olfactory system is extremely sensitive—far beyond the ability of any human or human-made instruments. Also, this ability to smell what they are interested in allows dogs to find targets that are not visible—such as victims who are fully covered by rubble. The use of dogs as a scent detecting systems is very common and used in various applications. Some examples of this include detecting explosives [19], drugs [20], and even cancer [21].

The challenge with using dogs for this type of work is that they cannot communicate the location of survivors if the handler cannot be physically present, nor can they indicate the condition of the survivors or the state of the supporting structures around the victims.

A more effective system would provide first responders with improved situational awareness—identifying dangerous areas, confirming the exact location of survivors and determining their conditions. If there are trapped people, a map of survivors’ locations and data concerning the state of surrounding structures would be very useful. Presumably, such a system would also allow first responders to interact with the environment and survivors within it.

### **1.3 Objectives**

The purpose of this research work is to develop new systems that on top of the ideas presented in the work of Tran and Ferworn [22-25]. The original concept is to merge the abilities of canines and robotic telepresence technologies to create systems that address the weakness of existing systems while leveraging the areas where such systems already perform quite well. The system called Canine Augmentation Technology (CAT) is a wireless video, audio, telemetry and sensing system intended to be worn by USAR dogs while actively searching for survivors in areas where their handler cannot follow [22, 25, 26]. The CAT system has been shown to be successful in imparting the sensed, real-time situation of a dog—through telepresence—to a remote human who can use this information to help plan a rescue. The Canine Remote Deployment System (CRDS) [27] was also successful in allowing canine handlers the ability to have limited but effective physical interaction with a disaster environment by being able to deploy packages from a search dog. Both systems have problems with wireless communication, since USARs environment are inherently difficult for wireless to penetrate.

The goal of this work is to address the limitations of both CAT and CRDS. This research makes the following contributions:

- A robust bark detecting algorithm and hardware that guarantees that the CRDS will deploy a package even if wireless signals are completely blocked,
- A new system that allows specially built robots to be deployed from a searching dog (as well as a design of such robot). This would significantly improve the level of interaction that a rescuer can have with the disaster environment and survivors within,
- A study of the use of RGB-D cameras for reconstructing a disaster scene in simulation,

- Visual odometry is the process of estimating the position and orientation of an agent with respect to a starting position. An algorithm that allows low-quality blurred images, recorded from a dog, to be used by computer vision visual odometry algorithms. With the use of visual odometry, it is possible to reconstruct the path of a search dog from outside the disaster in order to assist in locating of victims. In addition, it is also possible to build a 3-dimensional (3-D) model of what the environment looks like.

#### 1.4 Structure of this Dissertation

This dissertation is organized around the following chapters:

Chapter 2 – *Remotely Affecting Disaster Incidents*. This chapter is based on the contributions of the following published work [28]:

- J. Tran, A. Ferworn, “Bark Indication Detection and Release Algorithm for the Automatic Delivery of Packages by Dogs,” in 6th International Wireless Communications and Mobile Computing Conference (IWCMC 2010), June 28 – July 2, Caen, France, 2010

The primary author of this paper is Mr. Jimmy Tran. Mr. Tran’s involvement in the development of the manuscripts includes: primary research, algorithm development, hardware development, conducting experiments, and verifying results. Dr. Ferworn’s involvement in the manuscript includes: supervision of the research process, and review of the publication.

Chapter 3 – *Extending a Rescuer’s interaction Within a Disaster Environment*. This chapter is based on the contributions disclosed in the following published works [29, 30]:

- J. Tran, A. Ferworn, M. Gerdzhev, D. Ostrom, “Canine Assisted Robot Deployment for Urban Search and Rescue,” in IEEE International Workshop on Safety, Security & Rescue Robotics (SSRR-2010), July 26 – July 30, Bremen, Germany, 2010

- M. Gerdzhev, J. Tran, A. Ferworn, D. Ostrom, “DEX – A Design for Canine-Delivered Marsupial Robot,” in IEEE International Workshop on Safety, Security & Rescue Robotics (SSRR-2010), July 26 – July 30, Bremen, Germany, 2010

The main contributor of these papers is Mr. Jimmy Tran. Mr. Tran involvement in the development of the papers include: primary research, concept development, hardware development, conducting experiments, and verifying results. Mr. Gerdzhev’s involvement in the manuscripts includes: setting the Linux Operating System and Streaming software on testing prototype. Mr. Ostrom’s involvement in the manuscripts includes: machining some parts of the prototype. Dr. Ferworn’s involvement in the manuscript includes: supervision of the research process, and review of the publication.

Chapter 4 – *3D Disaster Scene Reconstruction with RGB-D Cameras*. This chapter is based on the contributions disclosed in the following published works [31, 32]:

- J. Tran, A. Ufkes, M. Fiala, A. Ferworn, “Low-Cost 3D Scene Reconstruction for Response Robots in Real-time,” in IEEE International Workshop on Safety, Security & Rescue Robotics (SSRR-2011), Nov 1 – 6, Kyoto, Japan, 2011
- A. Ferworn, J. Tran, A. Ufkes, A. D’Souza, “Initial Experiments on 3D Modeling of Complex Disaster Environment using Unmanned Aerial Vehicle,” in IEEE International Workshop on Safety, Security & Rescue Robotics (SSRR-2011), Nov 1 – 6, Kyoto, Japan, 2011

The main contributor of these papers is Mr. Jimmy Tran. Mr. Tran’s involvement in the development of the papers includes: primary research, concept development, algorithm development, hardware development, conducting experiments, and verifying results. Mr. Ufkes’ involvement in the manuscript includes: co-implementation of the software used to collect and test data. Dr. Fiala’s involvement in the manuscript includes: supervision of the research process. Dr. Ferworn’s involvement in the manuscript includes: supervision of the research process, and review of the publication.

Chapter 5 – *Visual Odometry from Searching Canines with RGB-D Sensors*. This chapter is based on the contributions explained in the following published work [33]:

- J. Tran, A. Ufkes, A. Ferworn, M. Fiala, “3D Disaster Scene Reconstruction Using a Canine-Mounted RGB-D Sensor,” in *Computer and Robot Vision (CRV)*, 2013 International Conference on, May 28 – 31, Regina, SK, Canada, 2013

The main contributor of this paper is Mr. Jimmy Tran. Mr. Tran’s involvement in the development of the paper includes: primary research, concept development, algorithm development, hardware development, conducting experiments, and verifying results. Mr. Ufkes’ involvement in the manuscript includes: co-implementation of the software used to collect and test data. Dr. Fiala’s involvement in the manuscript includes: supervision of the research process. Dr. Ferworn’s involvement in the manuscript includes: supervision of the research process, and review of the publication.

Chapter 6 – *Summary and Conclusion*. This is the concluding chapter that summarizes the results and key findings of the research work. It presents various conclusions, limitations within various areas, and discusses directions for possible future research.

## **CHAPTER 2     Remotely Affecting Disaster Incidents**

This chapter is based on the contributions of the following published work [28]:

- J. Tran, A. Ferworn, “Bark Indication Detection and Release Algorithm for the Automatic Delivery of Packages by Dogs,” in 6th International Wireless Communications and Mobile Computing Conference (IWCMC 2010), June 28 – July 2, Caen, France, 2010

The primary author of this paper is Mr. Jimmy Tran. Mr. Tran’s involvement in the development of the manuscripts includes: primary research, algorithm development, hardware development, conducting experiments, and verifying results. Dr. Ferworn’s involvement in the manuscript includes: supervision of the research process, and review of the publication.

### **2.1 Introduction**

Telepresence systems aim to replicate the experience of being in a remote location to its users—that is stimulating all the human senses: sight, touch, hearing, smell, taste. This is prohibitively difficult as some experiences are not easily replicated with the resources available, and therefore most telepresence systems focus what can be seen, as sight is the dominant sense for humans. For the application of USAR operations, rescuers would like to be able to influence the disaster environment, mostly interacting with the trapped patients and their surroundings. Some rescue robots have complex manipulators and two-way audio and video technologies, but as discussed in Section 1.2, these robots cannot reach the places that rescuers need to access and know about. This chapter describes an innovative approach that allows rescuers the ability to affect the disaster incident and interact with trapped patients.

### **2.2 Related Work**

This section discusses seminal work that pioneered the concept of dogs as autonomous vehicles. The worked done by Ferworn et. all [22-25] addressed the deficiencies common in Unmanned Ground Vehicles (UVG) in their very limited traversal capacity by utilizing search dogs equipped with telepresence robotics technologies. The work eventually developed the CAT system. The metaphor behind CAT is the realization that search dogs are a class of autonomous robot. A search dog has the ability to climb most rubble piles and is “trained/programmed” to use their incredible sense of smell to find trapped victims. Using this metaphor allowed the researchers

involved in CAT to ignore the mobility problem and concentrate on providing telepresence services.

CAT developed over many iterations starting with a single head-mounted, analog camera system affixed to a set of canine protective goggles with an effective wireless transmission range of approximately 25m. Figure 2.1 shows an image of the first version of CAT. This initial version suffered from not only a range issue but also the camera placement. Search dogs perform their searches with their noses pointing to the ground, making the camera view to miss important environmental information. CAT evolved into a twin camera system mounted on the shoulders of the dog and streamed over a wireless Institute of Electrical and Electronics Engineers (IEEE) 802.11.x network (also known as WiFi) [34] allowing an operating range of approximately 100m. Figure 2.2 is a picture of CAT version 2 where the twin camera system was tested. The system demonstrated that two cameras were more effective at capturing the dog's surroundings. Figure 2.3 shows an image of CAT version 3 where the camera both recorded on board and stream the images via WiFi to the operator.



Figure 2.1 – USAR canine wearing CAT dog goggles at exercise in Fergus, Ontario 2006



Figure 2.2 – USAR Dog Dare wearing CAT 2.0 prototype

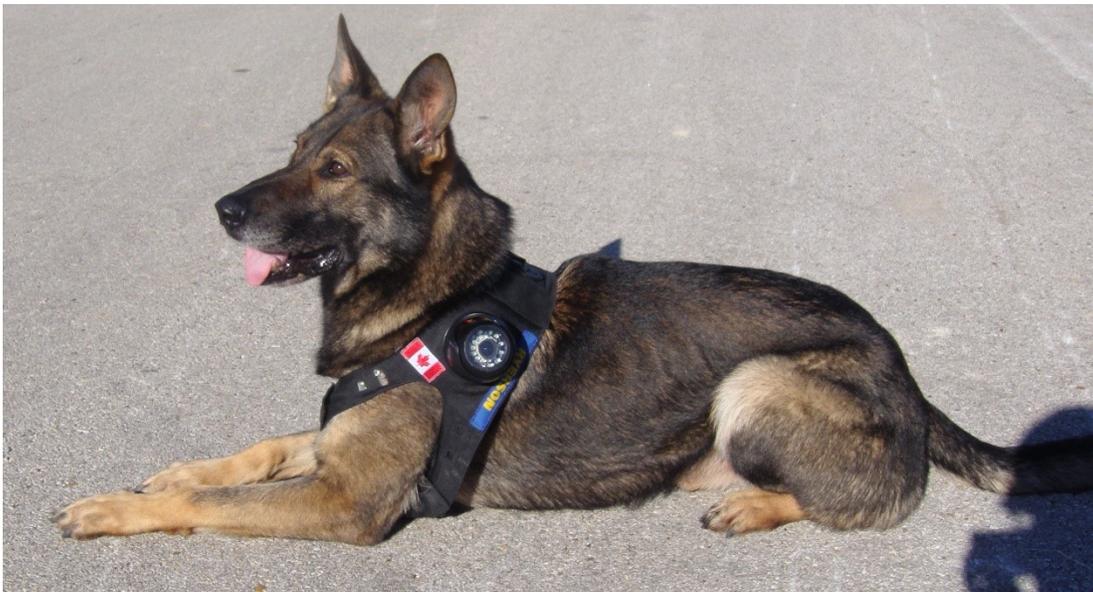


Figure 2.3 – CAT 3.0 worn by CA-TF1<sup>1</sup> USAR Dog Freitag

---

<sup>1</sup> FEMA designation meaning “California Task Force 1”.

CAT was tested and used in training by Provincial Emergency Response Team (PERT) which was renamed the Urban Search and Rescue (USAR) and Chemical, Biological, Radiological, Nuclear explosive (CBRNe) Response Team (UCRT) of the Ontario Provincial Police (OPP). All UCRT members are provincial police officers who have received specialized training in their respective job functions and have participated in exercises and emergency deployments making them, arguably, the expert practitioners in the Province of Ontario. Members are cross-trained as Hazardous Material and CBRN Technicians and USAR workers including training as structural collapse specialists, search methodology and technology and rope/technical and confined space rescue. Members are responsible for maintaining a high level of expertise in these specialized areas as well as a very high level of physical fitness to meet the demands of their team [35].

The response of the UCRT members who used the system on several exercises has been positive. The system provided them with much more information of the disaster than they would receive from just a canine team alone. One major complaint was that the videos from a searching dog was very fast moving, “bouncy” and generally disorienting. While operators could now see what was occurring around a trapped patient, the system did not provide them with any means of interacting with the patient.

Shortly after the development of CAT it was recognized by the research team and the UCRT that some form of interaction was necessary between the searching dog and the trapped victim that would provide some form of immediate assistance to the patient. Ferworn et. al. [27] devised and introduced the CRDS. The impetus for the CRDS was to provide the rescuers’ a tool to help increase victim survival rates during the often long interval between discovery and rescue. There have been cases where it has taken days to reach a victim [36]. As described in [27], the CRDS utilizes the agility and training of USAR dogs to deliver food, water, medical supplies and communications devices to victims trapped in areas of an urban disaster incident. The supplies are stored in a bright orange bag that the dog carries under its belly called the “underdog” as shown in Figure 2.4.



Figure 2.4 – Canine equipped with CRDS and underdog

USAR dogs are trained to provide a “bark indication” when they find a live victim. The sustained barking (called a “bark indication”) continues for some time as the dog remains where it “found” the victim. Taking advantage of the bark indication, the canine handler, equipped with a remote control, wirelessly activates the CRDS carried by the dog and releases the underdog, depicted in Figure 2.5. Figure 2.6 shows the components of the CRDS described in [27] with the exception of the remote.



Figure 2.5 – USAR Dog left underdog behind



Figure 2.6 – CRDS detail. From left to right: Remote, CRDS, harness, underdog

### 2.3 Continuing work on CRDS

Part of the work carried out in this dissertation was to take the original concept of the CRDS and develop it into a commercial product to be available for use by USAR responders wherever needed. The original CRDS prototypes were not designed for long and repeated use and did not stand up to the damaging conditions inevitably associated with disaster rubble. Thus, the final CRDS product was completely redesigned.

Since the device was meant to be worn by a dog, it was designed to be as small and compact as possible. The body is made from polyethylene terephthalate glycol-modified (PETG), an industrial plastic known for its durability, impact resistance, resistance to ultraviolet (UV) light, oils, and greases. A picture of the redesigned commercially available CRDS is shown in Figure 2.7. The release mechanism was also redesigned to release sideways instead of upward as shown in Figure 2.8. This means that when released, the device did not expand upward as in the original design, which could cause the dog that wears the device to have a greater chance of snagging the mechanism on the rubble and becoming “hung up”—unable to proceed to return to the handler.

Aside from the change in physical design, the electronics were also modified. The most notable change is within the wireless module. The final design uses a wireless XBee radio made by Digi. This radio uses the IEEE 802.15.4 [37] wireless protocol on the 900Mhz transmitting frequency. Part of the IEEE 802.15.4 protocol allows the wireless signal to be encrypted and the 900 Mhz signals are commonly known to have penetrating power.

The commercial version of the CRDS has a range of 140m in an indoor/urban environment and 3 km range line-of-sight or outdoor. The range can be greatly improved if a high gain antenna is used. While the release mechanism, the main unit worn by the dog, is powered by a small 3.7V lithium polymer (LiPo) battery with the capacity of 400mah. The CRDS can perform over 1000 releases on a single charge. A block diagrams of the CRDS main unit and remote unit are shown in Figure 2.9 and Figure 2.10.



Figure 2.7 – Latest version of CRDS in Retail Package



Figure 2.8 – Latest version of CRDS with tactical vest, demoed on model dog

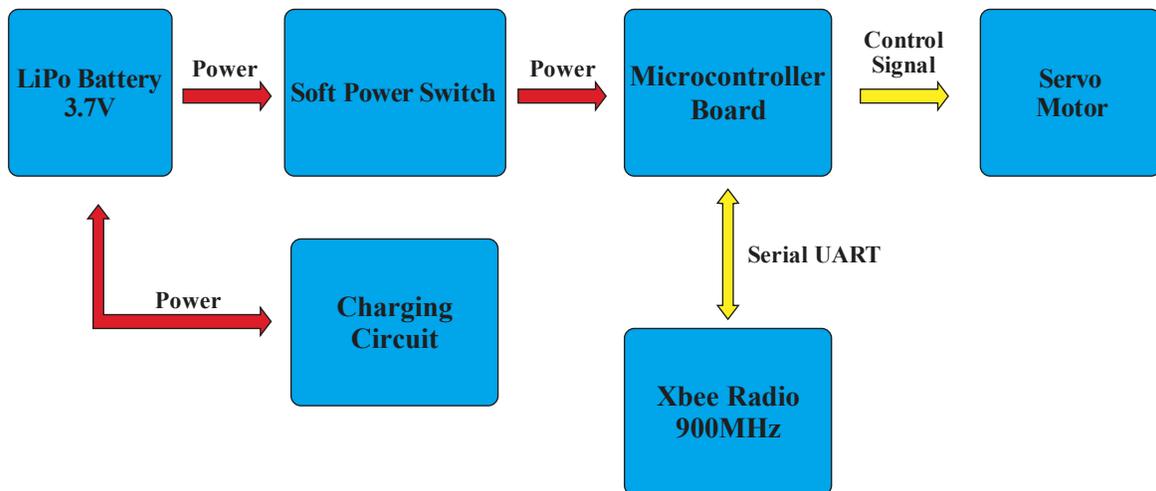


Figure 2.9 – Block diagram of the electronics components of the CRDS main unit

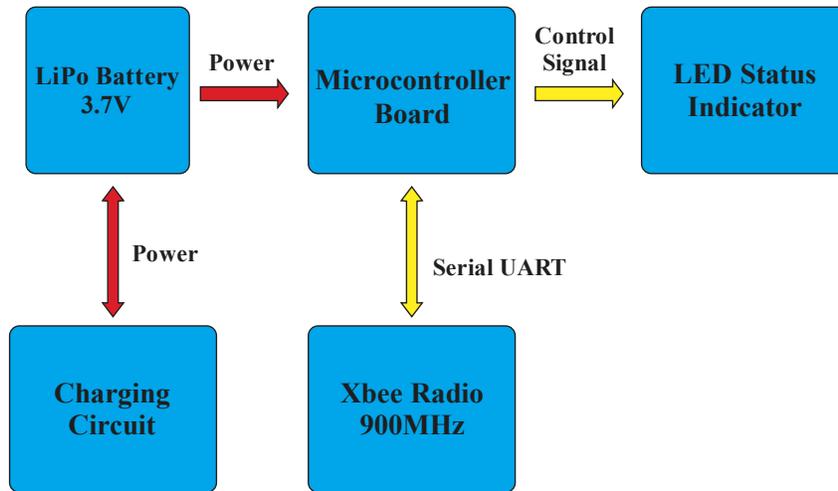


Figure 2.10 – Block diagram of the electronics components of CRDS remote unit

The CRDS has been tested by the UCRT in numerous USAR training exercises in Canada and the United States and has been used by a number of FEMA’s canine search teams. The wireless transmitter release mechanism has been demonstrated to work effectively in most cases. There are however two conditions which may cause this version of the CRDS to fail. The first is when the dog goes outside the audio range of its handler. If a handler cannot hear the dog’s barks, they do not know when to release the underdog. This situation may occur when the dog moves deep into inaccessible rubble where structural elements dampen the sounds of barking. The second condition is when the dog goes out of wireless range of the transmitter used to release the underdog. This includes cases where, for example, debris from a collapsed building may act as a barrier to wireless signals (Steel reinforced concrete usually exhibits this) [38].

In addition to these problems, it was observed that individuals who are designated to use the transmitter to release the bag often find the task very stressful. The reason being because, if they cannot see the dog, they do not know if their attempt to release the underdog was actually successful until the dog returns. In fact, a typical means of using the CRDS is for the handler to use an assistant to press the button on the transmitter as the handlers generally reported that they are stressed enough simply controlling the searching dog.

In order to resolve these problems, a bark detection system and deployment algorithm was devised for the CRDS—forming a bark release system (BARS). Such a system permits the deployment of the underdog without intervention from the handler in cases where the use of the transmitter system is undesirable or impractical.

## 2.4 Bark Detection

The proof of concept prototype of BARS was built from modifying the electronics circuits of a bark suppression collar—a training device for dogs. Later version was specifically designed with a microphone, signal amplification circuitry, microcontroller, wireless radio and LiPo batteries charging circuit. To filter environmental noise, the audio sensor is placed so that it presses against the dog's chest rather than being positioned outwards where it could receive spurious barking from other dogs and possibly misleading ambient noise. More importantly the sensor does not measure the frequency of the audio signal of a bark but instead measures the amplitude of the audio signal. This audio signal is not transmitted through the air but through the dog's body when it barks. Thus, the sensor only picks up signals related to barking from the dog wearing the collar.

However, a single bark is not a clear indication that the USAR dog has found a victim since all dogs routinely bark and often make other noises such as whining, whimpering or growling. The difference between USAR dogs and other dogs is that the USAR dogs are trained and rewarded for barking continuously when they have located the area of strongest human scent which usually gives a very good indication of the vicinity of the human victim.

To filter out false indications, an algorithm was developed that differentiates between single and multiple barks and between multiple barks in quick succession and random barks over long periods of time.

The following description of the algorithm use **X**, **Y** and **N** as parameters. After the first bark, a timer is started and will continue to run for **X** seconds. In the duration of **X** seconds if **N** more barks have not been detected then it reverts to the beginning state where it waits for another first bark again. If **N** barks are finally detected within the **X** seconds, then a delay of **Y** seconds is introduced followed by a signal to the CRDS to release the underdog. The algorithm is depicted in Figure 2.11.

Through testing, the following parameters seem to work best:

- **X** interval time = 8 seconds
- **N** MaxBarks = 4
- **Y** DelayTime = 5 seconds

The **Y** seconds delay was introduced as it was observed that the dogs would either stay where they are after barking or they actually try and move closer to the source of the scent. In either case,

extra delay only improves how closely the dog can deliver the package to the system and in no way diminish the performance of the system.

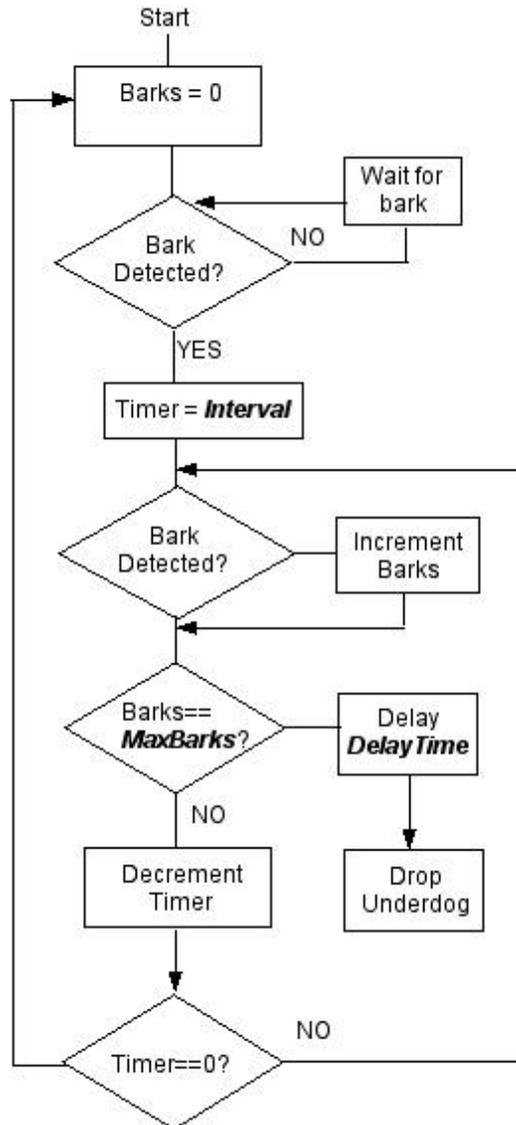


Figure 2.11 – Bark Release Algorithm

The algorithm was implemented on a microcontroller and introduced to the CRDS. The electric circuitry of an anti-bark collar was removed from the unit. In its originally intended use, the microphone circuit output was to drive a solenoid valve that was intended to spray a harmless but annoying puff of air or chemical to deter the dog from barking. After probing the circuit, it was discovered that each bark is detected by a microphone and amplification and signal filtering circuit. The output is a digital signal that activates a transistor driving the solenoid valve. For a quick

prototype, the circuit was to output to a microcontroller programmed with the described algorithm. Once the algorithm verifies the bark indication, a signal is sent to the CRDS release mechanism triggering the release.

Once the proof of concept was verified, a design was developed without modifying a commercial product. A block diagram shown in Figure 2.12 consists of a microphone and amplification circuit output an analog signal to a microcontroller. The amplification circuit utilize a standard op-amp and microphone combo (shown in Figure 2.13). The microcontroller only monitors the amplitude of the audio signal to see if it reaches a threshold value. When that threshold is reached, a single bark is registered. Similar to the previous system, after the algorithm determines that a true bark indication occurs, it would act as a transmitter and send the signal wirelessly to the main unit to trigger a release. With the wireless radio and battery, BARS is completely separated from the main unit, eliminating the use of cumbersome wires.

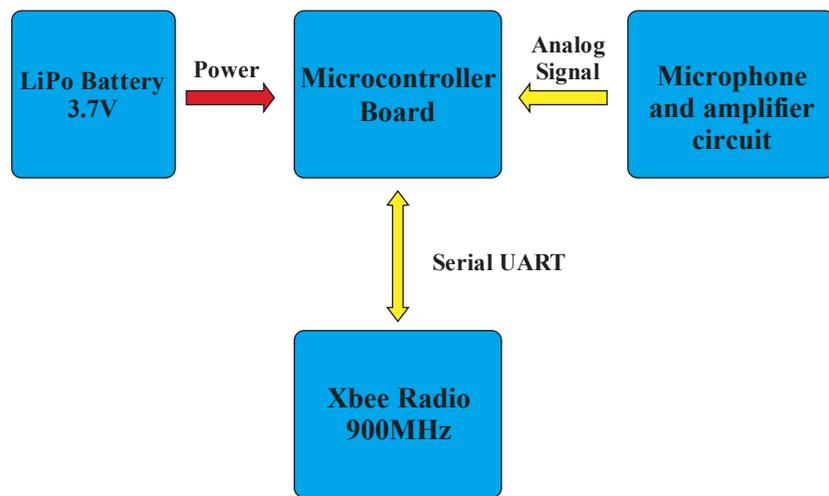


Figure 2.12 – Block diagram of the electronics components of Bark Release System

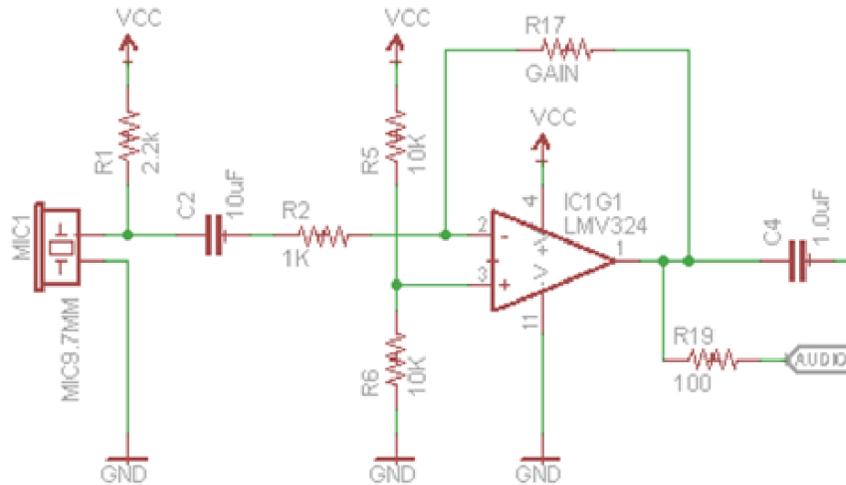


Figure 2.13 – Audio amplifier circuit schematic

## 2.5 Bark Release Experiments

Two sets of experiments were conducted to ensure the effectiveness and accuracy of the bark release CRDS. The first experiment was a proof of concept test and the second was designed to determine the expected accuracy of the bark release CRDS.

A typical exercise in the training of USAR dogs is called the “bark barrel”. A person to be “found” by the dog (called a “quarry”) is placed inside one of several large barrels and hidden from a subject USAR dog’s view as shown in Figure 2.14. The purpose of the bark barrel is to get a dog to bark when it smells a human in the barrel even though the dog cannot see the human. Eventually, the barrel is moved to a rubble environment and then removed altogether as the dog becomes used to finding hidden humans in voids and hidden spaces within rubble.

The bark barrel was used as the basis of the first set of tests as the barrel is a fundamental part of all USAR canine training and a good place to introduce the new equipment to the dogs for the first time. With the help of the canine handlers, the bark release version of the CRDS and underdog were worn by 4 different dogs while they performed the bark barrel exercise.



Figure 2.14 – Bark Barrel Test

The bark barrel exercises provided an opportunity to calibrate the CRDS to the barking characteristics of the dogs. This included determining the maximum number of barks that can reasonably be expected when a dog indicates the presence of a hidden human—MaxBarks, the determination of a reasonable interval within which that series of barks should be counted—Interval, and the delay to allow the dog time for finally positioning its body before the underdog is released—DelayTime.

The tests also provided an opportunity to properly fit the CRDS harness to the dogs and to ensure that each dog was not distressed by carrying the equipment or having a piece of it “mysteriously” fall off.

The second experiment was designed to ensure that the automatic bark release algorithm can deliver the underdog as close to the victim as possible employing the bark detection algorithm. The original CRDS performed well in previous tests with a human monitoring the barks and manually triggering the release. This experiment was setup to compare the performance of the manual release and automatic bark release. A quarry (human hiding from a dog) was placed in a rubble pile and concealed in a hole beneath a wooden pallet. The centre of the wooden pallet was marked with orange spray paint as shown in Figure 2.15 as a visual guide to the human observers on the rubble pile. Figure 2.16 shows a dog finding the quarry.



Figure 2.15 – Location on rubble pile where the quarry hid



Figure 2.16 – Dog finds victim

Three dogs were sent out, one at a time, to find the quarry. A human assistant was provided to the canine handler who was not able to see the dog searching or the pallet target. The job of the assistant was to listen for the dog barking and trigger the release of the underdog when they thought

the dog might close to the target. The distance between the centre of the target and the underdog was measured.

Following the manual release trial, the quarry was moved to a similar but different location on the rubble pile. This was done because the same dogs were used for the second trial and the dogs are smart and quickly learn the location of the quarry from the first trial and will run straight to them without actually searching in the second.

Similar to the previous trial, each of three dogs were sent to locate the victim except the dogs were equipped with the bark release CRDS. Again, each time, the distance between the centre of the pallet and the underdog was measured.

### 2.5.1 Bark Release Results

The bark barrel experiment was successful. The underdog was released at the appropriate moment (5 seconds after the 5th bark) on every test in all cases with all the dogs. The algorithm performed well under the different barking habits of the dogs. Some of them were excited and barked rapidly while others barked at a slower rate. The algorithm performed correctly in all cases.

In general, the dogs adapted to the unfamiliar harness and equipment quite well. One of the dogs had never worn the harness before and the handler had never seen the equipment. Yet the team was able to deliver the underdog successfully without any observable hesitation or distraction on the part of the dog.

The results of the accuracy experiments are presented in Table 2.1.

Table 2.1 – Distance from underdog to victim

	1st Dog	2nd Dog	3rd Dog
Manual Release	90cm	122cm	60cm
Automatic Release	75cm	75cm	360cm

The average distance from the victim to underdog in the manual release trial is 90.67cm while in the automatic release trial, the average distance is 170cm.

## 2.6 Conclusions

An important impetus for conducting these tests was to confirm that the equipment does not alter the search characteristics of USAR dogs. This is important because dogs do this task very well and, because they are so fast, they save lives by finding people in rubble very quickly. The tests confirmed that the CRDS did not hinder the dogs and provides them a new capability—automatically delivering supplies to victims.

This new functionality is significant in itself for a number of reasons. Firstly, this marks the transition of USAR dogs from the role of search to one of assisting in the rescue of humans. As it is unlikely that any artificial system is likely to move through rubble as quickly as a dog, this functionality makes dogs even more necessary for USAR work. Secondly, the dog now becomes part of the rescue management decision making process. It is the dog that decides where the trapped human is, makes the first contact and takes the first step towards rescue the victim. In effect, the dog is an agent assisting humans effectively using Biological Intelligence (BI).

It is also important to note the apparent anomaly in the presented data. It should be noted that the search dogs find their way to hidden humans using scent. They search for human scent and follow the scent trail until they find the highest concentration within the scent “plume”. This is when they give their bark indication. The scent plume can be drastically affected by wind and the configuration of the rubble. When examining the data, it may appear that the manual release of the underdog has a greater accuracy based on the longer distances from the target using automatic bark release. Drawing this conclusion would be incorrect. The hiding locations for quarries—called “hides”, are selected primarily for the ability of the quarry to actually occupy the position and remain hidden. The location of the first quarry was selected based on a convenient hole being near the centre of the rubble pile. There was no wind and the space occupied by the quarry was a culvert, sealed off from other parts of the rubble. Essentially, this meant that the dogs were not presented much of a challenge as the scent plume was emanating from a single location and was concentrated around the quarry’s location.

At the second hide the conditions were considerably different from the first. The second hide was located on a small hill formed by the rubble and the quarry was hidden in a hole that extended into the rubble pile and resurfaced at several locations around the pile. In addition, there was a brisk wind blowing across the pile that entered the hide and pushed scent through the holes in the pile. As the dogs were taking much longer to find the second hide, an investigation was conducted

to figure out why the dogs were having such a hard time finding the quarry. After the testing, a smoke grenade was thrown into the hide and the dissipation of the smoke was observed. It was discovered that the smoke was being discharged through at least four separate holes around the quarry's former location. In addition, the wind pushed the smoke (and presumably the human scent) over the hill and pooled it in a shallow bowl-like structure where one of the dogs eventually gave its bark indication and released its underdog. This situation is shown in the Figure 2.17.



Figure 2.17 – Circle indicates “hide” location, Square show pooling of smoke and, presumably, scent from the quarry

The test was considered valid by the canine handlers, as in their work environment, unpredictable wind conditions are a fact of life. The argument is that, if placed in the same situation, the human handler would still choose to release the underdog at the same place upon hearing the dog's barks. After reviewing the results of the experiments performed it can be concluded that the automatic bark detection feature is reliable and performs reliably with at least the same accuracy as manual release. At the same time, it works for situations where manual release would fail.

## CHAPTER 3     **Extending a Rescuer’s interaction Within a Disaster Environment**

This chapter is based on the contributions disclosed in the following published works [29, 30]:

- J. Tran, A. Ferworn, M. Gerdzhev, D. Ostrom, “Canine Assisted Robot Deployment for Urban Search and Rescue,” in IEEE International Workshop on Safety, Security & Rescue Robotics (SSRR-2010), July 26 – July 30, Bremen, Germany, 2010
- M. Gerdzhev, J. Tran, A. Ferworn, D. Ostrom, “DEX – A Design for Canine-Delivered Marsupial Robot,” in IEEE International Workshop on Safety, Security & Rescue Robotics (SSRR-2010), July 26 – July 30, Bremen, Germany, 2010

The main contributor of these papers is Mr. Jimmy Tran. Mr. Tran’s involvement in the development of the papers include: primary research, concept development, hardware development, conducting experiments, and verifying results. Mr. Gerdzhev’s involvement in the manuscripts includes: setting the Linux Operating System and Streaming software on testing prototype. Mr. Ostrom’s involvement in the manuscripts includes: machining some parts of the prototype. Dr. Ferworn’s involvement in the manuscripts includes: supervision of the research process, and review of the publications.

### **3.1 Introduction**

The approach of using dogs as autonomous vehicle analogs, as exemplified through the CRDS and CAT projects, provided rescuers access to victims in disaster rubble. One of the undesirable side effects of using dogs is that they must eventually leave the victims that they find. This may cause additional stress for the victims through loss of hope [39] as they witness the dog that just found them leave the scene. As the process of rescue can take quite a long time, the victim will have no indication that they may be saved.

Unlike dogs, rescue robots may linger with the victims, stay with them and provide a method for first responders to examine the scene close to the victims. The problem is that they cannot usually traverse the rubble to reach the victim. The chapter presents a method for extending the reach of robots through dogs.

## 3.2 Canine Assisted Robot Deployment

### 3.2.1 Unmanned Ground Vehicles in Urban Search and Rescue Operations

Rescue robots allow first responders to explore dangerous areas from a safe distance. The challenge for these robots is the unpredictable and harsh terrain. Some of these challenges were discussed in Section 1. This section describes these challenges in greater detail.

Various robotic designs have been proposed to address the mobility demands of rubble traversal. One popular design is called “variable geometry” in which the robot can shift its shape to accommodate its environment. Kenaf, developed at Tohoku University in Japan is typical of such robots. As shown in Figure 3.1, Kenaf is a 6 Degrees of Freedom (DOF) crawler track robot [40, 41]. Shape shifting is popular because it allows a robot to conform to a surface, allowing the operator to have fine control over which surfaces are in contact with the rubble. Examples of similar robots are common like Helios [42] and Souryu [43] as well as commercial robots like PackBot developed in various version by iRobot, VGTV-Extreme by Inuktun and Telemax by TelRob [44].

However, the disadvantage of these complex robots is their complex interface design used to control them. USAR operations inevitably involve complex terrain that is difficult for humans to traverse—much less ground robots. No level of autonomy has yet been demonstrated on a rubble pile. Therefore, response robots rely on teleoperation. Normally, the remote operator does not have a direct view of the robot. Because victims are buried in rubble, the working environment is usually total darkness. Illumination provides few visual cues concerning the environment the robot is intended to traverse and therefore it is difficult for an operator to determine an effective robot geometry to move through the environment.

Considering these factors, it is understandable that it is reported by Murphy [45] that effective operator situation awareness (SA) is difficult to achieve under the stressful conditions of USAR operations. With the addition of the pressure to avoid errors and operating while wearing personal protective equipment and gear, robot operators may experience cognitive overload [46].

Some of the problems that the robots that were deployed in the World Trade Center terrorist attack in September 11, 2001 experienced are described by Murphy [47]. Robots were actually lost in the rubble due to loss of wireless communication. Others were damaged due to the harsh environment. Typical damage and wear to robots is shown in Figure 3.2 and Figure 3.3.

A review of locomotion mechanisms of USAR rescue/response robots was conducted by Wang and Gu [48]. They indicated that for rescue robots in actual USAR situations, other factors that are problematic include finding methods for practical deployment, problems associated with travelling on aircraft<sup>2</sup>, decontamination and waterproofing. Their conclusion that developing a practical USAR robot will remain a great challenge for robot researchers remains true.



Figure 3.1 – Kenaf rescue robot with features including crawler enclosing its entire body and four sub-crawler arms that are each controlled independently [40, 41]

---

<sup>2</sup> Robots typically travel with spare batteries and parts and take up valuable space in aircraft that have limited space available for equipment.

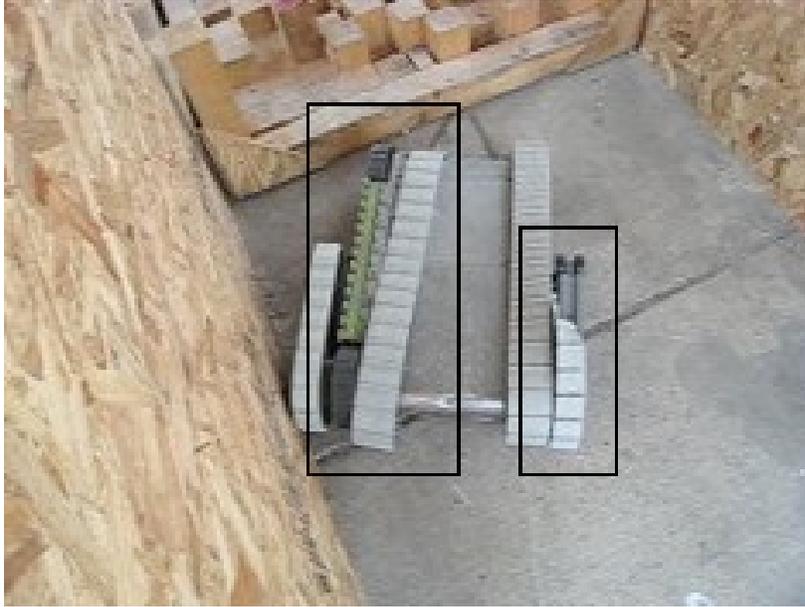


Figure 3.2 – Packbot having undergone track shedding after NIST Response Robot endurance testing



Figure 3.3 – Matilda robot with scratched and abraded armoured front visor. Damage caused by attempting to traverse rubble seen in background

### 3.2.1.1 Performance Metric of Rescue Robot for USAR

The Intelligent Systems Division of the National Institute of Standards and Technology (NIST) of the U.S. Department of Commerce, sponsored by the Federal Emergency Management

Agency (FEMA) and Department of Homeland Security (DHS), in a multi-year program, is investigating how to measure the performance of rescue robots. The main goal of the investigation is to determine how to evaluate robots in operation in an USAR environment [14] through the use of common metrics. NIST has developed performance standards for many categories of characteristics of robots [15] through the ASTM standards process, E54 Task Group [15-17].

Robots are evaluated on various performance characteristics including mobility, sensing capabilities and overall system performance (durability, communication, power requirements) as well as secondary physical characteristics like “Cache Packaging”—indicating the physical size of the robot and related equipment which must be transported by a task force to an operation.

While there are niche areas where robot mobility is quite impressive, their performance in traversing open rubble does not approach the mobility characteristics of dogs in and on rubble and is generally quite poor. This is unlikely to change in the near future. An idea of the difficulty of the problem can be gained by viewing a rubble field (see Figure 3.4) and its artificial equivalent NIST “step field” (see Figure 3.5). Typically, robot operators find traversing step field very challenging and rarely attempt traversing the actual rubble pile.



Figure 3.4 – Purpose built rubble pile for USAR training at “Disaster City”



Figure 3.5 – An example of a robot test “step field” in one configuration of a NIST mobility test

### 3.2.1.2 Marsupial Robot Operations

Another scheme for overcoming the rubble mobility problem is through the use of multiple robots acting in concert. The idea is that when one robot can no longer move because of a deficiency in its design, a partner robot may be able to succeed using the progress made by the previous robot. One common form of cooperation between response robots is marsupial operation [49, 50].

Marsupial operation seeks to take advantage of the fact that different robot designs have different strengths and weaknesses. By combining the strengths of one robot, the weaknesses of another robot can be overcome as multiple robots interact to complete complex tasks. For example, a large robot can scale terrain with big steps while a small robot can enter small openings. If the large robot carries the small robot and places it for entry in a small opening, the large robot is said to be using marsupial delivery when it places the small robot. The concept is explained in detail by Ferworn [49]. In practice, marsupial delivery is not commonly used in actual disasters because emergency task forces typically do not carry compatible robots, nor do the robot operators practice this rather complex operational skill.

### 3.3 Canine Assisted Robot Deployment Framework

This section introduces the Canine Assisted Robot Deployment (CARD) framework and why it was developed. CARD takes its inspiration from Canine Augmentation Technology (CAT) [23, 24] and the Canine Remote Deployment System (CRDS) [27].

The current version of CAT [51] is equipped with two cameras to transmit video wirelessly from the dog to operators, as well as on-dog recording. CAT was the original system of this set of technologies to use the biological intelligence of dogs to meet the challenges of USAR. Videos from CAT can provide rescuers valuable information about the disaster. However, CAT's video is often jittery since the camera is mounted on a dog bouncing around as it moves through rubble and it is not possible to focus on individual details for long periods of time as the dog is constantly moving. Another challenge is that search dogs do not stay with the casualty. After they bark to indicate live human scent, they are recalled to perform other searches—leaving the victim alone.

The CARD system essentially uses the CRDS to carry small teleoperated robots under a searching dog for deployment when a victim is found. The advantage of CARD is that after the robot is deployed from the dog, the system produces a very stable video feed for a long period of time in order to monitor the environment the victim is in. The robot stays with the victim until the victim is rescued and the robot retrieved.

The method of robot deployment depends on the CRDS. For CARD, the CRDS is used to deliver a small robot. An additional feature of the CRDS is its ability to detect a dog's sustained barking and to initiate the release mechanism after the constraints of an internal algorithm have been met [28].

The CRDS can also be used for the deployment of a wireless mesh network (WMN) made up of a number of mesh routers or “bread crumbs”. Previous work has suggested that WMNs would be appropriate to use in a structural collapse environment for USAR [19] where the ambient material of the collapse provides challenges for radio communications which cannot be overcome in any other way.

Using bark detection release of the CRDS mechanism is useful in situations where the dog is outside of the range of the wireless transmitter. It is especially useful in deploying the first

“breadcrumb” right by the casualty. A trail of “breadcrumbs” can be deployed along the path dogs<sup>3</sup> take to find a casualty—eventually creating a WMN from the rescuer outside the collapsed site to the casualty.

In summary, the framework for CARD begins with the deployments of mesh nodes creating a WMN. This process starts with the first node being deployed by a dog with the bark detection release at the casualty’s location. Once the WMN is established, another dog can be sent to the victims site carrying a robot. After being deployed from the dog, the robot can do its job of collecting data and sending sensed data back through the WMN.

Based on the framework described, there are four requirements for robots that could be deployed from search dogs in this manner. These are discussed below. In this paper, the class of robots that are designed to be deployed from search dogs are called Canine-Delivered Marsupial Robots (CDMR).

The first requirement is that the robot’s size must be small enough such that a dog can carry it underneath its belly between its hind and fore legs comfortably. Otherwise it would hinder the dog’s ability to perform its task.

The second requirement is that the robot must have some form of useful mobility when it is delivered. Even though the robot is not required to traverse far from where it is deployed, the environment is still rubble. The robot should be able to survey around it and travel short distances to establish a good observation point.

The third requirement is that the robot must be WMN compatible and that all on-board sensors must have their data communicated using this means.

Finally, any robot that is carried must be rugged. Dogs normally travel very quickly, will smash the robot against very hard objects and will rub against very abrasive surfaces. Any robot being carried must be able to withstand this means of transport.

---

<sup>3</sup> Dogs are in the plural because more than one search dog may be required to complete a WMN deployment.

### 3.4 Canine-Delivered Marsupial Robots

#### 3.4.1 Drop an Explore (DEX)–A Design for Canine-Delivered Marsupial Robot

A prototype of a small robot that is deployable from a search dog called “Drop and Explore” (DEX) was developed. DEX was used in a series of experiments designed to validate CARD as a viable method to enhance USAR search operations. This section is based on the work that describes the design concept and construction of DEX 1.0 [30], shown in Figure 3.6.

DEX was designed with several constraints in mind. The first is that it must be able to fit under a searching dog in a way that would not encumber the dog. It was hypothesized that a flat top would fit flush to the dog’s underbelly making it more acceptable for the dog to carry. The qualifier “acceptable” is preferred over “comfortable” to avoid anthropomorphizing the dog’s reaction to the robot.

The term *ergonomic* is used to describe the process of designing or arranging workplaces, products and systems so that they fit the people who use them. Ergonomic is directly related to the safety and comfort of people in work place. A useful term in this research is “dogonomics” which is ergonomics for dogs. This is important since the dog must concentrate on its task and not be distracted by the equipment it happens to be carrying.

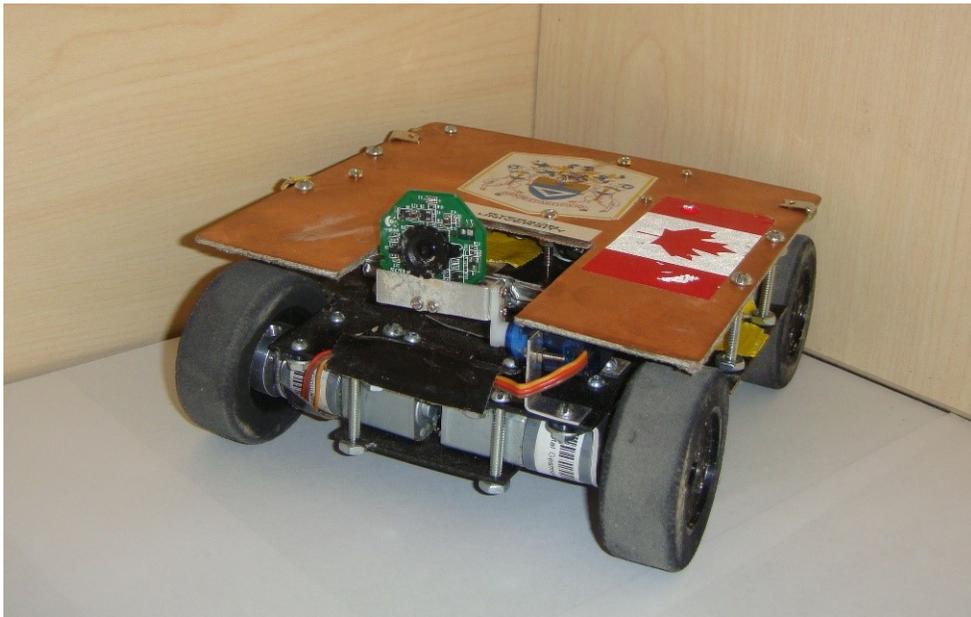


Figure 3.6 – DEX 1.0, the first prototype

These constraints limited the choices for all the other components such as wheel size. This reduced the robot’s mobility over rougher terrain but, since it was never the goal of DEX’s design to provide high-mobility, DEX’s mobility was acceptable since the robot would be dropped very close to the detected human victim. DEX also would not need to move very far to capture the data required. For simplicity and ruggedness, DEX was designed to be a 4-wheel drive vehicle to enhance its mobility as much as possible.

DEX’s small footprint also limits the electronics that can be used. Processing power is reduced because of this constraint as well as battery power. More details about the robot are presented below.

**Dimensions**

Dogonomics is an important consideration and thus part of the design is to minimize the physical size of DEX. In reality, not all dogs are suitable for this type of work. USAR dogs come in many shapes and sizes. As a rough gauge, Labrador and Shepherd breeds work very well in terms of size and build. Table 3.1 lists the dimensions of DEX. A common heuristic for this type of work is: the larger the size of a robot, the less accepting a dog intended to carry it is likely to be. Inevitably, robots strapped to a dog will encumber its mobility to some extent.

Table 3.1 – Dimensions of DEX 1.0

Length (cm)	Width (cm)	Height (cm)	Wheel Diameter (cm)
18	15.5	6.5	5.5

An important part of the deploying mechanism of CARD is how the CRDS, the dog and DEX would be attached to each other. Two methods were developed for attaching DEX to the release mechanism.

**Methods of attachment**

1) Method 1: retractable cable mechanisms, attached hooks to robot corners and mounted them on top of the robot. Figure 3.7 shows this on the robot and Figure 3.8 shows how the system is attached.



Figure 3.7 – DEX with retractable hook

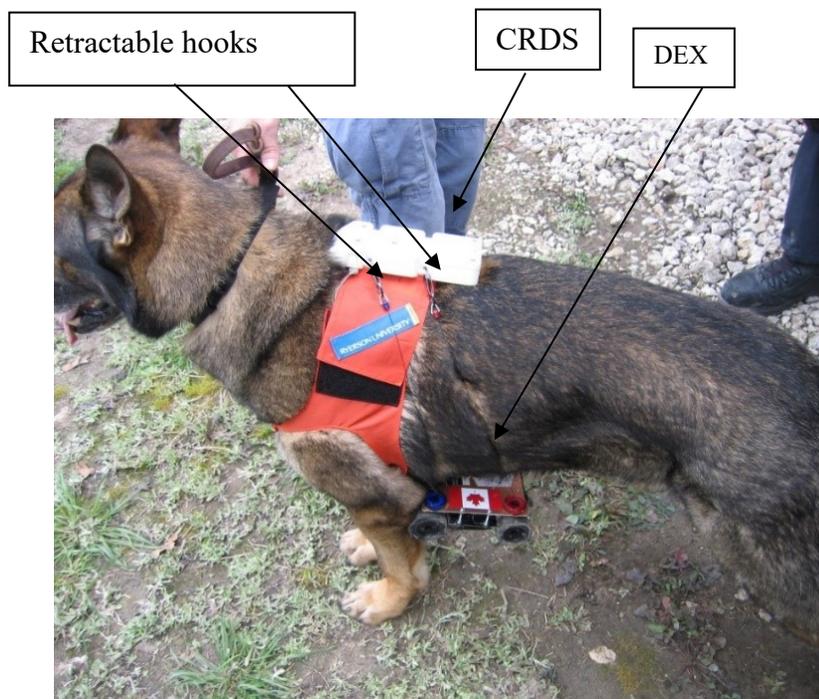


Figure 3.8 – DEX attached with hooks to CRDS

2) Method 2: The robot was nestled inside the underdog (the bag that goes under the dog that is designed to carry supplies to the victim). When the CRDS is triggered, the bag is released along with DEX Figure 3.9.

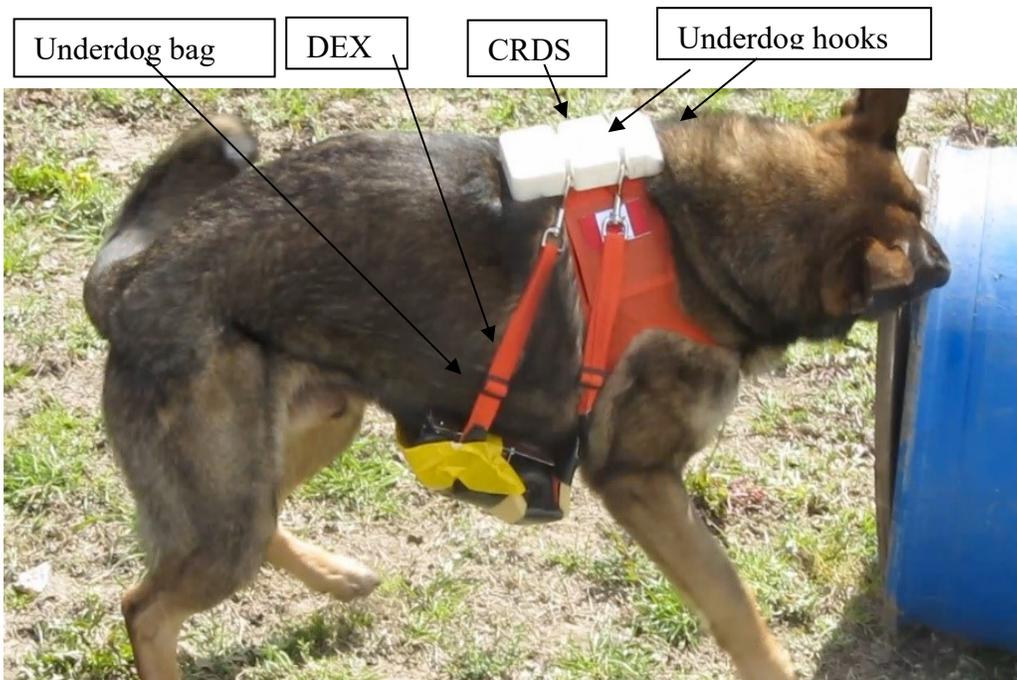


Figure 3.9 – DEX in underdog attached to CRDS

#### Electronics Hardware

The electronics components of DEX includes:

- Single board computer (SBC) – Beagleboard
- Camera – Logitech C200 webcam
- 4 ports USB Hub
- WiFi adapter – D-Link DWL-G122
- Switching voltage regulator
- Microcontroller board based on Atmel’s ATMega328
- Motor driver – Pololu TB6612FNG Dual Driver
- LiPo batteries

The Beagleboard is an ARM based 600MHz computer with a Digital Signal Processing chip. This is the main control unit that captures videos from the camera and transmits it over 802.11 (WiFi) WMN. It is also responsible for receiving control signals from another computer and passing it on to the microcontroller for mechanical functions.

The microcontroller is also connected to the motor driver which provides power to four geared motors. Aside from the drive motors, a servo motor was used to raise and lower the camera. The servo motor is controlled by the microcontroller.

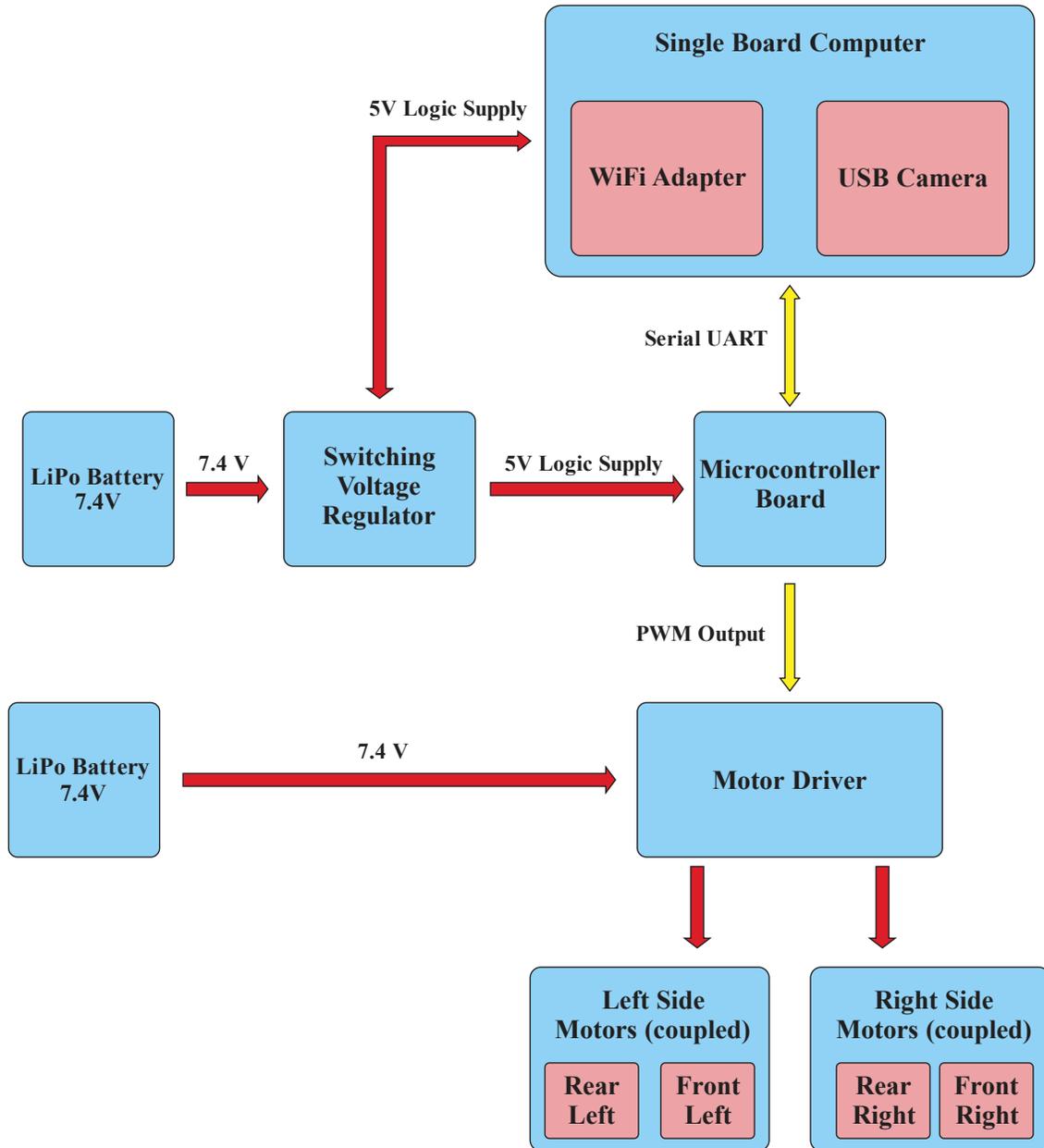


Figure 3.10 – Block Diagram of DEX 1.0 components

DEX is powered by two separate packs of Li-Po batteries. Each battery pack is made from two cells in series providing a nominal voltage of 7.4V and has the capacity of 2000mAh. Since

all the electronics operates on a 5V logic level, a switching power regulator was used. The second battery pack is used just to power the motors. This design of two separate packs helps reduce the electrical noise from the motors. A block diagram of DEX's system is shown in Figure 3.10.

### **Control and communication**

The robot is controlled through the 802.11 standard over a WMN. The reasoning that the network can be easily extended by dropping more nodes and thus allowing for connectivity in places where normally no wireless signals could penetrate, such as around bends in rubble made out of reinforced concrete.

The robot is controlled from a laptop computer on which the user interface displays a live audio/video stream from the camera. A joystick-like interface is used for easier control of the robot movement and 2 buttons for tilting the camera up and down. When a command is sent to move the robot, it goes through several “hops” of the mesh network to the robot's computer which, in turn, interprets the command and sends it to the microcontroller via a serial interface that controls the individual motors.

The video feed from the camera is recorded on a local flash drive—specifically a Secure Digital (SD) card on the Beagleboard. The video is also streamed through the WMN to be displayed on the laptop screen where search managers and first responders can analyze the information.

### **3.4.2 DEX Field Experiments**

A series of experiments were conducted at “Disaster City” [52] on the grounds of Texas A&M University. These experiments were designed to demonstrate a proof-of-concept for CARD and reported in [30]. Additional experiments were conducted on rubble elsewhere. These were end-to-end tests of the system in a realistic simulated environment. The test site is a training facility for the USAR canine unit of the OPP located in Bolton, Ontario.

#### **3.4.2.1 DEX Deployment Mechanism Experiment**

A typical exercise in the training of USAR dogs involves an apparatus called a “bark barrel.” A person to be “found” (called a “quarry”) by the dog is placed inside a large barrel hiding them from view. The purpose of the bark barrel is to get a dog to bark when it smells the human in the

barrel rather than when it sees them. Eventually, the barrel is moved to a rubble pile and then removed altogether as the dog becomes used to finding hidden humans in voids and spaces in rubble piles through their scent alone.

The initial experiment was designed to test the mechanics of a robot being dropped from a dog and its survivability. In the experiment, an USAR dog performs the bark barrel exercise wearing the CRDS equipment with a CDMR robot attached. After the dog barks, the robot is dropped from the dog.

### 3.4.2.2 Tunnel Experiment

Secondary experiments were conducted in tunnels under a rubble pile purposely made for USAR training at Disaster City, shown in Figure 3.11. This experiment was designed to test a CDMR's ability to capture videos of a victim within the tunnels. The scenario of the experiment required a quarry to hide somewhere in the tunnel under the rubble. Visual acuity stickers were placed around the location of the quarry for the purposes of determining if video would be stable and interpretable by humans. After the quarry was placed, the dog was sent in to find them. When the dog barked, the robot was deployed. The robot then explored the area as much as its limited mobility would allow. Figure 3.12 illustrates the size of the tunnel and placements of the mesh routers and victim.



Figure 3.11 – Rubble Pile at Disaster City

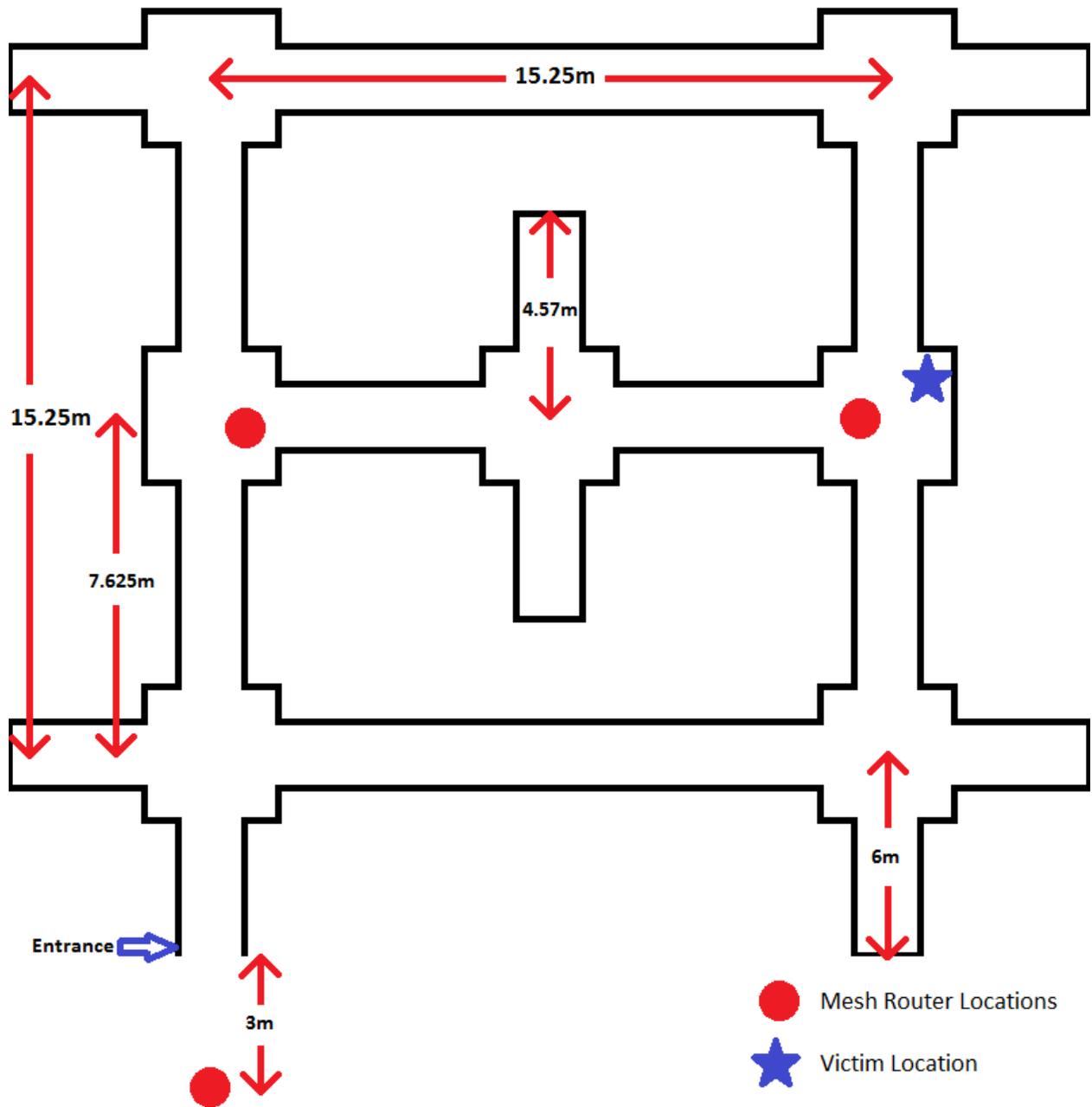


Figure 3.12 – Map of tunnel under rubble

Below are some samples of visual acuity stickers placed in the tunnels.



Figure 3.13 – Centre-fire Shooting Target

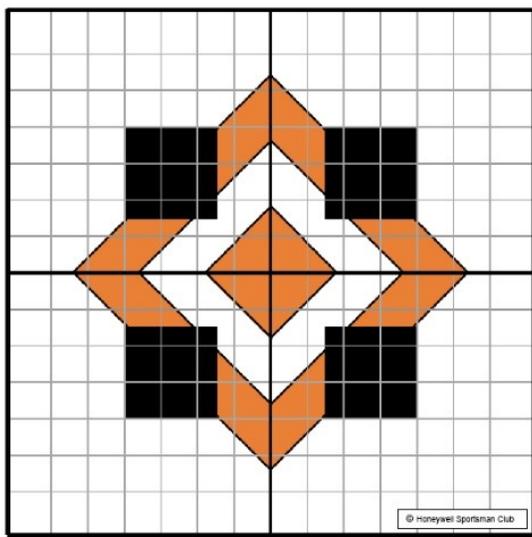


Figure 3.14 – Dot-mil Confidence Target

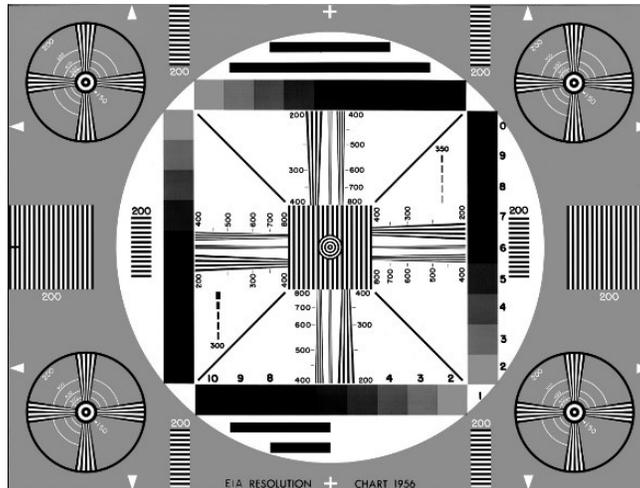


Figure 3.15 – IEA Resolution Chart 1956

### 3.4.2.3 Rubble Experiment

To validate that CARD can be used in USAR operations, an experiment was conducted in a similar fashion to how a professional USAR canine team training exercise would be conducted. A standard training exercise for the OPP's USAR canine teams is to have a quarry hide in a rubble pile inside a concrete tube completely covered and out of site. After the quarry is set in place, the handler brings the dog out to conduct the search. The handler remains on the ground away from the rubble while the dog searches on top of it. When the dog gives its bark indication, it is rewarded if the bark indication is given near the hidden quarry. The OPP has a purpose-built rubble pile in Bolton, Ontario, Canada. It is approximately 30m by 60m in size. Figure 3.16 shows a top view picture of the training site and Figure 3.17 the side view.

This experiment was conducted in the same way as a canine team training exercise, with a few additions. Each search dog is equipped with the CRDS and DEX strapped to its underbelly. When the dog barked, indicating a find, the robot operator would activate the CRDS, there by releasing DEX. The robot operator would be situated away from the rubble pile and have no line of sight to the dog. Using only the streamed video feed from DEX, the operator would attempt to explore the areas surrounding the deployment zone and indicate targets that could be seen. The video feed was recorded as well as the distance the robot was dropped from the quarry.



Figure 3.16 –Top view of OPP's Rubble Pile



Figure 3.17 – Side view of OPP's Rubble pile

### 3.4.3 DEX Experiment Results

#### 3.4.3.1 DEX Deployment Mechanism Experiment Results

In Section 3.4.2.1, two methods of attaching DEX to a dog were described. The results from the bark barrel test demonstrated a flaw in the release mechanism design of the first method.

The original concept was that the cables would retract into the robot after it was released so that the robot would not become entangled in them. In practice, this was a complete failure. When the dog ran towards the barrel at speed, the cables could not withstand the forces involved in canine transport and several of the cables failed causing DEX to be dragged and flung about. Eventually DEX was thrown from the dog, flew through the air and bounced on the ground several times. This is shown in Figure 3.17. The test did demonstrate that it is possible to construct a robot which can withstand this treatment as the robot was still working after being released and was able to move about. However, clearly the release mechanism did not work as intended.

As an alternative release mechanism, the underdog was modified so that it would wrap DEX snugly in between it and the dog, as shown in Figure 3.19. This configuration worked much better. The robot was released as planned. Figure 3.20 shows DEX dropping by a bark barrel. Several trials were performed to show that the configuration would work reliably.



Figure 3.18 – DEX being thrown from Freitag



Figure 3.19 – DEX wrapped by underdog



Figure 3.20 – DEX released by bark barrel

### 3.4.3.2 Tunnel Experiment Results

Testing under rubble proved successful in demonstrating CARD in a more realistic USAR environment. DEX was successfully dropped within a meter of the casualty. Figure 3.21 shows an image captured by DEX that includes the casualty and some of the visual acuity markers that were placed in the tunnel. The test also indicated that a CDMR's mobility is important. DEX was able

to view the casualty, but exploration around the area was hindered by several large beams blocking its path. Figure 3.22 illustrates the path the USAR canine “Freitag” took to the victim and the locations of DEX’s deployment. It also shows the location of the beam that blocked DEX from proceeding.



Figure 3.21 – Image from DEX



Unsurprisingly, the new experiment had the same results, with DEX deployed right on target or within 30cm of it.

Figure 3.23 shows how the area where the quarry was hidden. The quarry was right under the wooden lid with the handle. In Figure 3.24, on the left is a view of where DEX is deployed and on the right is a still picture taken from the robot's video feed.



Figure 3.23 – Quarry's location-hidden from view by orange cover



Figure 3.24 – DEX's deployment on rubble (left), image taken from DEX (right)

### **3.4.4 Conclusion and Future Work**

The experiments conducted at Disaster City and their results demonstrate that CARD is a viable technique for delivering a response robot across and through challenging terrain to a victim of an urban disaster. The results demonstrated that it is possible to capture video and other data that can be useful for first responders. The bark barrel experiment proved that a dog can comfortably carry a small robot under its belly and that the robot can be reliably deployed such that it will still function after the drop. The tunnel experiment showed that valuable information can be provided by the robot after a drop through a wireless network also deployed by dogs. The rubble experiment demonstrated that although DEX's mobility is limited, the CARD system as a whole allows it to reach its targeted destination. Figure 3.23 shows the challenging terrain that a dog could easily go over while any ground robot would have a lot of difficulty traversing.

## CHAPTER 4 3D Disaster Scene Reconstruction with RGB-D Cameras

This chapter is based on the contributions disclosed in the following published works [31, 32]:

- J. Tran, A. Ufkes, M. Fiala, A. Ferworn, “Low-Cost 3D Scene Reconstruction for Response Robots in Real-time,” in IEEE International Workshop on Safety, Security & Rescue Robotics (SSRR-2011), Nov 1 – 6, Kyoto, Japan, 2011
- A. Ferworn, J. Tran, A. Ufkes, A. D’Souza, “Initial Experiments on 3D Modeling of Complex Disaster Environment using Unmanned Aerial Vehicle,” in IEEE International Workshop on Safety, Security & Rescue Robotics (SSRR-2011), Nov 1 – 6, Kyoto, Japan, 2011

The main contributor of these papers is Mr. Jimmy Tran. Mr. Tran’s involvement in the development of the papers includes: primary research, concept development, algorithm development, hardware development, conducting experiments, and verifying results. Mr. Ufkes’ involvement in the manuscripts includes: co-implementation of the software used to collect and test data. Dr. Fiala’s involvement in the manuscripts includes: supervision of the research process. Dr. Ferworn’s involvement in the manuscripts include: supervision of the research process, and review of the publications.

### 4.1 Introduction

“Low-Cost 3D Scene Reconstruction for Response Robots in Real-time” discusses several methods for the creation of 3D models that can provide additional information to improve observer<sup>4</sup> situational awareness in USAR operations. 3D models can be constructed using spatial data gathered from an inexpensive, readily available video game sensor. In addition, the paper

---

<sup>4</sup> While the initial assumptions were that observers would be first responders involved in operations, it may also be true that secondary observers may find utility in more robust models presented in this work—enabling other public safety personnel to learn lessons from the more realistic representations of an incident.

introduces a new method for feature extraction as part of image registration in feature-sparse environments that operates in real-time.

“Initial Experiments on 3D Modeling of Complex Disaster Environments Using Unmanned Aerial Vehicles” studied the possible use of UAVs to capture 3D data at urban disaster sites.

## 4.2 Related Work

There are some environments that are simply too dangerous for humans to work in. Typically, such environments are formed when catastrophic events create local conditions that are intolerable for humans. Urban disasters often cause buildings to become unstable and quite dangerous. As work must still be completed within these structures, emergency first responders have increasingly turned to teleoperated response robots as effective tools. An example of such environment is provided by the Earthquake and Tsunami in Japan [53] which have created extremely dangerous environments in the areas within and surrounding the nuclear power facility at Fukushima. Here, ground robots have been used to monitor radiation levels [54] and UAVs have been used to monitor the situation from the air [55].

Response robots - controlled from a safe distance - act as surrogates for humans. The effectiveness of this control is heavily dependent on the remote operator’s perception and understanding of the situation that the robot is in when a command is issued. It has been shown that establishing this perception and understanding is typically a very challenging task [45, 46, 56-59].



Figure 4.1 – Traditional 2D Perspective from a robot’s camera

One of the limitations of working with response robots is the lack of effective mechanisms to promote spatial awareness in the human operator [60, 61]. Remote interaction usually occurs through an operator viewing a console showing various 2D camera views (see Figure 4.1) and then issuing movement commands based on the incomplete information these views provide. An improvement on this situation has been the inclusion of additional data from other sensors on the robot which may allow the human interface to depict the various joint angles of the robot arm or mechanisms for depicting the robot’s centre of gravity [62, 63].

3D information can provide the user with better understanding of the environment. Figure 4.2 depicts a 3D view of the same scene in Figure 4.1. The model can be examined and manipulated by the operator in support of gaining better situation understanding. Not to mention that important metrics such as measurements of ceiling heights, opening, etc. can provide valuable information to the rescue efforts.



Figure 4.2 – 3D viewpoint of Figure 4.1

There are many methods of performing 3D reconstructions, each have their appropriate application. One popular method is to use Light Detection and Ranging (LiDAR) sensors to obtain 3D information of the surrounding world and an algorithm to register the collected data into one

complete model. LiDAR sensors scatter its surrounding with laser pulses and measure the reflected pulses to obtain their distance. The result is a 3D point cloud, where each point is represented by X, Y, and Z coordinate with respect to the LiDAR. The LiDAR can make continuous scan producing point clouds, which then can be registered using an Iterative Closest Point (ICP) algorithm. ICP was first developed by Besl [64] and have been adapted and improve in both accuracy and computation efficiency [65-68]. The drawback with this approach is that ICP cannot handle large displacement of each scan and LiDAR sensors are very expensive.

Another way to obtain 3D information is to use stereo cameras. By tracking correspondence between the two images and computing their disparity, depth information can be gain by using the cameras' intrinsic properties. A comprehensive review was done by Scharstein and Szeliski [69]. To build larger model, an algorithm was originally introduced by Nister [70] call Visual Odometry. The trade-off with using stereo camera is that, although they are cheaper than LiDARs, they produce sparse depth information and their algorithms are computationally expensive.

The Microsoft Kinect gaming peripheral is the first commercial off the shelf sensor system that retails for less than \$200. Despite the low cost, the Kinect provides robust 3D visual data which makes it a compelling choice as a sensor. The Kinect is a depth camera, (also called an RGB-D camera). It provides both a 640x480 pixel color image from an RGB camera and a depth image provided by an infrared (IR) camera supported by an IR laser projector, all at 30 frames per second (fps) with each pixel correlated between sensors. In this paper the depth data will also be referred to as the depth image.

Depth is measured by emitting a pattern of structured infrared dots and determining the parallax shift of those dots for each pixel in the IR camera. Each pixel of the depth image has a resolution of 11 bits for up to 2048 raw disparity values. This provides a reliable range from approximately 0.5m up to 5m or more. Figure 4.3 below shows the Kinect, with cameras and IR laser emitter labeled.

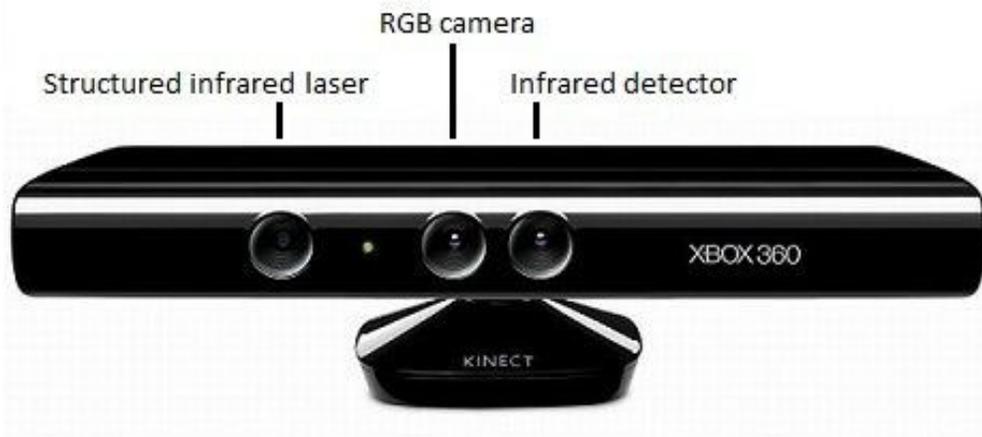


Figure 4.3 – Microsoft Kinect gaming controller

A more detailed description of the Kinect's inner workings can be found at [71]. All algorithms were implemented using the open source software package—OpenKinect OpenKinect [72] to access the Kinect's RGB and depth data streams via a standard USB hardware interface.

The propose approach is to use RGB-D camera and Visual Odometry to perform 3D reconstructions of the disaster area. The RGB-D camera is low cost and can produce dense 3D data, reducing computation cost while the RGB images can be used for visual odometry. This approach is described by Steinbrucker et. al. in [73]. However, one drawback of visual odometry is that the scene must contain a number of visually distinctive features to match between frames. This may be a problem in a collapsed structure as everything is covered in gray dust. This work explores various methods of feature extraction and matching and propose a method to overcome the distinctive feature problem.

### **4.3 Low-Cost 3D Scene Reconstruction for Response Robots in Real-time Technical Approach**

#### **4.3.1 RGB-D Visual Odometry**

The algorithm developed for the proposed system operates in a pair-wise fashion on a series of input frames, where the current input frame ( $F_{k+1}$ ) is compared to the previous frame ( $F_k$ ) to obtain the output result. In this paper, a “frame” consists of a single RGB image and its corresponding depth image.

The output data consists of two parts. First, the estimated location and orientation (pose) of the Kinect in  $F_{k+1}$  with respect to the global coordinate system established by the first frame,  $F_0$ . This information is represented by a rotation and translation transformation (RT) matrix. Second, an accumulated 3D point cloud model representing the environment that the system has moved through thus far.

The algorithm incorporates ideas similar to previous works reported in [73, 74], while introducing several new improvements and methods to increase both robustness and efficiency. There are four major steps in this algorithm:

Step 1 is to combine the depth and RGB images of the current frame in order to create a 3D point cloud relative to the Kinect camera. Although depth data is obtained directly from the IR camera, the IR and RGB images can be mapped to one another in the same manner as a traditional stereo camera setup. However, the cameras must first be calibrated. The calibration method used in the test system is the common checkerboard calibration method, adapted for the Kinect by Nicolas Burrus [75].

Additionally, the depth data is not usable in its raw state. Equation 1 introduced by Stephane Magnenat in a forum [76], is used to convert the raw 11-bit depth value to a depth value in meters. The equation utilizes mostly imperially derived numbers since Microsoft did not initially release this information.

$$depth = 0.123 \times \tan\left(\frac{rawDisparity}{2842.5} + 1.1863\right) \quad (1)$$

This formula yields the distance along the positive  $Z$  axis, which is perpendicular to the image plane. The  $X$  and  $Y$  axis values were found in a similar manner [75]. As with any stereo system, a re-projection error will be present after the calibration. For the particular Kinect used, it was measured to be approximately 1.5 pixels. Using the above methods, each depth data value is converted to an  $X, Y, Z$  triplet in a Cartesian coordinate system (where the depth camera lies at  $0,0,0$ ), transformed into the coordinate system of the RGB camera, and then projected onto the RGB image. At the end of Step 1, a list of  $u, v$  to  $X, Y, Z$  correspondences has been created for the RGB image. Any  $u, v$  in the RGB image without a 3D point correspondence is discarded.

Step 2 is performed in two phases: the extraction of interest points or features from the RGB image of the current frame (converted to gray scale), and then matching or tracking those points

back to the RGB image in the previous frame. An additional constraint must be added to the feature matching algorithm, that is, each matching pair of features must have a corresponding 3D point in their respective frames. The reason for this is made clear in Step 3.

Step 2 results in a list of 3D point correspondences, found by taking the 3D points of the matched features.

Step 3 utilizes the 3D correspondences from Step 2 to calculate an estimated pose for the Kinect. The system adopts the Random Sample and Consensus (RANSAC) algorithm [77]. by randomly selecting four correspondences at a time. Using these four correspondences, an estimate of the 3D rotation and translation from frame  $F_{k+1}$  to  $F_k$  is calculated using Singular Value Decomposition (SVD). Once the transformation is estimated, it is applied to all the features in  $F_{k+1}$ , and the result is checked against their matching points in  $F_k$ . The Euclidean distance between the point in  $F_k$  and the result of the transformation is calculated, and if the error is within a certain threshold, it is called an “inlier”. Through empirical experimentation, it was found that an error threshold of 3-5cm yielded the best results.

The RANSAC loop continues to select different sets of four correspondences at random until a transformation that yields more than 80% inliers is discovered, or the maximum number of iterations has been reached. Through experimentation, it was discovered that 200 maximum iterations were more than sufficient. Most often, the loop meets the 80% inlier condition within the first three iterations. In situations where the matching is poor, however, the RANSAC loop may perform a much higher number of iterations.

Once a suitable transformation estimate was found, the last step is to perform a least-squares SVD on all the inliers of that transformation, yielding a final refined estimate of the transformation matrix. This is then applied to all the points of  $F_{k+1}$  (not just the correspondences) to build up the 3D model of the environment. To avoid the final model from becoming overly dense, the system only adds point clouds after either a translation of 75cm or a rotation of 30 degrees. However, these parameters may be application specific.

Since the transformation is always applied backwards, that is, from  $F_{k+1}$  back to  $F_k$ , every frame is transformed into the coordinate system of the original frame,  $F_0$ .

A summary of the described algorithm is depicted in the block diagram in

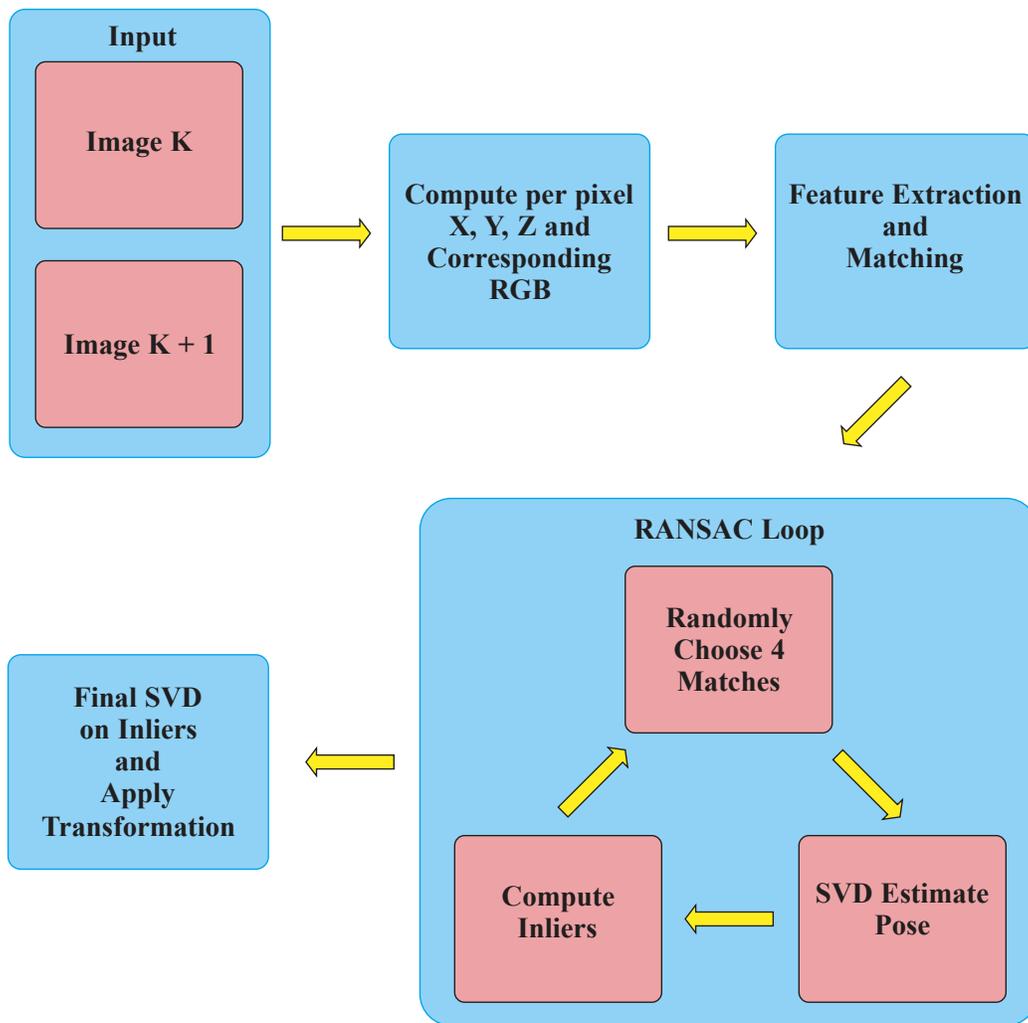


Figure 4.4 – Block diagram of visual odometry pipeline

### 4.3.2 pseudo-Random Interest Points Extractor Algorithm

The features extraction and matching is the most computationally expensive step in the algorithm. Therefore, the speed of the system, as a whole, is dependent on the methods used for feature detection and matching. Several well-known feature detection algorithms and feature tracking/matching methods were examined, including SURF [78], Shi-Tomasi corners [79], and FAST corners [80, 81]. Features were tracked using optic flow [82], and descriptor matching in the case of SURF. Although Graphics Processing Unit (GPU) accelerated implementations of feature detectors are becoming more and more popular, they were not included in the experiments. The reason for this is that smaller, single-board computer systems often found on mobile robots

cannot yet accommodate them. Instead, the methods chosen were such that they can be universally applied.

It is also important to note that in the case of optic flow, new features were found every 10 frames, or when fewer than half of the points were successfully being tracked, whichever occurred first.

The main contribution of this work is a novel feature extraction method called pseudo-Random Interest Points Extractor (pRIPE). pRIPE was used in conjunction with optic flow and designed to perform robustly in feature-sparse environments as well as areas in which limited depth data is available. These may include wide open areas, bland environments such as long hallways, or areas where ambient IR interference prevents the Kinect from obtaining reliable depth data.

In the case of feature-sparse data, the RGB image is divided evenly into sub-regions or ( $m \times n$ ) grid, the maximum number of desired interest points is equal to  $m \times n$ . A single point is then randomly selected from within each sub-region, creating a pseudo-random grid of points. These random points are then tracked by optic flow, where the size of the sub-regions can be used to roughly specify a search window for the optic flow algorithm (block diagram in Figure 4.5).

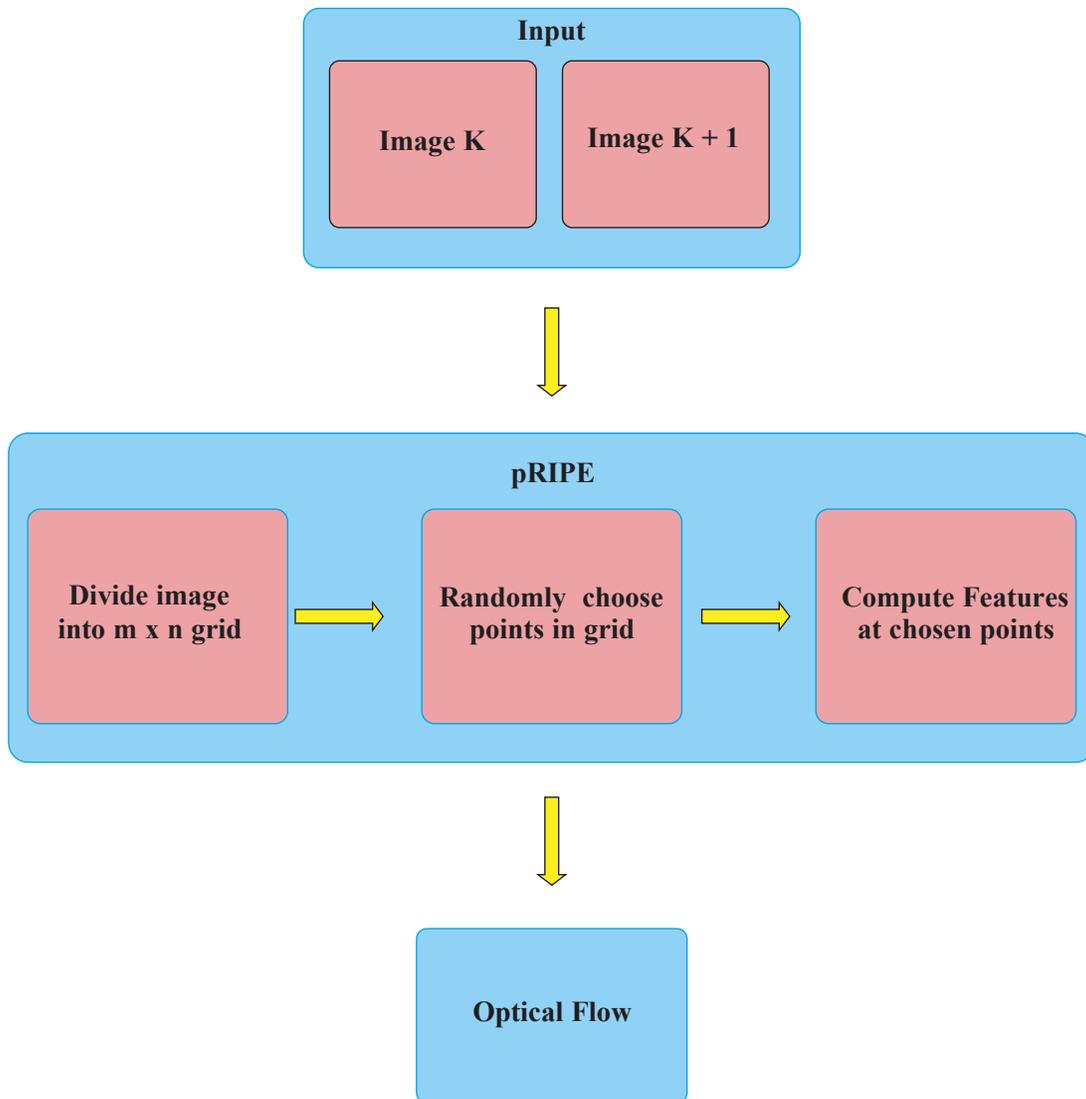


Figure 4.5 – pRIPE algorithm block diagram

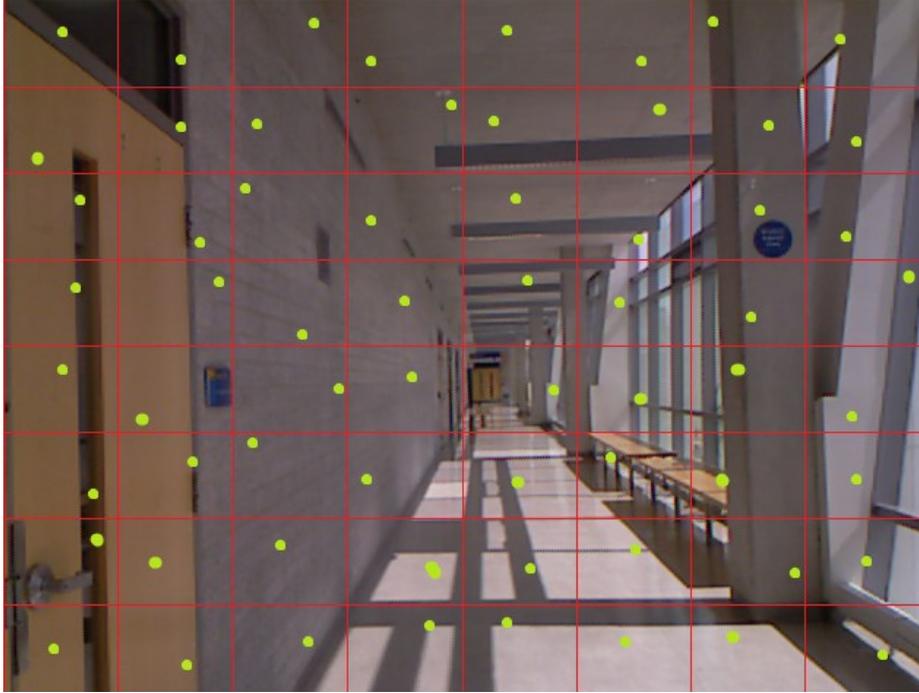


Figure 4.6 – pRIPE Feature Location Example

The rationale behind this random grid method is the assumption made that no scene is truly featureless, it simply does not favor the methods used by traditional corner and feature extraction algorithms. Unless the scene is completely plain in the most abstract sense, it will contain gradient changes that can be successfully tracked by optic flow. By dispersing the random points across the entire image (see Figure 4.6), there is an assurance that all regions of the image are sampled, and no potential area containing a traceable gradient change is missed. This method proved to be extremely effective when modeling the two feature-sparse hallway data sets.

In the case of sparse depth data, an additional constraint is placed on the selection step to only pick image points that have corresponding 3D information. This ensures that there are a sufficient number of 3D matches even if the majority of the image is lacking depth data. The resulting point cloud will still be sparse, but the transformation between the images will remain accurate.

### 4.3.3 3D Reconstruction Experiments

#### 4.3.3.1 Computing Hardware

The goal is to provide a usable 3D model of a robot and its surroundings to a human operator in order to improve control of the robot. In order for 3D data to be usable, it must be processed and displayed at an acceptable rate. Such a system would be pointless if the rate of data return was so slow as to impair the operator and decrease the effectiveness of the robot. Therefore, execution speed was of utmost importance when considering the various algorithms to be used in this system.

A system that could be adapted to different types of robots of various sizes and means of mobility has size restriction. Smaller size robots are limited by their payload capacity and will therefore have much stricter restraints on their computing power. The experiment was designed to test the speed of the system on different mobile computing platform as well as the accuracy of the output data. Table 4.1 shows the specifications of the four hardware configurations used in the experiments, which shows the contrast of computing power versus the weight of the devices. Additionally, the higher the computational power, the more demanding the electrical power, which requires more batteries (more weights) or reduced operation time.

Table 4.1 – Computing Hardware used

	Processor (cores)	Speed (Ghz)	Power Usage (W)	Weight (Kg)
Lenovo W520	Core i7-2720QM (4)	3.3	170	2.7
Panasonic CF-C1 Toughbook	Core i5-M520 (4)	2.4	80	1.8
fit-PC2	Atom Z550 (1)	2.0	8	0.37
Acer ZG5 Netbook	Atom N270 (1)	1.6	30	1.2

#### 4.3.3.2 Test Robot

The primary mobile platform used to collect data was a Pedesco RMI-9WT Remote Mobile Investigator. It provides a reasonable analogue for many of the mobile robots used in law enforcement and emergency response applications. Figure 4.7 shows the Pedesco refitted with the Kinect and the Panasonic Toughbook.



Figure 4.7 – Pedasco RMI-9WT Remote Mobile Investigator equipped with Kinect sensor

### 4.3.3.3 Test Environment

All vision-based navigation and mapping systems rely on feature detection and matching to determine camera pose. This can be problematic in environments that are not rich in natural features. A disaster scene can range from chaotic and feature-rich in the case of collapsed rubble, to plain and feature-limited in the case of interior voids, troughs, or tunnels. Hence, the algorithms were tested in three different environments that were both rich and lacking in features. Environment 1 is a lab where there are many features at depths within the Kinect's optimal range. Environment 2 is a long, narrow hallway with sparse features and good depth data. Environment 3 is a wide hallway along a windowed exterior wall. This not only challenges the feature detector, but also the Kinect's depth camera, which functions poorly in ambient sunlight (see Figure 4.8). Sample images of the different environments can be seen below in Figure 4.9 and Figure 4.10.



Figure 4.8 – The effect of ambient sunlight on Kinect depth data. black regions show where the Kinect failed to measure the depth



Figure 4.9 – Environment 1 (lab)



Figure 4.10 – Left: Environment 2 (narrow hallway), Right: Environment 3 (wide hallway)

#### 4.3.3.4 Test Method

To maintain consistency, raw data was recorded from each of the three environments using the Pedasco equipped with a Kinect. The data was then fed into the system as if it were live input. The recorded data was tested on all the combinations of feature extraction and matching algorithms and on each of the listed computing platforms. The results are presented in the following section.

#### 4.3.4 Results

Ideally, the output of the algorithms would be compared with the ground truth at every transformation. However, it is virtually impossible to find the ground truth in this manner. The most accurate points that can be measured are the starting and ending points of the recorded datasets. Although metrics are definitive, the accuracy of the algorithms can be judged by visually inspecting of the final 3D model created. This may seem backwards, but for a human operator it is enough to simply have a model that at least closely resembles the environment. Unlike robots operating on their own, human brains are capable of filling in any missing details relevant to the immediate task at hand.

Figure 4.11 shows the top and side view of a model created by using SURF descriptor and matching algorithm. It can be seen that the quality of the model is comparable to results of the model created by Shi-Tomasi and Optical flow in Figure 4.12. This provides some confidence that optical flow performance is more than adequate. This is further demonstrated when pRIPE was used in conjunction with optical flow. The model created in Figure 4.13 is comparable to the

one created using SURF in Figure 4.14. Finally, the model in Figure 4.15 was only possible with pRIPE.



Figure 4.11 – Top and side view of model created by SURF + descriptor  
matching in Environment 1 (lab)

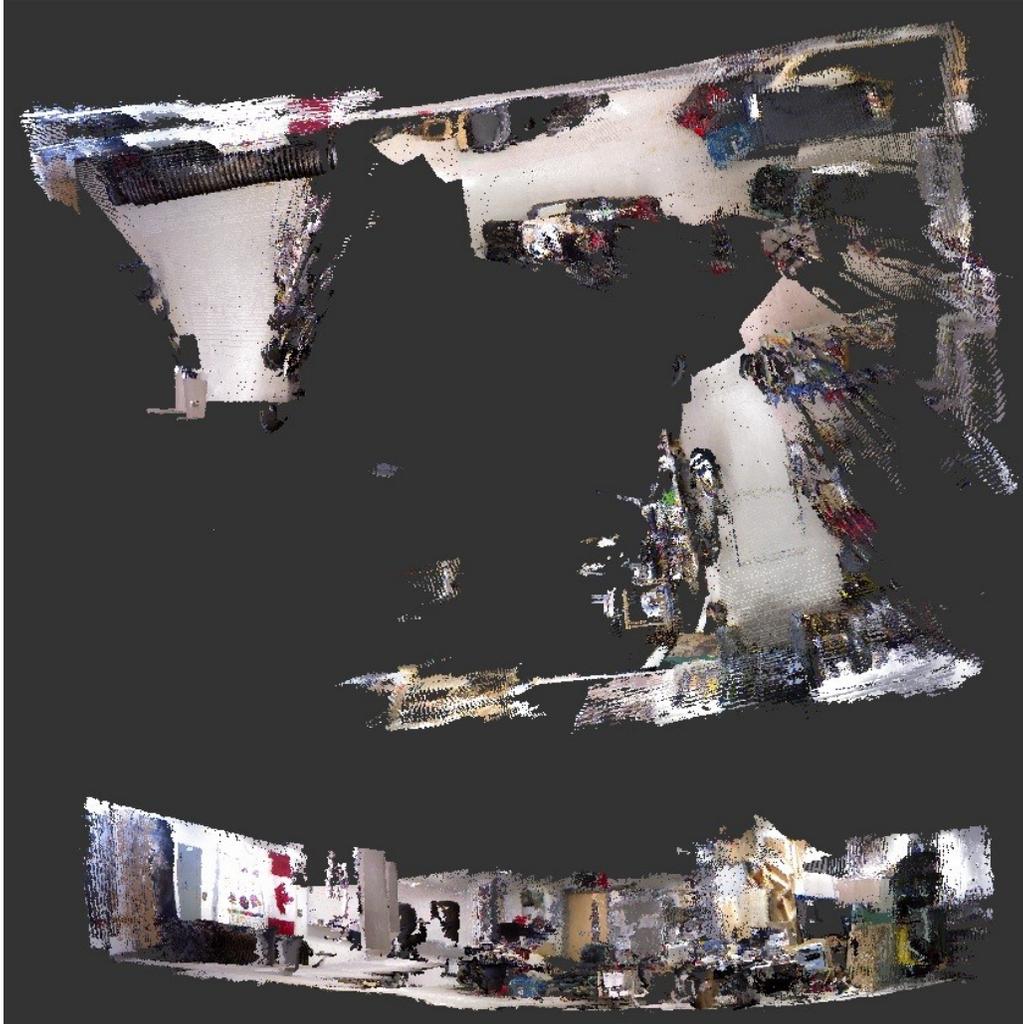


Figure 4.12 – Top and side view of model created by Shi-Tomasi corners tracked with optic flow in Environment 1  
(lab)

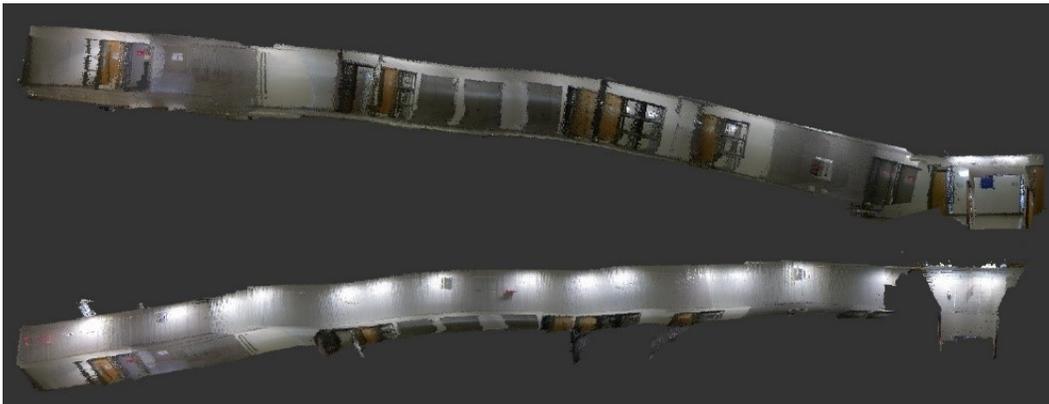


Figure 4.13 – Top and side view of model created by pRIPE in Environment 2 (narrow hallway)



Figure 4.14 – Top and side view of model created by SURF + descriptor matching in Environment 2 (narrow hallway)



Figure 4.15 – Top and side view of model created by pRIPE in Environment 3 (wide hallway)

Table 4.2 – Model Completeness

Dataset	Descriptor	Optic Flow			
	SURF	SURF	Shi-Tomasi	FAST	pRIPE
Environment 1	Complete	Complete	Complete	Marginal	Complete
Environment 2	Complete	Complete	Marginal	Fail	Complete
Environment 3	Fail	Fail	Fail	Fail	Complete

The figures above provide an idea of the type of model that the system can produce using different algorithms. Table 4.2 shows which algorithms were able to process the entire data set

without any failures. A score of marginal means the algorithm broke at one or more points, but still managed to create separate models representing different parts of the scene. A mark of fail means the algorithm failed to extract any sort of meaningful model from the data.

Contrary to the accuracy test, the speed test is easily measured. The detectors were tested on the feature-rich Lab data set. This ensures that the results are not affected by failed pose extraction and provides a good example of a situation where fast execution speed can be difficult to obtain due to the high number of features.

Additionally, the Netbook and the fit-PC were tested at a lower resolution of 320x240 and reduced the number of RANSAC iterations by half in order to try to improve speeds. With the exception of SURF descriptor matching, these changes had negligible effects on the Toughbook and the Lenovo and are not included in the results.

Table 4.3 – Feature Extraction and tracking Speed Comparison

Algorithm speed in frames per second (FPS)					
	Descriptor	Optic Flow			
	SURF	SURF	Shi-Tomasi	FAST	pRIPE
Lenovo W520	3.99	12.00	34.03	40.10 (2)	38.00
Panasonic CF-C1 Toughbook	2.75	7.33	17.67	20.00 (2)	19.49
fit-PC2	0.75	1.91	4.36	4.76 (2)	4.48
Acer ZG5 Netbook	0.47	1.26	3.10	3.50 (2)	3.13
fit-PC2 (320x240)	2.51	4.15	5.15	5.77 (3)	5.25
Acer ZG5 (320x240)	1.65	2.83	3.72	4.08 (2)	3.81

The results can be seen in Table 4.3 below and are measured in frames per second. A frame rate result followed by a number in brackets indicates the number of times the system failed to find a transformation. That is, it failed to find the required number of useable matches in order to compute the 3D transformation. In such a case, the system simply inserts a break point and begins a fresh model from that frame onward.

### **4.3.5 Discussion and Conclusion**

The experiments yielded interesting results. Optic Flow is a fast feature tracking algorithm that works well with any type of feature extractor as long as there are enough features.

SURF is a robust feature extractor that works in environments with dense or sparse features, and 3D models created using SURF features tended to be more accurate than other feature extractors. However, SURF is computationally expensive. It is significantly slower than other feature extractors, making it unsuitable for real-time mobile applications.

The Shi-Tomasi corner finder algorithm is fast but provides fewer features than SURF. When used with Optic Flow, the model accuracy is comparable to SURF with descriptor matching. It is far less effective in bland environments, however, often failing to find enough features for registration and therefore causing the modeling system to fail.

pRIPE is the only algorithm that was able to build a model in Environment 3, the wide hallway. Its strength is partially dependent on the robustness of Optic Flow, as it was able to match points even on plain surfaces. Even though the models created by pRIPE are slightly less accurate, it is also the fastest algorithm. FAST appears to be speedier only because it yields far fewer features, thus speeding up Optic Flow.

Balancing between speed and model accuracy, it appears that pRIPE is an appropriate choice to be used in modeling disaster environments. The astonishing factor is that it works in environments where all other methods failed.

## **4.4 Initial Experiments on 3D Modeling of Complex Disaster Environments Using Unmanned Aerial**

The simulated and experimental use of UAVs within USAR environments for rapid data gathering is described in [83]. UAVs are capable of carrying a diverse array of helpful sensors including cameras and other light-sensing technologies as well as various chemical and radiological sensors. This paper reports a set of experiments designed to determine the feasibility of using Commercial Off-The-Shelf (COTS) components to prototype an aerial modeling system that can be carried by UAVs at low altitudes.

The use of COTS components is well documented in diverse fields including telecommunications [84], space exploration [85] and power electronics [86] with benefits observed

including reduced cost, improved reliability and maintenance [87, 88]. The use of COTS is a well-established practice within USAR organization as well, with vast quantities of tools and other components purchased from hardware stores. The use of UAVs for USAR environments is in its infancy with many potential applications and little real experimental evidence.

This paper presents the results of several experiments conducted in the reference rubble pile (see Figure 4.16) of the Ontario Provincial Police (OPP) USAR, CBRN Response Team (UCRT) in Bolton, Ontario, Canada. The rubble pile is constructed of reinforced concrete, masonry, wood, vehicles and other debris. It provides slightly less than one acre of uneven surfaces, pits, voids and overhangs. It is a permanent, accessible representation of at least one form of collapsed structure.



Figure 4.16 – Reference Rubble Pile at UCRT HQ, Bolton, Ontario

## 4.4.1 Experiments

### 4.4.1.1 UAV

The test vehicle is the MK Hexakopter 2. It is a six-rotor UAV made by MikroKopter. The Hexakopter was especially chosen for its low cost (less than \$5000 USD) and high payload capacity of 1Kg. Most commercial, research-grade UAVs cost over \$30,000 and can only support a payload of 0.5Kg. At full payload, The Hexakopter has a flight time of over 8 minutes with the

payload. It also has very good flight stability, which makes it relatively easy for novices to fly under good weather conditions. Figure 4.17 is a photo of the Hexakopter 2.



Figure 4.17 – MikroKopter Hexakopter 2

#### 4.4.1.2 Sensor

A dismantled and repackaged Microsoft Kinect video game peripheral was used as the primary sensor. Kinect sensors are readily available for under \$200 (USD) and are capable of producing robust 3D data. The Kinect contains two cameras. The first is a standard RGB camera, and the second is an infrared camera used along with an onboard IR laser emitter to produce an accompanying depth image. The depth image is simply a matrix of distances from the Kinect along the Z axis. Both cameras return data at 30Hz at a resolution of 640x480.

The two cameras can be calibrated in the same manner as a stereo pair, allowing the depth data to be mapped onto the RGB image. This produces rich, full color 3D point clouds. The open-source libraries found at [72] were used to access the RGB and depth data streams,

and a common stereo calibration method adapted to the Kinect by Nicolas Burrus that can be found at [75]. Nister in his seminal paper [70]

The Kinect is not an ideal sensor as it has some severe limitations in outdoor environments. As part of this research work shed some insight into these limitations and discuss them along with the results.

#### **4.4.1.3 Computing Hardware**

The Kinect requires a computer to interface with. As this is a major component of the payload on the Hexakopter, the hardware was carefully chosen after several unsuccessful attempts. The computing platform finally chosen was a fit-PC2. It features a Z550 Atom processor clocked at 2.0 Ghz, 2 GB of RAM. The configuration of the fit-PC2 included an 80 GB solid state drive (SSD). Even with the additional weight of 300g, the fit-PC2 was a good choice as it was able to record the Kinect's data streams at the full 30 fps. The entire payload including the stripped-down Kinect, fit-PC2, and LiPo battery pack that provided power to both devices weighed 800g. It can be seen in Figure 4.18 below.



Figure 4.18 – UAV with hardware payload

#### 4.4.1.4 3D Point Cloud Stitching Algorithm

The software used to model the rubble pile is a 3D registration and scene reconstruction system developed in parallel to this research [32]. This is the same algorithm described in Section 4.3.1. It will be described in brief here, but for a more complete description see the cited work. Although the system was designed to run on mobile hardware, the UAV simply collected data and the modeling was done offline. This is because the single-board computer on the UAV is only capable of running the modeling algorithm at around 5fps. The speed and unpredictability of the UAV's motion necessitated a much higher frame rate in order to limit the translation between successive frames. The computer used to process the data was a Lenovo W520 laptop (3.3 GHz) and ran the algorithm at around 40fps.

The system operates in three main phases. First, temporally adjacent depth frames are mapped to their respective RGB frames. This depth data is converted from  $u, v, Z$  into  $X, Y, Z$  using the formulas found at [75]. Every  $X, Y, Z$  point in the depth image now maps to a pixel in the RGB image.

Second, features are detected and matched between adjacent RGB images after converting them to gray scale. The feature extraction method used was pRIPE, which will be described in more detail below. To match the features, optic flow was employed [82]. Optic flow lends itself very well to tracking features through minimal transformation, and since the data was so dense it was a clear choice. This results in a list of feature matches that translate into a list of 3D point correspondences.

The need to develop a new interest point detector was due to the fact that not only must points be matched between frames, but each point must also have an  $X$ ,  $Y$ ,  $Z$  coordinate from the depth image associated with it. When using the Kinect outdoors, especially during daylight hours, the depth data ranges from sparse to non-existent. This is due to the interference that bright sunlight causes to the sensor –severely limiting its ability to pick up depth data. Taking this into account, there may only be a small range of sub-regions within the image where usable feature matches may lie. It was discovered that as the completeness of the depth data decreased, the success of traditional feature detectors such as Shi-Tomasi [79] and descriptor based methods such as SURF [78] was lowered drastically. The pRIPE interest point detector identifies these regions, overlays them with a grid, and selects a point randomly from within each grid cell. The resolution of the grid is dependent on the size of the region. Through experimentations, the parameter of at least 200 points works well. The points are then tracked with optic flow, which was found to be incredibly effective even in bland regions containing little to no visible gradient changes.

By forcing the interest points to be within the regions containing valid depth data, the algorithm was able to find a sufficient number of matches with even when the depth data were only 3% of the frame.

The third and final step is to use the list of 3D correspondences to register the 3D point clouds of the two frames. This is done by using Singular Value Decomposition (SVD) within a Random Sample and Consensus (RANSAC) loop [77]. Each iteration of the RANSAC loop selects four 3D correspondences at random and performs SVD (a method that can be used to obtain a rotation matrix and a translation vector between two sets of matching 3D points) on them. This rotation and translation is then tested on all the points in the current frame, and the resulting points are compared with the matches in the previous frame. The Euclidean distance between the matching point and the result of the transformation is then calculated. If this distance is below some error threshold, it is considered an “inlier”.

Empirical experimentations showed that a 3-5cm error threshold produced the most reliable results. Once the RANSAC loop finishes, the inliers from the best rotation and translation are saved. A least-squares SVD is performed on these inliers, yielding a final, refined transformation. An additional exit condition was also implemented whereby the RANSAC loop would stop iterating if it found a transformation that yielded at least 80% inliers. Otherwise, it performed up to 200 iterations.

Once the refined transformation estimate is found, it is applied to all points in the current frame, thus placing them in the same coordinate system as the previous frame. The very first frame defines the global coordinate system since transformations are calculated and applied backwards rather than forwards. This provides a global position for the UAV rather than a relative one.

The system keeps track of relative transformations, however, using them to write points to the global cloud only when a certain minimum translation or rotation is achieved. This prevents the final model from being overly dense.

#### **4.4.1.5 Tasks**

Various waypoints were pre-marked on the rubble pile in fluorescent paint to form a flight path for the UAV. The UAV was controlled by an experienced operator standing on the second floor of a co-located repelling tower. All tests began and ended at the base of the repelling tower and consisted of short flights over the reference points collecting data. Eight flights were successfully recorded, taking place in both bright sunlight and overcast conditions. The recorded data was processed offline by the proposed system to form a model. An overhead image of the rubble pile can be seen in Figure 4.19.



Figure 4.19 – Overhead view of the rubble pile. Circled are two areas surveyed by the UAV

#### 4.4.2 Results

The flights recorded in direct sunlight were essentially useless. Though not entirely unexpected, this showed once and for all the limitations of the Kinect in direct sunlight. The suspicion was that ambient infrared light from the sun completely washes out the IR laser onboard the Kinect. The data showed that in shaded areas, depth data was successfully obtained. This effect can be seen in Figure 4.20 below, which shows a still frame over the rubble pile at a height of approximately three meters. The image on the right represents the depth data. Black areas show where no depth data was returned. This figure illustrates a very distinct performance separation between sunny and shaded areas.

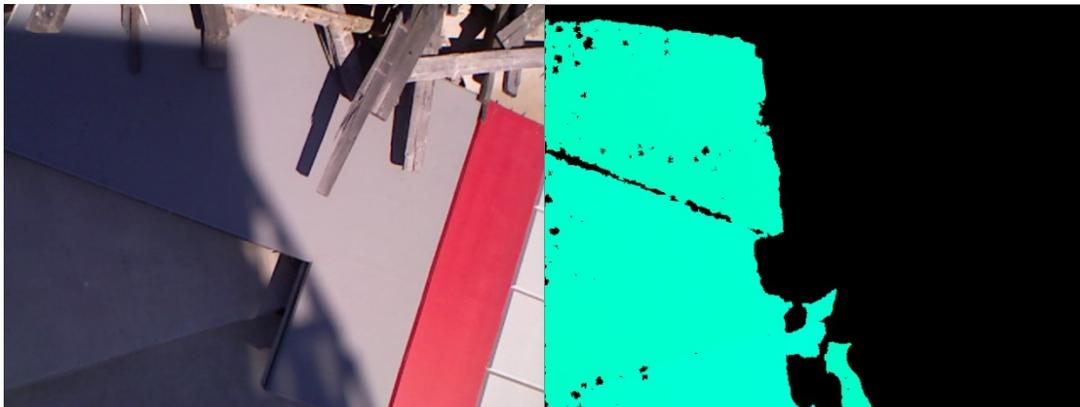


Figure 4.20 – Sunny vs. shady depth data

Flights performed during overcast periods yielded much better results. Shown below in Figure 4.21 is a swath of rubble approximately 10m long. The image on the right is a top-down view showing different types of debris and gives a sense of the structure of the pile. The left image is a side view showing the differences in terrain elevation (deepest on the left, highest on the right). A full aerial view of this area can be seen in region R1 (marked in red) in Figure 4.19.

Another section of the rubble pile is shown in Figure 4.22. This flight consisted of flying in a ‘U’ around a bowl-shaped section of the pile. As before, a top view is shown first followed by a side view below it. In the top view, the outline of a car has been highlighted. In the bottom view, a potential void large enough to contain a survivor has been found. This is something that could be difficult to recognize with top-down aerial photography alone. The aerial view of this region is marked as region R2 (blue) in Figure 4.19.

A more localized example of a void can be seen in Figure 4.23. The void was only viewed from straight on, but using the 3D model allows us to take a side view and clearly see that the void is wide and deep enough to contain a potential survivor. A close-up 2D image (see Figure 4.24) highlights the difficulty of determining the nature of the void without 3D data.

Another example of a void can be seen in Figure 4.25. This is a very clear example of the advantage of 3D data. The top down view reveals very little about the structure of the scene, even from such a close vantage point. When viewed from the side, however, a large gap opening into a potential void is seen underneath the car. A 2D aerial view (see Figure 4.26) provides no information about this void.

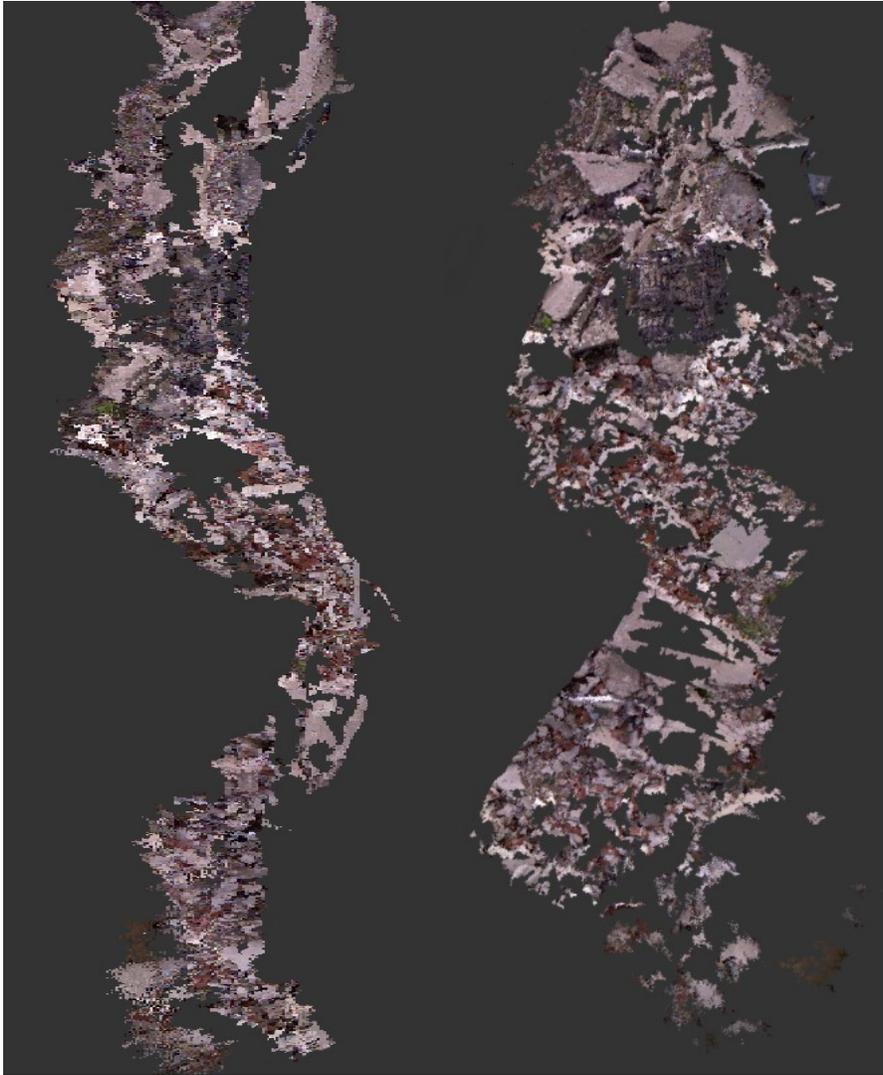


Figure 4.21 – 3D Rubble swath, side and top view

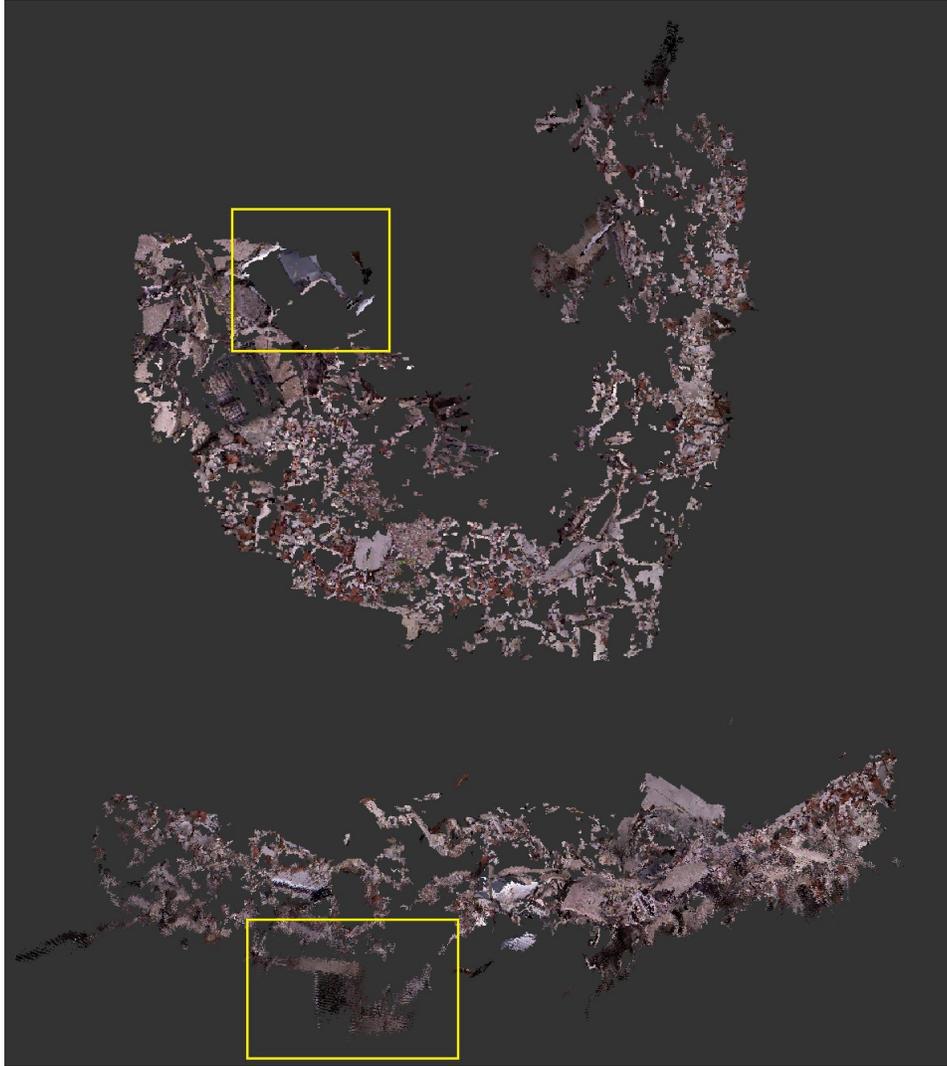


Figure 4.22 – 3D Rubble 'U', top and side views

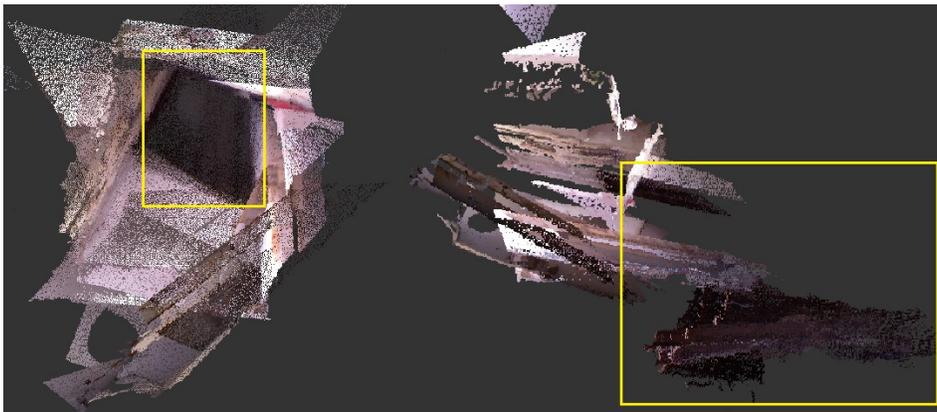


Figure 4.23 – Void, front and side 3D views



Figure 4.24 – Void entrance, 2D image

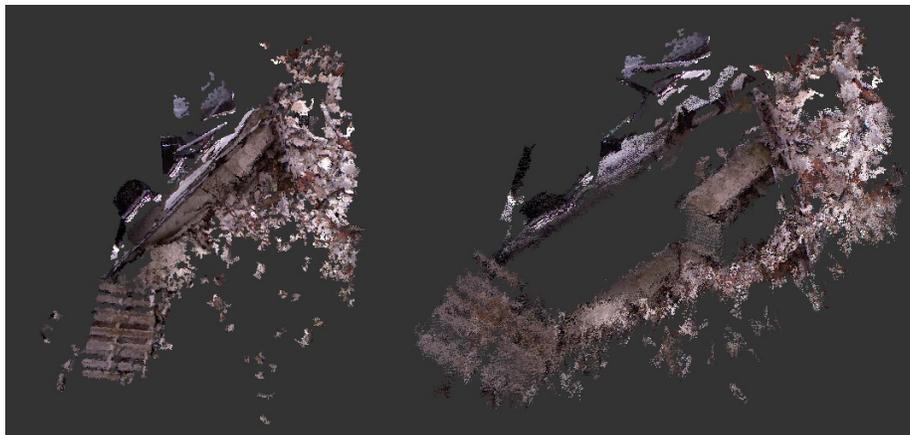


Figure 4.25 – Top and side views of hidden void



Figure 4.26 – Aerial image of hidden void

### 4.4.3 Discussion

In this paper it was shown that commercial, off the shelf hardware can be used to create a system that can be deployed in disaster response situations. This system can aerially survey a disaster site while recording data that can then be used to create 3D models of the site. These models can be used to complement existing disaster response methods and protocols.

Additionally, the experiments explored the limitations of the Kinect as an outdoor aerial sensor. Though ineffective in direct sunlight, it has otherwise proven effective for low altitude aerial modeling.

Finally, although post-processing takes only a matter of minutes, every second is critical for first responders. With the improvement of embedded systems with on-board GPUs the hope is to eventually develop this system to operate completely on the UAV with no need for post-processing. The system would create point cloud models online in real-time and transmit them wirelessly to a base station on site. This information could then be used to direct search and rescue efforts immediately, without any down time.

## CHAPTER 5 Visual Odometry from Searching Canines with RGB-D Sensors

This chapter is based on the contributions explained in the following published work [33]:

- J. Tran, A. Ufkes, A. Ferworn, M. Fiala, “3D Disaster Scene Reconstruction Using a Canine-Mounted RGB-D Sensor,” in *Computer and Robot Vision (CRV)*, 2013 International Conference on, May 28 – 31, Regina, SK, Canada, 2013

The main contributor of this paper is Mr. Jimmy Tran. Mr. Tran’s involvement in the development of the paper includes: primary research, concept development, algorithm development, hardware development, conducting experiments, and verifying results. Mr. Ufkes’ involvement in the manuscript includes: co-implementation of the software used to collect and test data. Dr. Fiala’s involvement in the manuscript includes: supervision of the research process. Dr. Ferworn’s involvement in the manuscript includes: supervision of the research process, and review of the publication.

### 5.1 Introduction

Visual Odometry is the process of estimating the position and orientation of an agent (e.g. human, robot, or vehicle) with respect to a starting position, using only input from a single or multiple camera mounted on it. This is also called “ego-motion estimation”. The word odometry is composed of the Greek words *hodos* meaning “travel” or “journey” and *metron* meaning “measure.” The term visual odometry was coined by Nister in his seminal paper [70] as it is similar to wheeled odometry—a process where the motion of an agent is incrementally estimated from integrating the number of turns of its wheels over time. In a similar fashion, visual odometry incrementally estimates both position and orientation of an agent by examining the apparent motion from image sequences captured by the onboard cameras. The nature of using visual data dictates a number of constraints: the environment is assumed to have sufficient illumination; the scene is static with enough texture to allow apparent motion to be extracted; and finally, the frame rate of the captured images must be fast enough to have sufficient scene overlap.

Visual odometry is applicable in domains such as robotics, wearable computing, augmented reality, and any application where there is a need to track the position of a mobile agent. Although

GPS provides location information with reasonable accuracy (depending on the grade of the hardware), the system can only work if there is an unobstructed line-of-sight to at least four or more GPS satellites. The importance of visual odometry is that it can provide a location estimation in GPS-denied environments such as urban areas with many tall buildings, under bridges, tunnels and inside buildings in general.

There are many examples of visual odometry systems used on ground vehicles with wheels or tracks and more recently aerial vehicles. These vehicles usually move in slow and/or smooth trajectories. The literature on adapting visual odometry to a multi-pedal platform is sparse. For USAR applications, it would be highly desirable to place a visual odometry system on a search dog. However, the work can be applied to other multi-pedal systems such as walking robots or even worn by humans.

The main challenge with using visual odometry on multi-pedal platforms is the rapid camera movement that causes motion blur—meaning that some images in the sequence are blurred. A critical part of the visual odometry pipeline is tracking key points over image sequences. If some of the images are blurred, then the tracking process can result in errors or complete failure. Even small errors can eventually be disastrous since all visual odometry systems suffer from the problem of cumulative error because these estimate of the final position/pose is dependent on the chained estimation of every other previous estimation. There will always be some small error associated with each estimation—causing the overall camera position to drift over long period of time. However, with the introduction of motion blur the system to rapidly fail.

The main objective of the proposed research is to develop a visual odometry system that accurately estimates camera motion while attached to a multi-pedal system. The focus is on developing a method of accurately tracking key points over image sequences with blurred images. The approach being developed will address motion blur and can be applied to a general visual odometry algorithm for improved accuracy. A detailed description of the proposed method is presented in Section 3.

## **5.2 Related Work**

Extensive research has been conducted in the field of visual odometry using stereo imagery, the majority of which utilizes wheeled or tracked robots as sensor platforms [70, 89, 90]. More recently visual odometry has been performed using RGB-D sensors (such as Microsoft’s Kinect)

in place of stereo cameras [73, 90, 91], and both stereo and RGB-D visual odometry has been performed on multi-rotor and fixed-wing aerial platforms [31, 92, 93].

As stated in the previous section, the literature remains essentially mute on work with multi-pedal platforms and the problems associated with using them. Howard [94] presented a visual odometry system using stereo cameras for autonomous ground robots. One of the intended applications described in the paper is terrain reconstruction for the “BigDog” robot to plan foot placement. The dataset was not collected using BigDog but by a human walking while holding a stereo camera mounted on a tripod in a sandbox. The intent was that the motion would simulate the motion of the robot. The visual odometry system adopts a standard frame-by-frame stereo approach and yielded good results with reconstruction error of around 40 cm over 200 m. The paper did not comment on motion blur but did note that a decrease in error was achieved when the frame rate was reduced. The author described this phenomenon as being counter intuitive—where better results were achieved using less information. The research work presented in this dissertation concurs and can shed some lights to the reasoning.

Aside from Howard’s work, other work that involved multi-pedal platforms are that of Takaoka et. al. [95], and Chestnutt et. al. [96]. Although their approaches differ in that Takaoka used stereo-based visual odometry and Chestnutt used a monocular camera combined with a time-of-flight ranging sensor mounted to perform reconstruction, their similarity is that they both used bipedal humanoid robots. Data collected by both papers were from slow-moving robots with stable, unnatural gaits in a flat indoor environment. These works do not address the issue of motion blur, as it would unlikely be a problem under those conditions. For visual odometry to have broader application, it should be able to work on humans, animals, and robots that can cope with unstructured terrain.

The problem of motion blur in data collected from legged robots was addressed by Pretto et. al. in [97]. They utilize the Point Spread Function (PSF) of the motion blur in local image patches to reliably extract and match features in the blurred frames. They are able to successfully match through periods of motion blur frames in artificially blurred data, as well as data gathered by a small humanoid robot. Although Pretto’s work addresses a number of problems, it is hypothesized that the evaluations can be more thorough to ensure that this can be used in a greater variety of situations. Pretto’s approach tries to resolve motion blur by improving blurred frames.

Aside from visual odometry, there has been myriad work done with respect to 3D environment modeling both in general, and using the Kinect, an RGB-D sensor developed by Microsoft. Early work with the Kinect began with visual odometry and mapping on both ground and air platforms [90, 93] and evolved into full visual Simultaneous Localization and Mapping (SLAM) systems [98]. Some of the latest, most impressive results have been achieved by the KinectFusion project [99]. This system utilizes the Iterative Closest Point (ICP) algorithm to register the 3D depth data directly, without making use of the Kinect's color camera outside of texturing the finished model. One restriction of this system is that, due to memory constraints, it is spatially limited to a region approximately  $7\text{m}^3$ . This problem was solved by Kintinuous [100], which removes the spatial limitation by modifying the KinectFusion algorithm to allow environment sizes that can vary dynamically. However, preliminary tests with the KinectFusion algorithm implemented in the PCL library [101] showed that it was unable to model the data gathered by the canine mounted sensor. The reason is that for ICP to work, there can only be a small transformation between frames, but dogs move too fast and erratic causing the transformation to be relatively large.

### **5.3 Contributions**

There are three contributions explained in this paper. First, a method of mounting an RGB-D sensor on a moving dog for collecting visual data. To the best of the author's knowledge, this is the first set of RGB-D data collected using a dog as a platform. Second, experiments were conducted, and the observations made provided an analysis correlating the dog's gait with blurred images. Third, an algorithm was developed that can look ahead and skip over noisy frames to select appropriate frames for registration. The results presented in this paper show that the algorithm is useful in creating models that more accurately depicting the environment captured by the dog.

### **5.4 Technical Approach**

One of the goals of canine search in USAR applications is to determine how to track and recreate the path of a searching dog. This is because the searching dog will provide a path that is guaranteed to be passable—at least to dogs—allowing access to any victims the dog finds and clues to where they might be found hidden under tons of rubble.

The proposed solution is to use a canine-mounted RGB-D sensor on a searching dog. The proposed methodology is to mount an RGB-D sensor on the dog's back, pointing up and the dog (see Figure 5.1). The sensor will record the journey of the dog. If the recorded data can be registered, then the path can be recreated. The methods employed to test this hypothesis can be described in three phases:

- data collection – recording of on-dog data
- data analysis – the examination of how motion blur affects the registration process
- data processing – experiment with different strategies to perform registration on noisy data.



Figure 5.1 – RGB-D Sensor mounted on search dog Dare

#### 5.4.1 Data Collection

The sensor chosen is the Microsoft Kinect. Since its release, the Kinect has become ubiquitous among roboticists and researchers due to its low cost and ability to produce dense 3D point cloud data. For a complete description of the Kinect's capabilities, see [71]. The system records RGB-D data from the Kinect at a rate of 30Hz and at a resolution of 640x480.

The Kinect sensor is mounted on a rigid platform attached to a custom designed harness. The Kinect is angled upwards just below the dog's shoulders, ensuring that the dog obscures as little of the Kinect's view as possible. The recording unit consists of a battery pack and single board computer that interfaces with the Kinect sensor and stores the recorded data. It is attached

to a bag that is strapped to the dog's chest. Figure 5.1 shows a picture of the system worn by a USAR dog.

### 5.4.2 Data Analysis

Even though efforts were made to ensure that the Kinect sensor stays as rigid on the dog's back as possible, swaying of the sensor from side to side still occurs. Through observations, Figure 5.2 was produced to demonstrate the motion of the sensor on the dog's back. As the dog's speed increases, the speed and severity of the sensor's swaying motion increase as well.

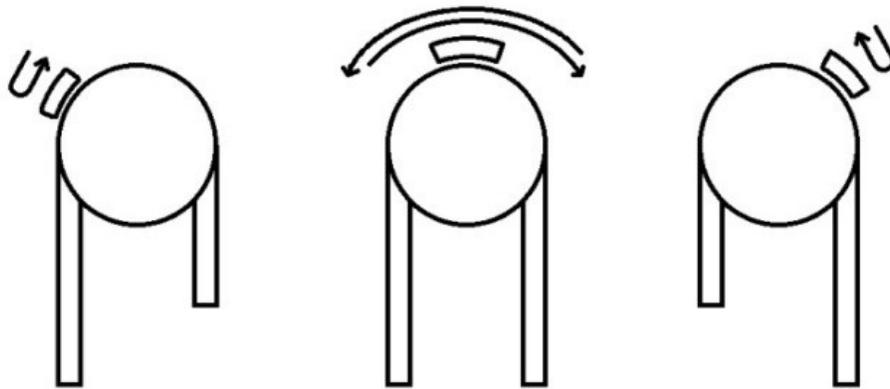


Figure 5.2 – Depiction of sensor motion on a moving dog

To identify usable frames within the motion blurred data, analysis of the motion of the sensor with respect to the gait of the dog. It was observed that the sensor tends to oscillate very consistently with the dog's stride. When one of the dog's front legs reaches its highest point in the stride, the sensor is rolled the farthest away from that leg. This creates an inflection point where the sensor has rolled farthest to one side and begins rolling back in the other direction as the dog lowers that leg.

At these inflection points in the data, there are typically a handful of relatively useful frames with minimal motion blur. The downside is that these clear frames are also the most severely rotated and spatially separated. It also followed from this that the frames midway through the dog's gait were the most blurred and therefore unusable despite the fact that at these points the sensor is mostly level without any rotation. Figure 5.3 shows what happens when a clear frame is matched to a blurry frame. The matching is poor, and there is no way to obtain a robust registration.

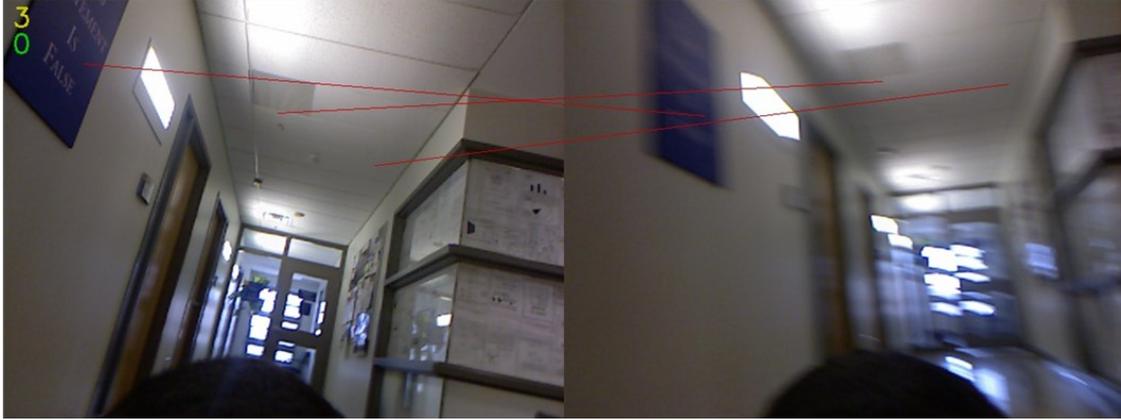


Figure 5.3 – Example of feature matching between a clear and blurry image

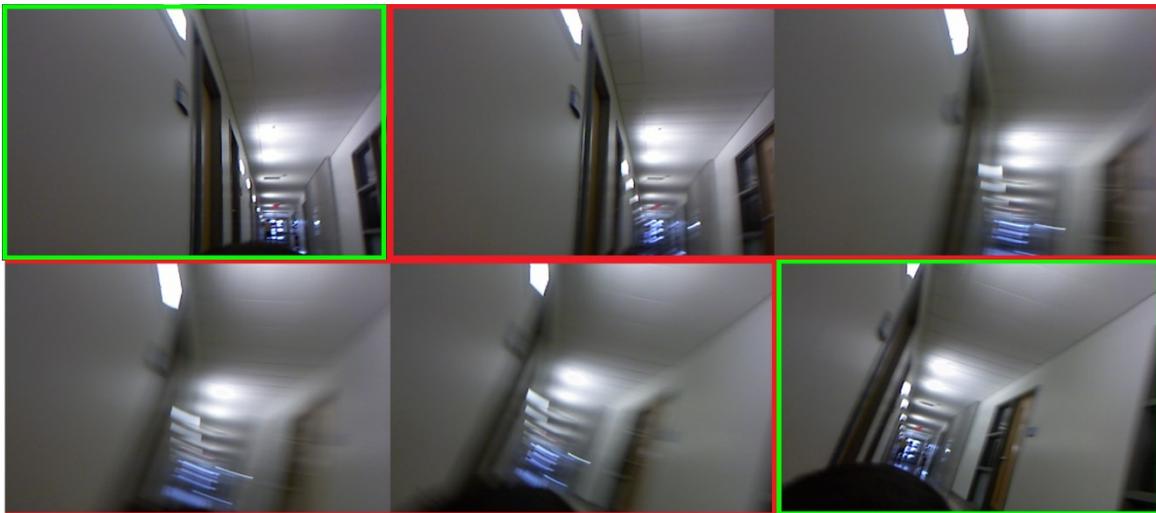


Figure 5.4 – Consecutive series of blurry frames (red) between clear frames (green) on the inflection points.

### 5.4.3 Data Processing

The method used to register frames is similar to the method presented in [102], which highlights the efficacy of the optic flow tracking algorithm [103] even under extreme conditions. Here, however, the pairs of usable image frames contain too much spatial separation for optic flow to be used reliably. Hence, descriptor matching is used to obtain image point correspondences. The proposed system utilizes a GPU implementation of the popular Speeded-Up Robust Features (SURF) [78] algorithm in the Open Computer Vision Library. From each pair of input images, SURF features are extracted and matched to produce a list of 2D image correspondences. These matching pairs of image points are projected into 3D using the depth data for each pixel provided by the Kinect, resulting in a new set of 3D correspondences.

The transformation between these sets of corresponding 3D points can be determined using SVD. To remove outliers such as poor or false feature matches, a RANSAC loop [77] was used. This loop randomly selects a set of four matching pairs of points and computes a hypothesis transformation. This hypothesis is then applied to all the matched points from one frame. If the result of applying the transformation to a given point is within a certain threshold of the matched point, the matching pair is considered to be an inlier. After a predetermined number of iterations, the transformation that yields the highest number of inliers is selected. A least-squares SVD is then performed on the set of inliers matches which yields the final, accepted 3D transformation between the two frames.

Initially the collected data was used into the described visual odometry system without any selective rejection of blurry frames. As expected this produced poor, inaccurate models.

The system was then modified to use a human operator to manually choose the set of frames to match with each other. This experiment was a proof of concept to see if human intuition and recognition of clear frames would help the system produce better models.

Finally, an algorithm was developed to automate the manual process of using the aid of a human operator. The algorithm was called Intelligent Frame Selector (IFS). The first step in the algorithm is to produce a large matrix of matches between every single frame and the next  $N$  number of frames where  $N$  is the window size. This creates a histogram of the number of match inliers found in each frame. The second step is to generate a list of local maxima. Figure 5.5 shows a sample histogram of matches and the red bars highlights the local maxima. The local maxima is defined as the frame with a higher inlier counts than its neighbours (frame before and frame after).

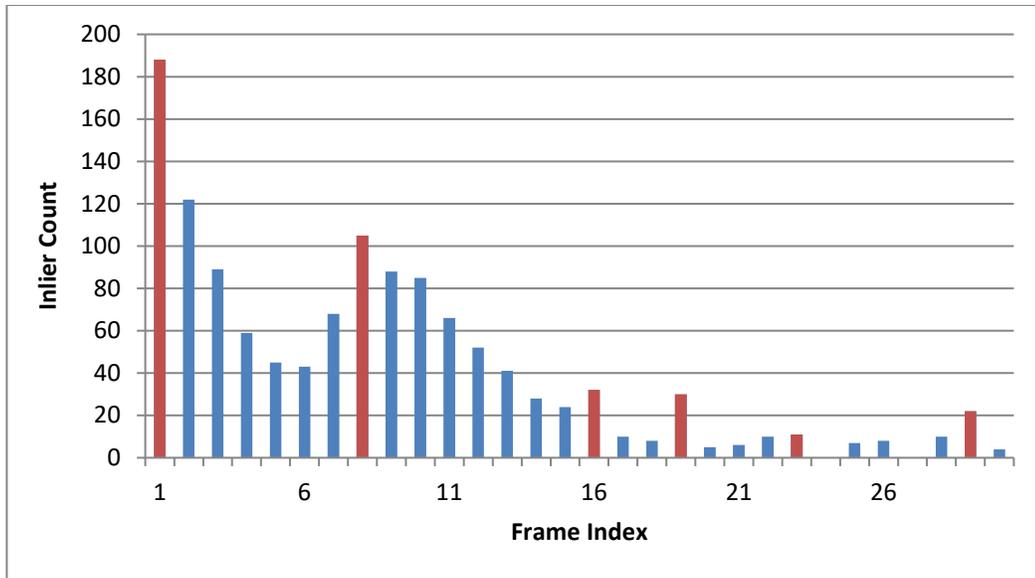


Figure 5.5 – Sample histogram of a window of matches

It was observed that these local maxima each lie across a series of non-blurred frames from the inflection point of the sensor motion described in Section 5.4.2. This observation forms the basis of the IFS algorithm. The algorithm selects the frame lying on the furthest local maximum that has a greater number of inliers than an acceptable threshold. The acceptable threshold is dynamically calculated. It is the average inlier count of every frame in the window with an inlier count higher than four. The IFS algorithm is depicted in a block diagram shown in Figure 5.6 below.

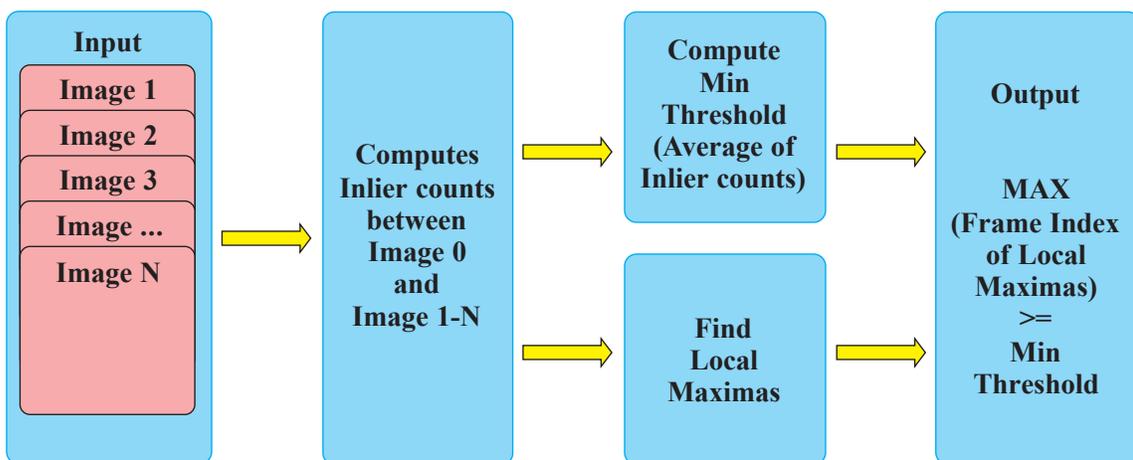


Figure 5.6 – IFS Algorithm Block Diagram

Registration is performed between the chosen frame and the origin frame. The process is then repeated from this new frame. Essentially the algorithm steps from local maximum to another local maximum, avoiding blurry frames as much as possible.

One of the parameters that can be adjusted is the size of the search window. This would depend on the data and the environment. From experimentation, the window size should be adjusted based on the speed of the dog. For all but the slowest tests, within a span of 30 frames the dog had moved a distance greater than the ideal range of the Kinect's depth camera, making it impossible to match beyond 30 frames.

Searching for matches across an entire window for each new frame may seem expensive but in reality, there is only a small amount of redundant computation. Features are extracted once per frame and then stored until the window passes over that frame. Matching is the only process performed multiple times per frame, but it typically takes under 2ms with GPU acceleration.

## 5.5 Experiments

The test subject for this paper was Dare, a FEMA-certified OPP USAR dog. Under the guidance of Dare's handler, Constable Dan Baily, several trials were conducted.

At each trial, Dare walked/ran down a long, straight, interior hallway (see Figure 5.7) at three different speeds, slow, medium and fast. In the slow and medium speed trials, Dare was led by his handler to control the pace. In the fast speed trial, Dare started at one end of the hallway and was called by his handler at the other end. Dare's handler advised us that the fast speed trials show the typical pace for Dare when performing a search in a real scenario.

There are several reasons for using the hallway. The straight path makes it easy to calculate the speed of the dog and to visually judge the quality of a 3D model. Additionally, the hallway is visually plain and the little textures and features that it has are repeated consistently. The scarcity of visual features and geometric 3D features makes it one of the most challenging environments for visual odometry systems. The motivation is that the interior of a disaster site is unpredictable. The presence of large amounts of dust and debris tend to obscure defining visual characteristics, which puts severe strain on feature detection and matching methods.

It is difficult to obtain ground truth data with the experimental setup without an expensive motion capture system. Only visual inspection can be used to evaluate the model produced by the proposed system. To create a baseline comparison, the recorded dataset was created from the

hallway with an RGB-D sensor mounted on a wheeled-cart. The model created from this dataset represents what the data should look like if it was on an ideal platform (smooth and slow moving). Table 5.1 shows the details of the six dog trials and one cart trial.



Figure 5.7 – Plain interior hallway at George Vari Engineering Building at Ryerson University

Table 5.1 – Hallway trials

	Distance (m)	Speed (m/s)	Accompanied
Slow1	44	1.00	Yes
Slow2	44	1.08	Yes
Medium1	44	1.60	Yes
Medium2	44	1.65	Yes
Fast1	44	2.35	No
Fast2	44	2.37	No
Cart	44	0.65	N/A

## 5.6 Results

In the slow and medium speed trials, the proposed system was able to generate four complete models from one end of the hallway to the other. In the high-speed trials, the system produced a model that represented approximately two thirds of the hallway. There was a section of the hallway where there was a long stretch of plain wall and at the fastest speed of the dog, the system failed to produce a suitable registration.

From top to bottom, Figure 5.8 shows a comparison between the model created from Cart trial data with models created from the Slow1, Medium1 and Fast1 trials. These models were created using the IFS algorithm.

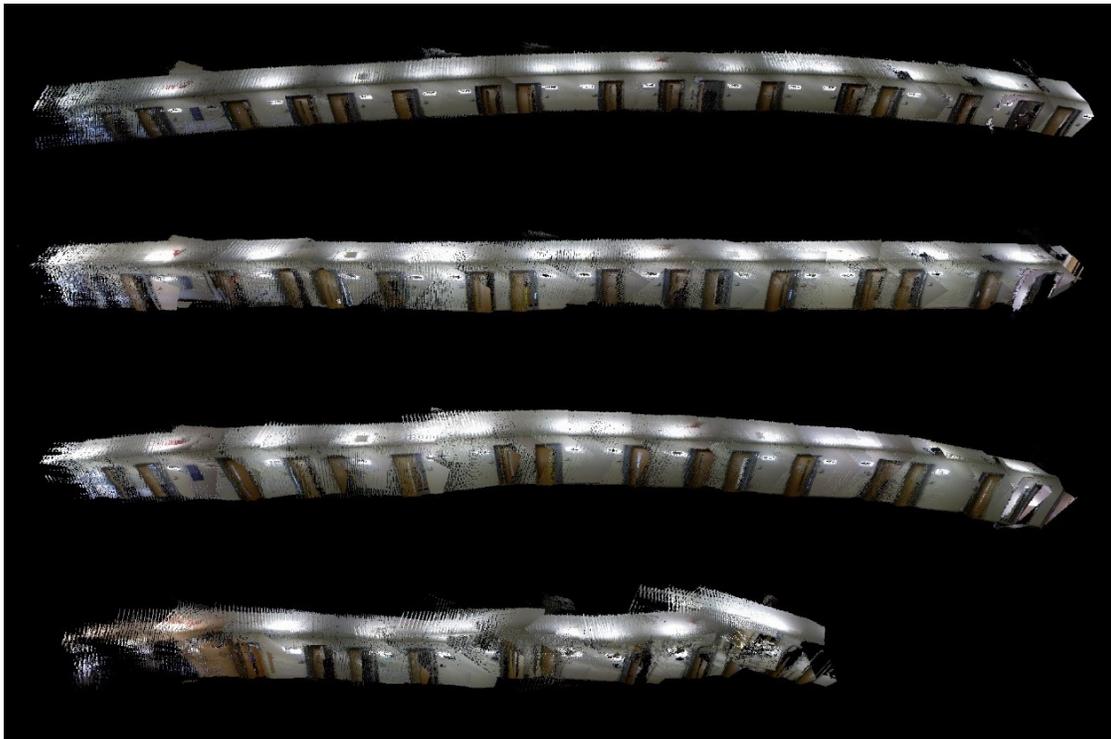


Figure 5.8 – From to top bottom, model generated from Cart, Slow1, Medium1 and Fast1 datasets

As a comparison, the proposed system was tested without the IFS algorithm. The models produced from the slow trials have several badly broken parts and the models from the fast trials are not even recognizable. This is shown in Figure 5.9.

A comparison between IFS and the human assisted process was made. As shown in Figure 5.10 the models created using human assistant in the Slow1 actually suffers from more drift than the IFS generated model. In the Fast1 dataset, the IFS and human assisted models look similar.



Figure 5.9 – Models generated without IFS (top-Slow1, bottom-Fast1 dataset)



Figure 5.10 – Models generated by human assisted process (top-Slow1, bottom-Fast1)

Table 5.2 shows the process time of each dataset using the RGB-D system without IFS, with IFS and with human assistance.

Table 5.2 – Process time comparison

	Video Length (s)	Frame Count	Process Time (mm:ss)		
			Without IFS	IFS	Human Assisted
Slow1	44	1320	0:56	4:25	~30:00
Slow2	44	1320	0:54	4:32	~30:00
Medium1	27.5	825	0:37	2:45	~25:00
Medium2	26.7	801	0:30	2:46	~25:00
Fast1	18.7	561	0:22	1:30	~20:00
Fast2	18.6	558	0:22	2:28	~20:00

To further validate IFS, it was compared against another system that use ICP, which does not rely on a crisp RGB images. The datasets were ran through on a PCL implementation of KinectFusion but it failed to produce a model.

## 5.7 Discussion and Conclusion

This paper presents work on recreating the path of a search dog in a GPS denied environment while building a 3D map. The challenges of using visual odometry on a canine mounted platform were presented and the results demonstrated that the IFS algorithm can help a visual odometry system produce better models in these extreme conditions. While IFS does increase process time, this is acceptable in a USAR scenario. Due to the logistics of using a canine-mounted system, all processing must be done offline. Thus, the increase in processing time is negligible when compared to the timeline of the rescue operation. The IFS algorithm is also significantly faster than the tedious human assisted process.

The positive results from the experiments warrant further study in the canine-mounted visual odometry problem. The next challenge would be to determine limitations of this system in confined space. The obvious limitation is the minimum workable distance of the Kinect. The system should also be tested on winding paths and under various lighting conditions.

Finally, there need to be further study of other methods of mounting the sensor on the dog. It is hypothesized that mounting the Kinect further down the dog's back may significantly reduce the rolling back and forth of the sensor. An additional mounting technique being considered is to use two sensors, one on each shoulder looking sideways rather than backwards. It was observed that during the fast trials, the dog would often move very close to the wall causing large portions of the image to be inside the minimum range of the depth camera. Mounting cameras on each side ensures that at least one of them has a clear view.

Currently there is an emphasis being placed on multi-pedal mobile platform research such as the quadropedal robot, BigDog [104] (see Figure 5.11) and bipedal robot, PETMAN [105] (see Figure 5.12) for military and public safety applications. These robots have gaits and motion similar to that of humans and animals and are able to traverse a variety of surfaces. Future visual odometry system must be able to cope with this type of platform operating in the real world.

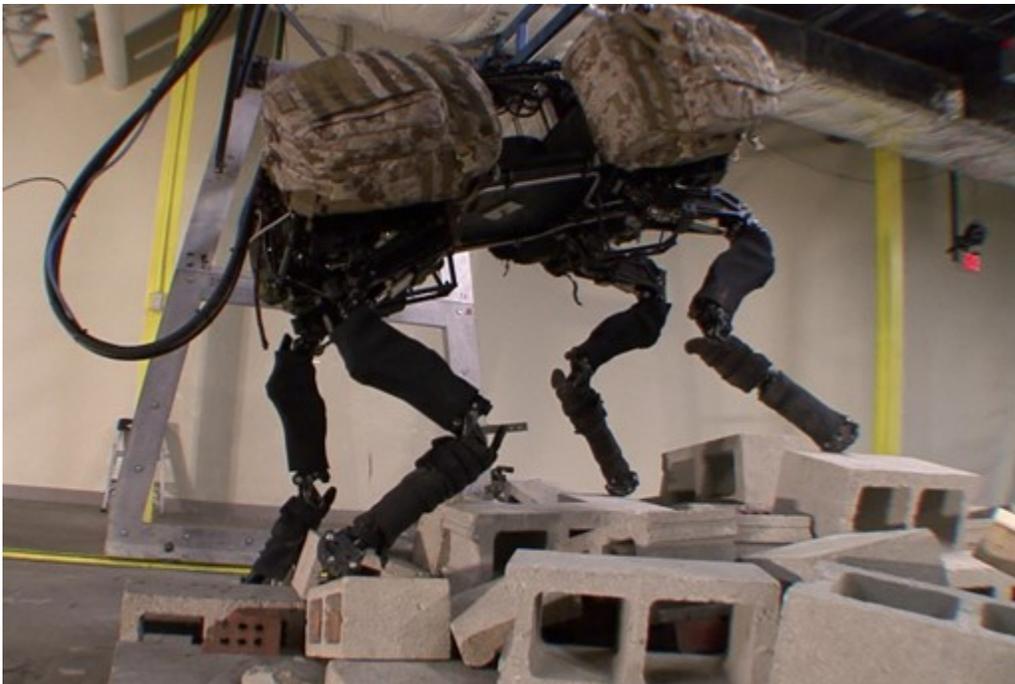


Figure 5.11 – Boston Dynamics' BigDog robot walks on cinder blocks as reported in [104]



Figure 5.12 – Boston Dynamics' Petman robot with and without suit as reported in [105]

Aside from being attached to robots, wearable computing is a growing field. With the launch of devices like Google Glass [106] (see Figure 5.13), ego-motion estimation can play an important role in our daily lives as these devices become ubiquitous. It is not hard to imagine some form of Head-Up Display (HUD) coupled with visual odometry complementing GPS-based location-based services in GPS-denied environments. Search dog may not be the only biological agent that wears computing, but humans may do so in the future.



Figure 5.13 – Picture of Google Glass made by Google [106]

The impact of the proposed research would expand the application of visual odometry to the realm of multi-pedal agents that inherently inject motion blur through natural motion leading to spasmodic and rapid camera movement. Furthermore, this work provides a way to improve the robustness of many state-of-the-art approaches to visual odometry.

## CHAPTER 6 Summary and Conclusions

### 6.1 General Conclusion

The applicability of robotic technology to USAR operations has always been evident. The working environment is inevitably dangerous, unpredictable and rife with challenges that would seem to be addressable by technology—if only that technology existed. In practice, hundreds of robots have been conceptualized, tested and deployed with limited to no success. By and large, USAR operations throughout the world are carried out by human first responders and their dogs.

This is not to underplay the role of technology in this work environment. Optical, heat, GPR and other sensors have played an important role in finding people, but their application has always been hindered by their inability to deal with the inherent difficulties of traversing rubble and getting the sensor to a point where it might do some good.

In the work presented in this thesis sought to use the tools already available to first responders working in USAR environments and augment them with elegant applications of technology guided by the computational thinking in the field of computational public safety.

With few exceptions, there are no purely robotic solutions that have been successfully applied to the basic problem of rubble traversal. One of the exceptions is the case of the Fukushima reactor melt down, robots were used to gain access and activate controls. This would be impossible for dogs. However, this is a very limited application and although there was a disaster, the structure was generally intact. The bar is high for finding people quickly. Most structural collapses happen on a massive scale where radiation is not an issue, but time is of the essence. Dogs meet this bar. The work demonstrated complete systems based on augmenting search dogs with technology, founded in theory and demonstrated in practice.

Given the reality of the problems associated with USAR operations including structural stabilization, gathering situational awareness, searching for, finding and rescuing survivors, the argument is that the ideal solution for a telepresence system used in USAR operations can be described as a mobile robotic platform—about dog-sized—with the agility of a dog, able to traverse the rough terrain of rubble. Equally important, the operator of such a system must be able to control the robot to go to designated positions. This ideal robot should also have a set of sensors that allows the operator to seek out a human victim with great accuracy. Additionally, the robot

must be able to report its position within an unmapped and GPS-denied environment. Finally, the operator needs to be able to manipulate components of the disaster environment through the robot.

Currently, this ideal solution does not exist as every feature of this type of robot are open research topics in robotics. While the presented research work is not the optimal solution, it can be argued that the contributions made provide first responder in USAR operations with better solutions than current robotics technology allows. These contributions are detailed in the next section.

## **6.2 Contributions**

A major problem in Human-Robot-Interaction is how to gain enough perception to allow a robot to be controlled remotely in such a way that a useful task is accomplished. Biologically inspired snake robots [107-109] and variable geometry track robots [40, 110] have high degrees of dexterity offering potential solutions to being able to traverse rubble, but they are difficult to operate remotely when it is difficult or impossible to determine their configuration relative to their environment. Usually, the more configurations that a robot can exhibit, the more difficult it is to control [45, 47, 59]. The work done by Ferworn et. al. [22, 23, 29, 51, 111, 112] exploits the Biological Intelligence of search dogs as a kind of metaphor replacement for autonomous robots is central to this work. Dogs are trained or “programmed” to search for live victims and use their own perception to perform the search and automatically traverse rubble. Ferworn’s work treats dogs like “black boxes” exhibiting exactly the performance required for a task but without explanation. The work reported in this thesis builds on Ferworn’s work.

One contribution is the extension of the Canine Remote Deployment System with an additional bark detection and activation system BARS. With BARS, the CRDS is guaranteed to release a package from a dog regardless of whether wireless signals can reach the release mechanism. The importance of this is that significant buildings tend to be constructed from concrete with steel reinforcing rods (rebar) which block wireless signal. The CRDS forms the basis for other contributions.

The next contribution of this work is the Canine Assisted Remote Deployment framework with the design guide for DEX. The ability to deploy a robot from a dog mitigates some challenges of using search dogs. While a search dog is fully capable of doing its programmed task, the handler lacks the ability to finely control the dog. This includes the lack of a way to stop and

examine scenes of interest in the disaster environment. The dog cannot linger with the victim indefinitely and is usually in constant motion—making stable collection of visual data challenging or impossible. CARD and DEX extend the operator’s/canine handler’s telepresence in USAR operations by providing a systematic way to bring robots much deeper into disaster rubble where victims may be found. Without the protocols associated with CARD, no robot would be able to reach a victim in most scenarios

Another contribution is the pioneering work in using RGB-D cameras in USAR operational sensing. It has been demonstrated that the limitation of RGB-D cameras is their ineffectiveness in direct sunlight. However, they can be used in shaded areas or during twilight or dusk. pRIPE was shown to be a robust algorithm where visual features are sparse and depth data is spotty. Disaster environment are unpredictable as are the visual features in them. While the pRIPE algorithm accuracy is not state-of-the-art, its robustness can be crucial in providing much needed information.

The last contribution is the concept of using visual odometry on search dogs and the IFS algorithm. Due to the nature of canine movement, any visual data collected from a moving dog contains many unstable and blurry images. However, the ability to successfully perform visual odometry from a search dog allows an observer to track the dog in a GPS-denied environment. In addition, the system allows the 3D reconstruction of the interior of disaster rubble that leads directly to live victims found by the dog. The impact of the IFS algorithm is not only to improve visual odometry on search canines, but it reduces the cumulative error effects of general visual odometry.

### **6.3 Limitations of this work**

Although the contributions of this work have made significant improvements to the tools available to USAR first responders, there are limitations in each system presented. This section discusses these limitations.

The Bark Release System’s limitation is the limitation of the search dogs themselves. The dogs are incredibly apt at searching for live victims and are successful the majority of the time. However, there are situations that are similar to the results of the experiments discussed in Section 2 and [28]—when one of the search dogs barked and indicated that it found a victim. In reality, due to the way scent travels in rubble, the actual location of the victim was about 3.6M away. Dogs

perform their search using scent and thus are prone to this problem. The robot carried by a dog using CARD and released through bark detection would fail to find such a victim.

A major problem with deploying ground robots in USAR operations is communications. USAR robots currently in operation are teleoperated through tethered control lines or radio frequency (RF) links [15, 113]. Tethered robots typically require additional operators to physically manage the tether. Tether design eliminates signal loss, but they are difficult to manage since they can be caught, snagged, become tangled or break. Wireless communications are susceptible to interference from a wide variety of causes and signals have difficulty in penetrating concrete and other debris. Wireless robots may be susceptible to inadvertent jamming by other operators controlling other robots in the vicinity. Although CARD and DEX solves certain mobility problems, they would still have the same communication limitations.

The pRIPE algorithm trades accuracy for the ability to complete a model by pushing through difficult parts of a data set. Since pRIPE utilizes Optic Flow, the transformation between frames must be small for it to work. The cumulative error is also a problem—especially when working with a data set collected from a camera mounted on a search dog. Using a SURF-based visual odometry along with IFS improves performance greatly. The limitations of all visual odometry is that the images used from the video stream must have sufficient illumination, have enough visual features and the captured images must have sufficient scene overlap to match with each other. In real life applications this cannot be guaranteed. Additionally, to obtain depth information, objects are required to be a minimum distance away. Some parts of a disaster environment can be very confined and thus there will always be areas where visual odometry cannot work.

#### **6.4 Future Work**

As mentioned in the previous sections, remotely controlling robots in disaster sites often encounter communication problems. One solution to expand wireless range is to use WMN technology. WMNs can be applied to different wireless technologies, but since robots need to stream rich data, WiFi protocol is appropriate. WMNs are self-forming, self-configuring, and self-healing. Self-forming means as each mesh router node goes on-line, it automatically connects to other mesh routers to form a network. Self-configuring is the ability of each mesh router node to configure the shortest path to route the data from one client to another. Self-healing is the ability of the WMN to be aware when a mesh router node is down and to re-route data through a different

path [114]. WMNs are flexible network, with fault tolerance can be deployed rapidly with relative ease.

In most applications mesh nodes are deployed with much deliberation in order to maximize their performance. In USAR operations, this is not a luxury that can be afforded. The difficulty in deploying WMNs is that usually one cannot simply walk onto disaster rubble. It is hypothesized that as a CRDS-equipped search canine is conducting its search a trail of “breadcrumbs” of mesh nodes could be deployed to form a WMN allowing the dog to create a form of “wireless tether” supporting communications. Much work must be done to determine the appropriate hardware would be applicable and how automatically deployment might be achieved through some algorithm yet to be discovered.

Aside from using a pure wireless solution, some preliminary work has been done on a hybrid approach employing wired and wireless networks as suggested in [115]. The propose of such a solution is a system developed to establish communications through the delivery of appropriate equipment by dogs. Search dogs with a bark-release-capable CRDS can deploy a communication node upon locating a patient. The node is a WiFi access point tethered to an exterior router that can be used as a command centre.

The work in [115] describes a proof of concept system, where a modified Category-5 (CAT-5) Ethernet cable was used as a tether. The cable was modified by stripping the shielding, reducing four twisted pair wires to two, which reduces the size and weight of the cable greatly. A custom container was designed to act as a magazine for the deploying tether. Since the tether is deployed from the dog, it avoids any snagging or entanglement issues. The dog carries both the tether and the access point with it. The results from this work demonstrated that a tether can be deployed successfully from a dog like it might be deployed from a robot. However, in the preliminary results, the tether must be wound in a specific way, which was both time consuming and included limitations on how long a tether could be. Future work would include revisiting this system in order to design a miniature spool system fed by a cable winding algorithm. Such a system could be carried by a searching dog and, on deployment, might serve to provide better wireless communication from a located victim to searching first responders.

The future work of a canine-mounted visual odometry system involves exploring different ways of mounting the sensor on the dog. It may be possible to reduce the rolling back and forth of the sensor by mounting the sensor further down the dog’s back. Another mounting technique

maybe to use two sensors, one on each shoulder of the dog looking sideways rather than backwards. When a search dog moves close to a surface, causing large portions of the image to be inside the minimum range of the depth camera, the camera on the other side of the dog would have a better view and better model reconstructions might result.

The current IFS algorithm uses a fixed size window as a tunable parameter. The algorithm may benefit from a window size that can be dynamically changed. In some cases, where the dog moves very quickly, the window size should be small. In the case where a dog loiters around an area for some time, the window size should be larger.

## **6.5 Concluding Remarks**

I would like to leave the reader with a final thought. There is no direct evidence that this work has ever contributed to saving lives in USAR environments. However, work in this domain is critical as the need for improved technologies becomes more pressing every day. It is my hope that one day, the work presented here, may contribute to saving someone's life.

## REFERENCES

- [1] M. Minsky, "Telepresence," *Omni*, 1980.
- [2] P. Goldmark, "Tomorrow we will communicate to our jobs," *The Futurist*, vol. 6, pp. 35-42, 1972.
- [3] K. M. Tsui, M. Desai, H. A. Yanco, and C. Uhlik, "Exploring use cases for telepresence robots," in *Proceedings of the 6th international conference on Human-robot interaction*, 2011, pp. 11-18.
- [4] R. W. Perry and M. K. Lindell, *Emergency planning*: Wiley, 2007.
- [5] D. A. McEntire, *Disaster Response and Recovery*: Hoboken, New Jersey: John Wiley & Sons, 2007.
- [6] M. K. Lindell, C. Prater, and R. W. Perry, *Introduction to emergency management*: Wiley, 2007.
- [7] "Population density of the United States from 1790 to 2017 in residents per square mile of land area," U.S. Bureau of the Census, Statista, 2018, Available: <https://www.statista.com/statistics/183475/united-states-population-density/>
- [8] B. Plumer. "Nine facts about terrorism in the United States since 9/11," *The Washington Post*, 11 September, 2013 Available: [https://www.washingtonpost.com/news/wonk/wp/2013/09/11/nine-facts-about-terrorism-in-the-united-states-since-911/?noredirect=on&utm\\_term=.4d004d96ea37](https://www.washingtonpost.com/news/wonk/wp/2013/09/11/nine-facts-about-terrorism-in-the-united-states-since-911/?noredirect=on&utm_term=.4d004d96ea37)
- [9] J. Daniell. "Sichuan 2008: A disaster on an immense scale," *British Broadcasting Corporation*, [News], May 9, 2013 Available: <https://www.bbc.com/news/science-environment-22398684>
- [10] "Haiti quake death toll rises to 230,000," *British Broadcasting Corporation*, [News], February 11, 2010 Available: <http://news.bbc.co.uk/2/hi/americas/8507531.stm>
- [11] "Deaths, people missing set to top 1,600: Edano," *Japan Times*, March 13, 2011
- [12] "Canadian Urban Search and Rescue (USAR) classification guide," Public Safety Canada, December 8, 2015,
- [13] Chief Raymond Downey. "Types of Collapse," *With New York Firefighters*, pp. 10-11, January, 2001

- [14] A. Jacoff, E. Messina, and J. Evans, "A standard test course for urban search and rescue robots," *Proceedings of the Performance Metrics for Intelligent Systems Workshop*, 2000.
- [15] E. Messina and A. Jacoff, "Performance standards for urban search and rescue robots," in *Proceedings of SPIE*, 2006, pp. 639-650.
- [16] E. Messina, J. M. E. Llc, and K. T. Consulting, "Standards for Visual Acuity," ASTM International Task Group on Performance Measures for Robots for Urban Search and Rescue, 2006.
- [17] E. R. Messina and A. S. Jacoff, "Measuring the performance of urban search and rescue robots," in *IEEE Conference on Technologies for Homeland Security, 2007*, Woburn, MA, USA, 2007, pp. 28-33.
- [18] M. E. Cablk, J. C. Sagebiel, J. S. Heaton, and C. Valentin, "Olfaction-based Detection Distance: A Quantitative Analysis of How Far Away Dogs Recognize Tortoise Odor and Follow It to Source," *Sensors*, vol. 8, pp. 2208-2222, 2008.
- [19] A. R. Ford and S. W. Reeve, "Sensing and characterization of explosive vapors near 700 cm," in *SPIE - International Society for Optical Engineering*, 2007, p. 65400Y.
- [20] V. D. Acree and U. S. D. o. Justice, "Customs tailored enforcement techniques: trouble for federal lawbreakers," *FBI Law Enforcement Bulletin*, vol. 45, p. 5, 1976.
- [21] D. Pickel, G. P. Manucy, D. B. Walker, S. B. Hall, and J. C. Walker, "Evidence for canine olfactory detection of melanoma," *Applied Animal Behaviour Science*, vol. 89, pp. 107-116, 2004.
- [22] A. Ferworn, A. Sadeghian, K. Barnum, D. Ostrom, H. Rahnama, and I. Woungang, "Canine as Robot in Directed Search," in *IEEE International Conference of Systems of Systems (SoSE'07)*, San Antonio, TX, USA, 2007, pp. 1-5.
- [23] J. Tran, A. Ferworn, C. Ribeiro, and M. Denko, "Enhancing canine disaster search," in *IEEE International Conference of Systems of Systems (SoSE'08)*, Monterey, CA, USA, 2008, pp. 1-5.
- [24] J. Tran, "A Telepresence System for Canine Search in an Urban Search and Rescue Environment," MSc Thesis, Computer Science, Ryerson University, Toronto, Canada, 2009.
- [25] A. Ferworn, A. Sadeghian, K. Barnum, H. Rahnama, H. Pham, C. Erickson, *et al.*, "Urban search and rescue with canine augmentation technology," in *IEEE International Conference of Systems of Systems (SoSE'06)*, Los Angeles, CA, USA, 2006, p. 5.

- [26] A. Ferworn, A. Sadeghian, K. Barnum, D. Ostrom, H. Rahnama, and I. Woungang, "Rubble Search with Canine Augmentation Technology," in *IEEE International Conference of Systems of Systems (SoSE'07)*, San Antonio, TX, USA, 2007, pp. 1-6.
- [27] A. Ferworn, D. Ostrom, K. Barnum, M. Dallaire, D. Harkness, and M. Dolderman, "Canine Remote Deployment System for Urban Search and Rescue," *Journal of Homeland Security and Emergency Management*, vol. 5, p. 9, 2008.
- [28] J. Tran and A. Ferworn, "Bark Indication Detection and Release Algorithm for the Automatic Delivery of Packages by Dogs," in *6th International Wireless Communications and Mobile Computing Conference (IWCMC 2010)*, Caen, France, 2010.
- [29] J. Tran, A. Ferworn, M. Gerdzhev, and D. Ostrom, "Canine Assisted Robot Deployment for Urban Search and Rescue," in *Safety Security and Rescue Robotics (SSRR), 2010 IEEE International Workshop on*, 2010, pp. 1-6.
- [30] M. Gerdzhev, J. Tran, A. Ferworn, and D. Ostrom, "DEX - A design for Canine-Delivered Marsupial Robot," in *Safety Security and Rescue Robotics (SSRR), 2010 IEEE International Workshop on*, 2010, pp. 1-6.
- [31] A. Ferworn, J. Tran, A. Ufkes, and A. D'Souza, "Initial Experiments on 3D Modeling of Complex Disaster Environments Using Unmanned Arial Vehicles " in *9th IEEE International Symposium on Safety, Security, and Rescue Robotics SSRR 2011*, Kyoty, Japan, 2011.
- [32] J. Tran, A. Ufkes, M. Fiala, and A. Ferworn, "Low-Cost 3D Scene Reconstruction for Response Robots in Real-time " in *9th IEEE International Symposium on Safety, Security, and Rescue Robotics SSRR 2011*, Kyoto, Japan, 2011, pp. 1-6.
- [33] J. Tran, A. Ufkes, A. Ferworn, and M. Fiala, "3D Disaster Scene Reconstruction Using a Canine-Mounted RGB-D Sensor," in *International Conference on Computer and Robot Vision (CRV), 2013* Regina, Canada, 2013, pp. 23-28.
- [34] "IEEE Standard for Information technology--Telecommunications and information exchange between systems Local and metropolitan area networks--Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007)*, pp. 1-2793, 2012.
- [35] *UCRT (USAR CBRNE) - Team Name and Description*. Last Update Date: Oct. 18, 2018, Available: <http://www.opp.ca/index.php?id=115&lng=en&entryid=572769db8f94ac7e408a7252>

- [36] C. HUTCHISON and J. BROWNSTEIN. "Girl's 15 Days in Haiti Rubble Raises Survival Questions," *ABC News*, January 29, 2010 Available: <https://abcnews.go.com/Health/HaitiEarthquake/girls-15-day-wait-haiti-rubble-raises-questions/story?id=9691529>
- [37] "IEEE Standard for Low-Rate Wireless Networks," *IEEE Std 802.15.4-2015 (Revision of IEEE Std 802.15.4-2011)*, pp. 1-709, 2016.
- [38] A. Ferworn, N. Tran, J. Tran, G. Zarnett, and F. Sharifi, "WiFi repeater deployment for improved communication in confined-space urban disaster search," in *IEEE International Conference of Systems of Systems (SoSE'07)*, San Antonio, TX, USA, 2007.
- [39] M. Başoğlu, C. Kiliç, E. Şalcioğlu, and M. Livanou, "Prevalence of posttraumatic stress disorder and comorbid depression in earthquake survivors in Turkey: an epidemiological study," *Journal of Traumatic Stress: Official Publication of The International Society for Traumatic Stress Studies*, vol. 17, pp. 133-141, 2004.
- [40] T. Yoshida, E. Koyanagi, S. Tadokoro, K. Yoshida, K. Nagatani, K. Ohno, *et al.*, "A high mobility 6-crawler mobile robot iKenafi," in *4th International Workshop on Synthetic Simulation and Robotics to Mitigate Earthquake Disaster (SRMED2007)*, Atlanta, USA, 2007, p. 38.
- [41] K. Ohno and T. Yoshida, "RoboCupRescue 2009-Robot League Team< Pelican United (JAPAN)>," *RoboCup Team Description Paper*, 2009.
- [42] M. Guarnieri, P. Debenest, T. Inoh, E. Fukushima, and S. Hirose, "Helios VII: a new vehicle for disaster response-mechanical design and basic experiments," *Advanced Robotics*, vol. 19, pp. 901-927, 2005.
- [43] M. Arai, Y. Tanaka, S. Hirose, H. Kuwahara, and S. Tsukui, "Development of iSouryu-IVi and iSouryu-V: i Serially connected crawler vehicles for in-rubble searching operations," *Journal of Field Robotics*, vol. 25, pp. 31-65, 2007.
- [44] B. Bruggemann, B. Gaspers, A. Ciossek, J. Pellenz, and N. Kroll, "Comparison of different control methods for mobile manipulation using standardized tests," in *Safety, Security, and Rescue Robotics (SSRR), 2013 IEEE International Symposium on*, 2013, pp. 1-2.
- [45] R. R. Murphy and J. L. Burke, "Up from the rubble: Lessons learned about HRI from search and rescue," in *Proceedings of the 49th Annual Meetings of the Human Factors and Ergonomics Society*, Orlando, USA, 2005.
- [46] J. Burke and R. Murphy, "RSVP: an investigation of remote shared visual presence as common ground for human-robot teams," in *Proceedings of the ACM/IEEE international*

- conference on Human-robot interaction (HRI '07), Arlington, Virginia, USA, 2007, pp. 161-168.
- [47] R. R. Murphy, "Trial by fire [rescue robots]," *Robotics & Automation Magazine, IEEE*, vol. 11, pp. 50-61, 2004.
- [48] Z. Wang and H. Gu, "A review of locomotion mechanisms of urban search and rescue robot," *Industrial Robot: An International Journal*, vol. 34, pp. 400-11, 2007.
- [49] A. Ferworn, G. Hough, R. Manca, B. Antonishek, J. Scholtz, and A. Jacoff, "Expedients for Marsupial Operations of USAR Robots," in *IEEE International Workshop on Safety, Security and Rescue Robotics (SSRR06)*, Gaithersburg, MD, USA, 2006.
- [50] R. R. Murphy, "Marsupial and shape-shifting robots for urban search and rescue," *Intelligent Systems and Their Applications, IEEE [see also IEEE Intelligent Systems]*, vol. 15, pp. 14-19, 2000.
- [51] J. Tran, M. Gerdzhev, and A. Ferworn, "Continuing Progress in Augmenting Urban Search and Rescue Dogs," in *6th International Wireless Communications and Mobile Computing Conference (IWCMC 2010)*, Caen, France, 2010.
- [52] Texas Engineering Extension Services. *Disaster City*. Last Update Date: Oct. 18, 2018, Available: <https://teex.org/Pages/about-us/disaster-city.aspx>
- [53] N. Mimura, K. Yasuhara, S. Kawagoe, H. Yokoki, and S. Kazama, "Damage from the Great East Japan Earthquake and Tsunami-A quick report," *Mitigation and Adaptation Strategies for Global Change*, pp. 1-16, 2011.
- [54] E. Fish. "Radiation Detector Robot Deployed in Fukushima," *PCWorld (Online)*, 2011 Available: [http://www.pcworld.com/article/222947/radiation\\_detector\\_robot\\_deployed\\_in\\_fukushima.html](http://www.pcworld.com/article/222947/radiation_detector_robot_deployed_in_fukushima.html)
- [55] W. Heller. (2011). *Disaster Robots 2011 - Report from Japan*. Last Update Date: Oct. 18, 2018, Available: <http://robotland.blogspot.com/2011/05/disaster-robots-2011-report-from-japan.html>
- [56] C. M. Humphrey and J. A. Adams, "Robotic Tasks for Chemical, Biological, Radiological, Nuclear and Explosive Incident Response," *Advanced Robotics*, vol. 23, pp. 1217-1232, 2009.
- [57] T. Johnson, J. Metcalfe, B. Brewster, C. Manteuffel, M. Jaswa, and T. Tierney, "Human-robot interaction modeling and simulation of supervisory control and situational awareness

- during field experimentation with military manned and unmanned ground vehicles," in *SPIE - The International Society for Optical Engineering Proc. SPIE Int. Soc. Opt. Eng.*, Orlando, Florida, USA, 2010, p. 76920A.
- [58] J. Scholtz, B. Antonishek, and J. Young, "A field study of two techniques for situation awareness for robot navigation in urban search and rescue," in *IEEE International Workshop on Robot and Human Interactive Communication, 2005. ROMAN 2005.*, 2005 pp. 131-136.
- [59] J. Riley, R. Murphy, and M. Endsley, "Situation awareness in the control of unmanned ground vehicles," *Human factors of remotely operated vehicles*, vol. 7, pp. 359-371, 2006.
- [60] D. Dahn and M. Gacy, "Human control of unmanned systems under stressful conditions," in *1st Joint Emergency Preparedness and Response/Robotic and Remote Systems Topical Meeting*, Salt Lake City, UT, USA, 2006, pp. 479-486.
- [61] G. S. Virk, Y. Gatsoulis, M. Parack, and A. Kherada, "Mobile Robotic Issues for Urban Search and Rescue," in *Proceedings of the 17th World Congress The International Federation of Automatic Control*, Seoul, Korea, 2008, pp. 3098-3103.
- [62] Z. Aiyun, Y. Kui, Y. Zhigang, and Z. Haibing, "Research and application of a robot orientation sensor," in *IEEE International Conference on Robotics, Intelligent Systems and Signal Processing, 2003. Proceedings. 2003*, 2003, pp. 1069-1074 vol. 2.
- [63] H. A. Yanco and J. Drury, "Where am I? Acquiring situation awareness using a remote robot platform," in *IEEE International Conference on Systems, Man and Cybernetics, 2004* 2004, pp. 2835-2840 vol. 3.
- [64] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," in *Sensor Fusion IV: Control Paradigms and Data Structures*, 1992, pp. 586-607.
- [65] T. Jost and H. Hügli, "Fast ICP algorithms for shape registration," in *Joint Pattern Recognition Symposium*, 2002, pp. 91-99.
- [66] A. Nuchter, K. Lingemann, J. Hertzberg, and H. Surmann, "6D SLAM with approximate data association," in *Advanced Robotics, 2005. ICAR'05. Proceedings., 12th International Conference on*, 2005, pp. 242-249.
- [67] A. Nuchter, K. Lingemann, and J. Hertzberg, "Cached kd tree search for ICP algorithms," in *3-D Digital Imaging and Modeling, 2007. 3DIM'07. Sixth International Conference on*, 2007, pp. 419-426.

- [68] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on*, 2001, pp. 145-152.
- [69] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International journal of computer vision*, vol. 47, pp. 7-42, 2002.
- [70] D. Nister, O. Naroditsky, and J. Bergen, "Visual odometry," in *IEEE Computer Society Conf. Computer Vision and Pattern Recognition CVPR 2004*, Washington, DC, 2004, pp. I-652-I-659 Vol.1.
- [71] Wikipedia. (2011). *Kinect*. Last Update Date: Oct. 18, 2018, Available: <http://en.wikipedia.org/wiki/kinect>.
- [72] OpenKinect. (2011). *OpenKinect project*. Last Update Date: Oct. 18, 2018, Available: [http://openkinect.org/wiki/Main\\_Page](http://openkinect.org/wiki/Main_Page)
- [73] F. Steinbrucker, J. Sturm, and D. Cremers, "Real-time visual odometry from dense RGB-D images," in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, 2011, pp. 719-722.
- [74] M. Fiala and A. Ufkes, "Visual Odometry Using 3-Dimensional Video Input," in *poster session of the 8th Canadian Conference on Computer and Robot Vision (CRV 2011)*, St. John's, Newfoundland, Canada, 2011.
- [75] N. Burrus. (2011). *Kinect Calibration*. Last Update Date: Oct. 18, 2018, Available: <http://nicolas.burrus.name/index.php/Research/KinectCalibration>
- [76] S. Magnenat. (2010). Last Update Date: Oct. 18, 2018, Available: [https://groups.google.com/group/openkinect/browse\\_thread/thread/31351846fd33c78/e98a94ac605b9f21?lnk=gst&q=stephane&pli=1](https://groups.google.com/group/openkinect/browse_thread/thread/31351846fd33c78/e98a94ac605b9f21?lnk=gst&q=stephane&pli=1)
- [77] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, pp. 381-395, 1981.
- [78] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," *Computer Vision–ECCV 2006*, vol. 3951, pp. 404-417, 2006.
- [79] J. Shi and C. Tomasi, "Good features to track," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994*, Seattle, WA, USA 1993, pp. 593-600.

- [80] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," *Computer Vision–ECCV 2006*, pp. 430-443, 2006.
- [81] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," in *Tenth IEEE International Conference on Computer Vision (ICCV'05)*, Beijing, China, 2005.
- [82] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of Imaging Understanding Workshop*, 1981, pp. 674–679.
- [83] M. Onosato, F. Takemura, K. Nonami, K. Kawabata, K. Miura, and H. Nakanishi, "Aerial Robots for Quick Information Gathering in USAR," in *SICE-ICASE, 2006. International Joint Conference* Busan, 2006, pp. 3435-3438.
- [84] U. Olsson, "Data management in telco networks: From costly duckling to profitable swan," in *14th International Conference on Intelligence in Next Generation Networks (ICIN), 2010* Berlin, 2010, pp. 1-4.
- [85] R. Ramesham, "Environmental testing of COTS components for space applications," in *Reliability, Packaging, Testing, and Characterization of MEMS/MOEMS and Nanodevices VIII*, San Jose, CA, USA, 2009, pp. 72060H-1 to 72060H-11.
- [86] R. Bolla, R. Bruschi, and A. Ranieri, "Performance and power consumption modeling for green COTS Software Router," in *First International Communication Systems and Networks and Workshops, 2009. COMSNETS 2009*, Bangalore 2009, pp. 1-8.
- [87] P. Lister, "The impact of COTS on a ground support system procurement," in *The AIAA/IEEE/SAE Digital Avionics Systems Conference, 1998. Proceedings., 17th DASC*, Bellevue, WA , USA 1998, pp. D27/1-D27/7 vol. 1.
- [88] T. Thames, "Using commercial off-the-shelf (COTS) equipment to meet joint service requirements," in *1998 IEEE AUTOTESTCON '98. IEEE Systems Readiness Technology Conference.,* Salt Lake City, UT , USA 1998, pp. 204-209.
- [89] Y. Cheng, M. Maimone, and L. Matthies, "Visual odometry on the Mars exploration rovers," in *Systems, Man and Cybernetics, 2005 IEEE International Conference on*, 2005, pp. 903-910.
- [90] M. Fiala and A. Ufkes, "Visual Odometry Using 3-Dimensional Video Input," in *Computer and Robot Vision (CRV), 2011 Canadian Conference on*, 2011, pp. 86-93.

- [91] F. Endres, J. Hess, N. Engelhard, J. Sturm, D. Cremers, and W. Burgard, "An evaluation of the RGB-D SLAM system," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, 2012, pp. 1691-1696.
- [92] V. Guizilini and F. Ramos, "Visual odometry learning for unmanned aerial vehicles," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, 2011, pp. 6213-6220.
- [93] A. S. Huang, A. Bachrach, P. Henry, M. Krainin, D. Maturana, D. Fox, *et al.*, "Visual odometry and mapping for autonomous flight using an RGB-D camera," in *International Symposium on Robotics Research (ISRR)*, 2011.
- [94] A. Howard, "Real-time stereo visual odometry for autonomous ground vehicles," in *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, 2008, pp. 3946-3952.
- [95] Y. Takaoka, Y. Kida, S. Kagami, H. Mizoguchi, and T. Kanade, "3D map building for a humanoid robot by using visual odometry," in *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, 2004, pp. 4444-4449.
- [96] P. Michel, J. Chestnutt, S. Kagami, K. Nishiwaki, J. Kuffner, and T. Kanade, "Online environment reconstruction for biped navigation," in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, 2006, pp. 3089-3094.
- [97] A. Pretto, E. Menegatti, M. Bennewitz, W. Burgard, and E. Pagello, "A visual odometry framework robust to motion blur," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*, 2009, pp. 2250-2257.
- [98] H. A. Wurdemann, E. Georgiou, L. Cui, and J. S. Dai, "SLAM Using 3D Reconstruction via a Visual RGB and RGB-D Sensory Input," 2011.
- [99] R. A. Newcombe, A. J. Davison, S. Izadi, P. Kohli, O. Hilliges, J. Shotton, *et al.*, "KinectFusion: Real-time dense surface mapping and tracking," in *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*, 2011, pp. 127-136.
- [100] T. Whelan, M. Kaess, M. Fallon, H. Johannsson, J. Leonard, and J. McDonald, "Kintinuous: Spatially Extended KinectFusion," in *RSS Workshop on RGB-D: Advanced Reasoning with Depth Cameras*, Sydney, Australia, 2012.
- [101] "Point Cloud Library (PCL)," 1.6.0 ed: Willow Garage, 2012

- [102] J. Tran, A. Ufkes, M. Fiala, and A. Ferworn, "Low-cost 3D scene reconstruction for response robots in real-time," in *2011 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, , 2011, pp. 161-166.
- [103] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th international joint conference on Artificial intelligence*, 1981.
- [104] M. Raibert, K. Blankespoor, G. Nelson, and R. Playter, "Bigdog, the rough-terrain quadruped robot," in *Proceedings of the 17th World Congress*, 2008, pp. 10823-10825.
- [105] *PETMAN*. Last Update Date: Oct. 18, 2018, Available: [http://www.bostondynamics.com/robot\\_petman.html](http://www.bostondynamics.com/robot_petman.html)
- [106] *Google Glasses*. Last Update Date: Oct. 18, 2018, Available: <http://www.google.com/glass/start/>
- [107] C. Wright, A. Buchan, B. Brown, J. Geist, M. Schwerin, D. Rollinson, *et al.*, "Design and architecture of the unified modular snake robot," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, 2012, pp. 4347-4354.
- [108] C. Wright, A. Johnson, A. Peck, Z. McCord, A. Naaktgeboren, P. Gianfortoni, *et al.*, "Design of a modular snake robot," in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, 2007, pp. 2609-2614.
- [109] T. Kamegawa, T. Yamasaki, H. Igarashi, and F. Matsuno, "Development of the snake-like rescue robot," in *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, 2004.
- [110] B. M. Yamauchi, "PackBot: a versatile platform for military robotics," in *The International Society of Photo-Optical Instrumentation Engineers*, Orlando, FL, USA 2004, p. 228.
- [111] A. Ferworn, "Canine Augmentation Technology for Urban Search and Rescue," in *Canine Ergonomics: The Science of Working Dogs*, ed, 2009, p. 205.
- [112] C. Ribeiro, A. Ferworn, M. Denko, and J. Tran, "Canine Pose Estimation: A Computing for Public Safety Solution," in *2009 Canadian Conference on Computer and Robot Vision (CRV 2009)*, Kelowna, BC, Canada., 2009, pp. 37-44.
- [113] H. G. Nguyen, H. R. Everett, N. Manouk, A. Verma, Space, and C. A. Naval Warfare Systems Center San Diego, *Autonomous Mobile Communication Relays: Defense Technical Information Center*, 2002.

- [114] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer Networks*, vol. 47, pp. 445-487, 2005.
- [115] A. Ferworn, J. Tran, A. Ufkes, S. Herman, and C. Kong, "Establishing network connectivity under rubble using a hybrid wired and wireless approach," in *Safety, Security, and Rescue Robotics (SSRR), 2012 IEEE International Symposium on*, 2012, pp. 1-2.