

Collaborative Micro Aerial Vehicle Exploration of Outdoor Environments

by

David Pitman

S.B., Massachusetts Institute of Technology (2006)

Submitted to the Department of Electrical Engineering and Computer Science

in partial fulfillment of the requirements for the Degree of

Master of Engineering in Electrical Engineering and Computer Science

at the Massachusetts Institute of Technology

February 2010

©2010 Massachusetts Institute of Technology

All rights reserved.

Author
Department of Electrical Engineering and Computer Science
February 10, 2010

Certified by
Mary L. Cummings
Associate Professor of Aeronautics and Astronautics
Thesis Supervisor

Accepted by
Arthur C. Smith
Professor of Electrical Engineering
Chairman, Department Committee on Graduate Theses

Report Documentation Page		Form Approved OMB No. 0704-0188
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.		
1. REPORT DATE FEB 2010	2. REPORT TYPE	3. DATES COVERED 00-00-2010 to 00-00-2010
4. TITLE AND SUBTITLE Collaborative Micro Aerial Vehicle Exploration of Outdoor Environments		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER	
	5e. TASK NUMBER	
	5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, 02139		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited		
13. SUPPLEMENTARY NOTES		

14. ABSTRACT

Field personnel, such as soldiers, police SWAT teams, and first responders, face challenging, dangerous environments, often with little advance knowledge or information about their surroundings. Currently this Intelligence, Surveillance & Reconnaissance (ISR) information is provided by satellite imagery and prior or second-hand experiences. Although satellite imagery is currently the preferred method for gaining Situational Awareness (SA) about an outdoor environment, it has many shortcomings. Unclassified satellite imagery maps available to these field personnel are flat images, with no elevation information and fixed points of view. These maps are often outdated, and, due to shadows and shading, give false impressions of elevations and details of the environment. Critical features of buildings, such as doorways and windows are hidden from view. Combined, these flaws often give field personnel a false mental model of their environment. Given the need of these personnel to simultaneously perform a primary task, such as finding a Person of Interest (POI), as well as explore the environment, an autonomous robot would allow these groups to better perform ISR and improve their SA in real-time. Recent efforts have led to the creation of Micro Aerial Vehicles (MAVs), a class of Unmanned Aerial Vehicle (UAV), which are small and have autonomous capabilities. At most a few feet in size, a MAV can hover in place, perform Vertical Take-Off and Landing, and easily rotate with a small sensor payload. The compact size of these vehicles and their maneuvering capabilities make them well-suited for performing highly localized ISR missions with MAV operator working within the same environment as the vehicle. Unfortunately, existing interfaces for MAVs ignore the needs of field operators, requiring bulky equipment and the operator's full attention. To be able to collaboratively explore an environment with a MAV, an operator needs a mobile interface which can support the need for divided attention. To address this need, a Cognitive Task Analysis (CTA) was performed with the intended users of the interface to assess their needs, as well as the roles and functions a MAV could provide. Based on this CTA, a set of functional and information requirements were created which outlined the necessities of an interface for exploring an environment with a MAV. Based on these requirements, the Micro Aerial Vehicle Exploration of an Unknown Environment (MAV-VUE) interface was designed and implemented. Using MAV-VUE, operators can navigate the MAV using waypoints, which requires little attention. When the operator needs more fine-grained control over the MAV's location and orientation, in order to obtain imagery or learn more about an environment, he or she can use the Nudge Control mode. Nudge Control uses Perceived First Order (PFO) control to allow an operator to effectively "nudge" a MAV with no risk to the

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:

a. REPORT
unclassified

b. ABSTRACT
unclassified

c. THIS PAGE
unclassified

17. LIMITATION OF
ABSTRACT

**Same as
Report (SAR)**

18. NUMBER
OF PAGES

124

19a. NAME OF
RESPONSIBLE PERSON

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

1	Introduction	19
1.1	Micro Aerial Vehicles	21
1.1.1	MAV Operators	22
1.1.2	MAV GCSs	23
1.2	Mobile Devices	26
1.3	Problem Statement	26
1.4	Research Objectives	27
1.5	Thesis Organization	27
2	Background	29
2.1	Human-Robot Interaction	29
2.1.1	Human Supervisory Control	29
2.1.2	Human-Robot Interaction Roles	31
2.1.3	Urban Search and Rescue (USAR)	32
2.2	Research Relevant to MAV Interaction	33
2.3	Research in Hand-held Devices	36
2.4	Summary	37
3	Interface Design	39
3.1	Introduction	39
3.2	Cognitive Task Analysis	39
3.3	MAV Interface Requirements	40
3.3.1	Desired MAV Roles & Requirements	40
3.3.2	Interface Requirements	42
3.4	MAV-VUE Displays and Interaction	45
3.4.1	Navigation Mode: Map & Waypoints	45
3.4.2	Mini Vertical Altitude and Velocity Indicator	47
3.4.3	Health & Status Monitoring	48
3.4.4	Nudge Control Flight Mode	48

3.5	Design Principles	50
3.5.1	Direct Manipulation	50
3.5.2	Multimodal Interaction	50
3.5.3	Order Reduction	51
3.5.4	Perceived First Order Control	53
3.6	Architecture	59
3.7	Summary	60
4	Usability Evaluation	63
4.1	Study Objectives	63
4.2	Research Questions	63
4.3	Participants	64
4.4	Test Bed	64
4.4.1	Apparatus	64
4.5	Experiment Metrics	66
4.5.1	Task Performance Time	66
4.5.2	Spatial Reasoning	66
4.5.3	Additional Mission Performance Metrics	66
4.5.4	Qualitative Metrics	67
4.6	Procedure	67
4.7	Data Collection	69
4.8	Summary	69
5	Results and Discussion	71
5.1	Introduction	71
5.2	Score Task Performance Time	71
5.3	Eye Chart Identification	72
5.4	Person of Interest Identification	74
5.5	Nudge Control Analysis	75
5.5.1	Webcam Imagery and Frame Rate	76
5.6	Participants' Navigation Strategies	76
5.6.1	Participant A	77
5.6.2	Participant B	78
5.6.3	Participant C	79
5.7	Subjective Responses	80
5.7.1	Navigation Mode	80
5.7.2	Nudge Control	80

5.8	Experiment Observations	82
5.9	Summary	83
6	Conclusions	85
6.1	Research Objectives and Findings	86
6.2	Future Work	87
A	Cognitive Task Analysis Interview Questions	89
B	Demographic Survey and Statistics	91
C	Configuration of the Micro Aerial Vehicle (MAV)	95
D	Usability Survey	97
E	Consent to Participate Form	103
F	Maps	109
G	Snellen Eye Chart	111
H	Person of Interest Used and Identification Sheet	113
I	Debriefing Interview Questions	117
J	Scored Task Descriptive Statistics	119
	References	121

THIS PAGE INTENTIONALLY LEFT BLANK

List of Figures

1-1	Comparison of overhead satellite imagery of a building with Three Dimensional (3D) model of the same building, courtesy of Google Earth™.	20
1-2	A Honeywell RQ-16 T-Hawk in flight, courtesy of U.S. Navy.	21
1-3	Examples of commercial MAVs available on the market.	22
1-4	Examples of software interfaces for commercial MAVs.	23
1-5	Examples of MAV Ground Control Stations.	24
1-6	Video goggles for AeroVironment's Ground Control Station (GCS), photo by author. . . .	25
1-7	An Army operator controlling a RQ-16 with a stylus, courtesy of U.S. Army.	25
2-1	General Human Supervisory Control, adapted from Sheridan [7].	29
2-2	Hierarchical Control Loops for a Single MAV performing an Intelligence, Surveillance & Reconnaissance (ISR) mission, adapted from Cummings et al. [11]	30
2-3	MAV control interfaces created by Durlach et al. [29]	35
3-1	The map display and general interface of MAV-VUE.	45
3-2	The inset webcam view within the map display.	46
3-3	Details of the MAV-VUE map display.	46
3-4	Waypoint states in the map display	47
3-5	The Mini-VAVI shown on the map display.	47
3-6	Health and Status monitoring display in MAV-VUE.	48
3-7	Overview of Nudge Control interface.	49
3-8	Details of Nudge Control directional interface.	49
3-9	1 st order feedback control with a pulse input.	51
3-10	Interface feedback as a result of the operator performing a natural tilt gesture with the device.	53
3-12	Using a swipe touch gesture to move the MAV left	54
3-11	Diagram of Perceived First Order Control. Delays and lags are represented by τ , gains represented by k	55
3-13	Discrete step function initially used for gain, k	56
3-14	Linear step function used for gain, k	57

3-15	Gestures to change the z coordinate (altitude) of the MAV	58
3-16	Using the VAVI to change the z coordinate of the MAV.	58
3-17	Swiping a finger across the screen causes the device to rotate (yaw) right or left.	59
3-18	Communication architecture between Micro Aerial Vehicle Exploration of an Unknown Environment (MAV-VUE), MAVServer, and MAV.	60
4-1	Modified Ascending Technologies Hummingbird AutoPilot used in study.	65
4-2	Annotated version of the map used in the study showing the layout of the environment.	68
5-1	Box plot of participants' times to complete the scored task.	72
5-2	Images which participants saw while trying to read a line of the eye chart.	73
5-3	Examples of blurred imagery seen by participants while trying to identify the eye chart.	73
5-4	Examples of imagery seen by participants while finding and identifying the POI.	74
5-5	Participant A's flight path.	77
5-6	Participant B's flight path.	78
5-7	Participant C's flight path.	79
5-8	Intended and observed interaction to issue a Nudge Control yaw command.	81
5-9	Resting poses observed while participants used the interface, courtesy of Nadya Peek.	82
F-1	The map displayed on the subject's iPod.	109
F-2	Supplementary map given to subject for their scored task.	110
G-1	Example of the standard Snellen eye chart used in the study, by Jeff Dahl. Licensed under the Creative Commons Attribution-Share Alike 3.0 Unported License.	112
H-1	Person of Interest (POI) used in the scored task of the usability study.	114
H-2	POI identification sheet used in usability study.	115

List of Tables

2.1	Sheridan and Verplank Levels of Automation [8]	30
2.2	Scholtz’s Human-Robot Interaction (HRI) roles, with Goodrich and Schulz’s additions, adapted from [12, 13]	31
3.1	Potential roles for an operator and a MAV collaboratively exploring an outdoor environment.	40
3.2	Potential MAV outdoor ISR missions	41
3.3	System requirements for a MAV performing an outdoor ISR mission.	42
3.4	Functional requirements for a HRI interface to support collaborative exploration of an unknown environment.	43
3.5	Information requirements to support collaborative exploration of an unknown environment.	44
B.1	Responses for self-assessed sleep quality	94
B.2	Responses for self-assessed video game experience	94
B.3	Responses for self-assessed Remote Control (RC) vehicle experience	94
B.4	Responses for self-assessed comfort with Google Maps™, Google Earth™, and iPhone	94
D.1	Responses for self-assessed confidence in actions during scored task.	101
D.2	Responses for self-assessed performance during scored task.	101
D.3	Responses for self-assessed understanding of interface.	101
J.1	Performance Descriptive Statistics.	119
J.2	Times Descriptive Statistics.	120
J.3	Descriptive statistics of Nudge Control commands performed by participants during the scored task.	120

THIS PAGE INTENTIONALLY LEFT BLANK

Collaborative Micro Aerial Vehicle Exploration of Outdoor Environments

by

David Pitman

Submitted to the
Department of Electrical Engineering and Computer Science

February 10, 2010

In partial fulfillment of the requirements for the Degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

Field personnel, such as soldiers, police SWAT teams, and first responders, face challenging, dangerous environments, often with little advance knowledge or information about their surroundings. Currently, this Intelligence, Surveillance & Reconnaissance (ISR) information is provided by satellite imagery and prior or second-hand experiences. Although satellite imagery is currently the preferred method for gaining Situational Awareness (SA) about an outdoor environment, it has many shortcomings. Unclassified satellite imagery maps available to these field personnel are flat images, with no elevation information and fixed points of view. These maps are often outdated, and, due to shadows and shading, give false impressions of elevations and details of the environment. Critical features of buildings, such as doorways and windows are hidden from view. Combined, these flaws often give field personnel a false mental model of their environment.

Given the need of these personnel to simultaneously perform a primary task, such as finding a Person of Interest (POI), as well as explore the environment, an autonomous robot would allow these groups to better perform ISR and improve their SA in real-time. Recent efforts have led to the creation of Micro Aerial Vehicles (MAVs), a class of Unmanned Aerial Vehicle (UAV), which are small and have autonomous capabilities. At most a few feet in size, a MAV can hover in place, perform Vertical Take-Off and Landing, and easily rotate with a small sensor payload. The compact size of these vehicles and their maneuvering capabilities make them well-suited for performing highly localized ISR missions with MAV operator working within the same environment as the vehicle. Unfortunately, existing interfaces for MAVs ignore the needs of field operators, requiring bulky equipment and the operator's full attention.

To be able to collaboratively explore an environment with a MAV, an operator needs a mobile interface which can support the need for divided attention. To address this need, a Cognitive Task Analysis (CTA) was performed with the intended users of the interface to assess their needs, as well as the roles and functions a MAV could provide. Based on this CTA, a set of functional and information requirements were created which outlined the necessities of an interface for exploring an environment with a MAV. Based on these requirements, the Micro Aerial Vehicle Exploration of an Unknown Environment (MAV-VUE) interface was designed and implemented. Using MAV-VUE, operators can navigate the MAV using waypoints, which requires little attention. When the operator needs more fine-grained control over the MAV's location and orientation, in order to obtain imagery or learn more about an environment, he or she can use the Nudge Control mode. Nudge Control uses Perceived First Order (PFO) control to allow an operator effectively "fly" a MAV with no risk to the vehicle. PFO control, which was invented for MAV-VUE, utilizes a 0th order feedback control loop to fly the MAV, while presenting 1st order controls to the operator.

A usability study was conducted to evaluate MAV-VUE. Participants were shown a demonstration of the interface and only given three minutes of training before they performed the primary task. During

this task, participants were given search and identify objectives, MAV-VUE installed on an iPhone[®] and an actual MAV to explore a GPS-simulated urban environment. Participants performed well at the task, with thirteen of fourteen successfully performing their objectives with no crashes or collisions. Several statistically significant correlations were found between participants' performance and their usage of the interface. Operators who were more patient and had higher scores on a spatial orientation pretest tended to have more precise MAV control. Future design and implementation recommendations learned from this study are discussed.

Thesis Supervisor: Mary L. Cummings

Title: Associate Professor of Aeronautics and Astronautics

Acknowledgements

The success of this thesis is only in part to due my efforts, with many other individuals providing help, their time, advice and friendship throughout its course.

First, I would like to thank Missy Cummings, my research advisor at MIT. Your hard work ethic, unrelenting attention to details, and ingenuity have been a constant source of inspiration for me. I have always appreciated the time you take to sit down and answer my questions or explain concepts. HAL has been a great place to be, and the question you are trying to answer will always be of interest me.

Research at MIT is not true research without UROPs. Thank you to Paul Quimby for always finding a solution to seemingly impossible problems, your hardwork and persistence will serve you well, and to Carine Abi Akar, for calmly handling our participants and doing whatever needed to be done.

Often while collaborating with the ACL, I found myself with a question or quad rotor in need of repair. I would like to thank Buddy for his patient assistance and hard efforts to help us, as well as the rest of the ACL for making themselves available to answer questions.

Gracious thanks are necessary to Thomas Villaren, who helped me to first develop MAV-VUE, and to Pierre Maere, who endured long hours of soldering and quad rotor duty, always with a smile.

Special thanks go to Cam, for acting as an endless source of Course 16 information. Also, thank you to Jeff for all the great discussions about new ideas.

My friends in HAL, Anna, Andrew, Christin, Dan, Farsan, Jackie, and many others, thank you for all the conversation, fun, and happy hours at the R&D Pub.

I would like to thank my parents, who have always served as my best role models. Thank you for the good cooking and skiing, the good advice and understanding when I was too busy too talk. Thank you to my sister, Lisa, for showing me what being truly dedicated to your passions mean and always having a good joke on hand.

Lots of thanks to the Slugs and my friends at the Media Lab.

Emily, a few sentences can in no way describe how you have helped me over the years, so I will not try. Instead, know all of the appreciation I have and the fact that every day is more fun because you are around.

I would also like to thank the Office of Naval Research and The Boeing Company for providing the opportunity for me to work on this research.

THIS PAGE INTENTIONALLY LEFT BLANK

Nomenclature

2D	Two Dimensional	54
3D	Three Dimensional	7
AOI	Area of Interest	22
ch	channel	64
CT	Conventional Touch	48
CTA	Cognitive Task Analysis	27
DARPA	Defense Advanced Research Projects Agency	21
DoD	United States Department of Defense	19
DOF	Degree of Freedom	
dpi	dots per inch	26
fps	frames per second	65
ft	foot	56
FOV	Field of View	33
GB	Gigabytes	64
GHz	Gigahertz	64
GCS	Ground Control Station	21
GPS	Global Positioning System	22
GUI	Graphical User Interface	32
HCI	Human-Computer Interaction	32
HRI	Human-Robot Interaction	26
HSC	Human Supervisory Control	29
Hz	Hertz	57
IMU	Inertial Measurement Unit	48

in	inches.....	21
IP	Internet Protocol	60
ISR	Intelligence, Surveillance & Reconnaissance	19
JSON	JavaScript Object Notation	60
km/s	kilometers/second.....	35
lb	pound	21
LIDAR	Light Detection and Ranging	
LOA	Level of Automation	29
mA	milli-amperes	64
MAUI	Minimal Attention User Interface	36
MAV	Micro Aerial Vehicle	20
MAV-VUE	Micro Aerial Vehicle Exploration of an Unknown Environment	39
Mb	Megabits	64
MHz	Megahertz	64
MOLLE	Modular Lightweight Load-carrying Equipment	21
MRT	Mental Rotation Test	66
MVC	Model-Viewer-Controller	59
NG	Natural Gesture	48
NBC	Nuclear, Biological and Chemical	
NFZ	No Fly Zone	56
OOI	Object of Interest.....	35
PDA	Personal Digital Assistant.....	26
PFO	Perceived First Order	53
POI	Person of Interest.....	20
PTSOT	Perspective Taking and Spatial Orientation Test	66
px	pixel.....	45
RAVEN	Real-time indoor Autonomous Vehicle test ENvironment	64
RC	Remote Control.....	64
SA	Situational Awareness.....	20
sd	standard deviation	64

sec	second	36
SLAM	Simultaneous Localization and Mapping	56
SWAT	Special Weapons and Tactics	39
TCP	Transmission Control Protocol	60
UAS	Unmanned Aerial System	21
UAV	Unmanned Aerial Vehicle	19
UGV	Unmanned Ground Vehicle	32
UDP	User Datagram Protocol	65
UI	User Interaction	45
UV	Unmanned Vehicle	26
USB	Universal Serial Bus	65
USAR	Unmanned Search and Rescue	20
VAVI	Vertical Altitude and Velocity Indicator	47
VTOL	Vertical Take-Off and Landing	21
WYSIWYG	What-You-See-Is-What-You-Get	34

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 1

Introduction

Obtaining Intelligence, Surveillance & Reconnaissance (ISR) information in real-time is a top priority for the United States Department of Defense (DoD), as well as other first responder and homeland defense agencies. Satellite imagery, Unmanned Aerial Vehicles (UAVs) and other advances have revolutionized how the military maintains situational awareness about a battlefield. These advances have allowed the military to obtain much more information while simultaneously removing the need for soldiers to perform risky *in situ* ISR missions. However, many of these technologies only provide a larger overview of a situation at periodic points in time. For the warfighter on the ground, means of obtaining real-time intelligence and information about their environment have lagged behind the sophisticated techniques used by battlefield commanders. Recently, however, advances in autonomous unmanned vehicles and mobile computing have created new opportunities to outfit a soldier with a personal robot capable of performing customized, detailed exploration of a local environment.

Soldiers entering a dense city have a poor understanding of the topography and urban environment they will encounter. Currently the best tool available to help soldiers understand a city environment is unclassified satellite imagery. While satellite imagery provides an understanding of the urban layout of a city, it lacks information which is critical to an understanding of the environment at a ground level. A soldier may need to enter a building, jump a fence without knowing what is on the other side, or determine where potential secondary entries and exits are located in a building. Overshadowing all of these potential actions is the need to have a good understanding of the immediate vicinity and tactical advantages such as low walls, grassy openings, clear sight lines, and streets or alleys which dead-end. None of these questions can easily be answered from a satellite map alone, and often soldiers simply rely on past or second hand experience with the environment.

Satellite maps, which are currently the standard for environment ISR, still have many inherent flaws.

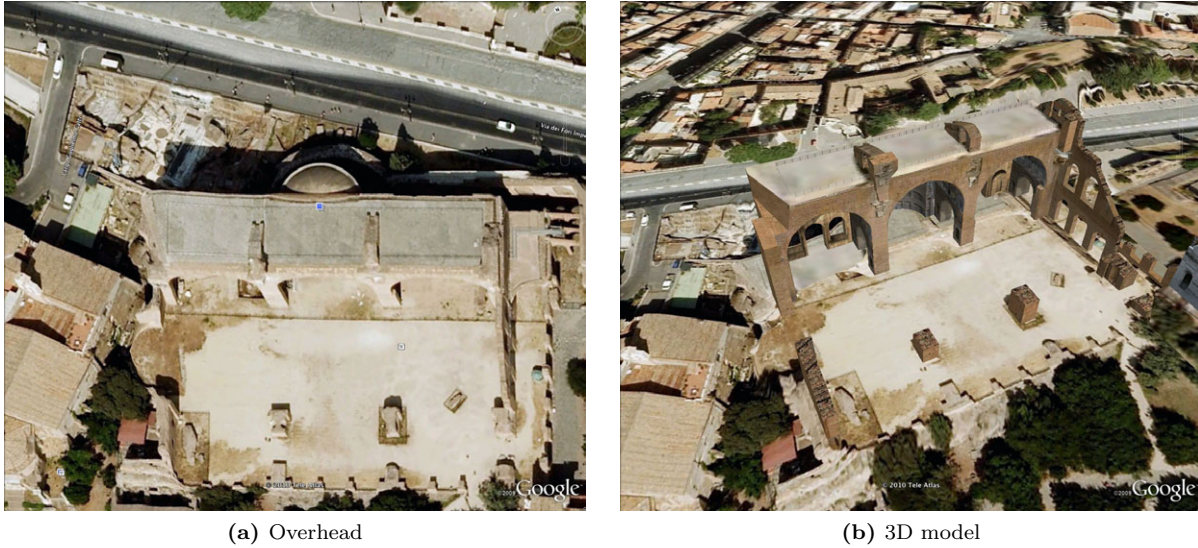


Figure 1-1: Comparison of overhead satellite imagery of a building with 3D model of the same building, courtesy of Google Earth™.

A flat image, these maps give no elevation information, and often, due to shadows and shading, give false impressions of elevation. For example, while it can be safely assumed that roads approximate a level plane, the rest of an urban environment is often closer to a series of blocks of varying heights or depths with shadows cast by adjacent buildings. Building entrances and exits are hidden due to the birds-eye view of a satellite image, with little to no information about a building's exterior. An example of this problem is shown in 1-1. This imagery is often outdated or relevant only to the season in which the image was taken. Combined, these flaws often give soldiers a false mental model of their environment. Many of these flaws could be addressed by having personnel on the ground use a robot to explore and map their environment. Given the need of these personnel to simultaneously perform a primary task such as finding a Person of Interest (POI), an autonomous robot would allow these groups to better perform ISR and improve their Situational Awareness (SA) in real-time. However, performing an ISR mission aided by an autonomous vehicle will require an interface which allows the user to easily transition between high-level control of the robot (e.g. moving via waypoints) and a low-level, fine-grained control to align the robot for obtaining the best view.

Recent advances in several fields have led to a new type of unmanned autonomous vehicles known as Micro Aerial Vehicles (MAVs). Given their compact size, low cost, and flight capabilities, MAVs are primarily marketed and designed for ISR-type missions. In a conventional ISR mission, people interviewed for this effort envisioned that soldiers and law-enforcement personnel would use a MAV to improve their SA of an urban environment or building. MAVs are also ideal for Unmanned Search and Rescue (USAR) missions, where they can easily traverse rubble and other obstacles which would normally prove challenging for ground-based USAR robots. Finally, MAVs can be used in a variety of

civilian applications, such as structural inspections or environmental/farming surveys.

1.1 Micro Aerial Vehicles

MAVs were first investigated in a Defense Advanced Research Projects Agency (DARPA) program in the late 1990s which examined the feasibility of creating small aircraft less than 6 inches (in) diameter [1]. Although the study was successful and several MAVs were subsequently created according to the original DARPA specifications (i.e., TU's Delfy [2], AeroEnvironment MicroBat, etc), the commercial sector has largely pursued MAV helicopters. Simultaneously, the Army has pursued their own class of MAVs, known as Class I Unmanned Aerial System (UAS), which are larger and designated for platoon-level support [3]. Class I UASs are required to weigh less than 51 pounds (lbs) (including the Ground Control Station (GCS)) and should fit into two custom Modular Lightweight Load-carrying Equipment (MOLLE) containers, which are approximately the size of a large backpack. Currently, the only deployed Class I system is the Honeywell RQ-16 "T-Hawk," seen in Fig. 1-2.

For the purposes of this thesis, the term MAV will be used to refer to the commercial sector MAVs that are available or in development. These helicopters may have two, four, or six rotors, are typically less than two feet across, and can carry payloads of up to a kilogram. Several off-the-shelf MAVs are available from a variety of companies, and two such examples are shown in Fig. 1-3.



Figure 1-2: A Honeywell RQ-16 T-Hawk in flight, courtesy of U.S. Navy.

These MAVs all share a common set of features which are critical to their proposed use case for ISR missions. First and foremost, these vehicles are capable of Vertical Take-Off and Landing (VTOL), which

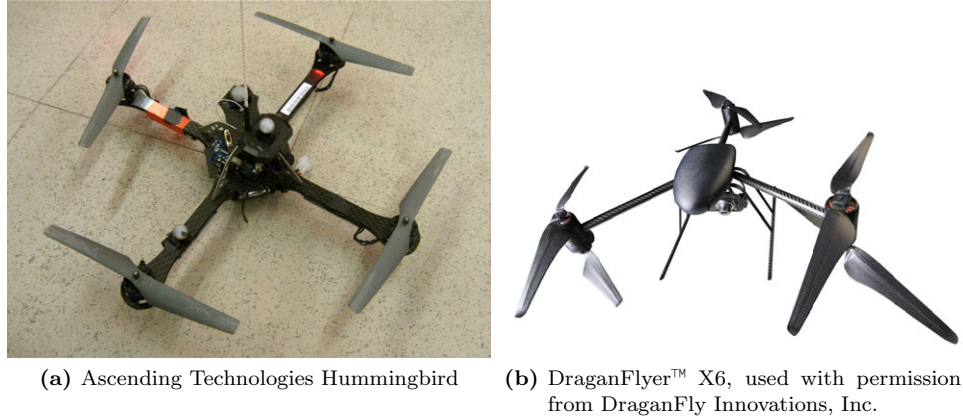


Figure 1-3: Examples of commercial MAVs available on the market.

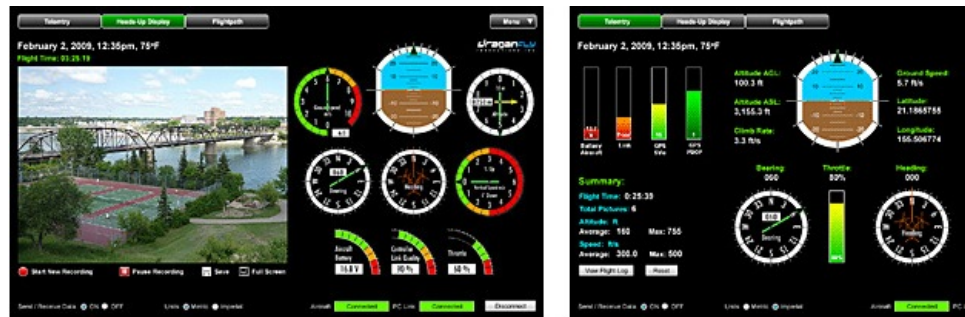
allows them to be launched and recovered in confined spaces or urban environments which may not have the physical space to allow for a traditional takeoff/landing. Complementing their VTOL capability, these MAVs are able to precisely hover and move to a fixed point in space. This allows them to easily survey from a fixed vantage point, without the need to make repeated passes of an Area of Interest (AOI), a capability referred to as perch and stare. To support these capabilities, MAVs range from semi- to fully-autonomous. Even the most basic MAVs have complex flight dynamics which require a low level of automation to maintain vehicle stability in-flight. More advanced MAVs are fully autonomous and capable of flying a route of Global Positioning System (GPS) waypoints with no human intervention [4].

1.1.1 MAV Operators

Given their short flight time, MAVs are operated by field personnel located in the vicinity of the MAV. These field personnel may be emergency first responders, police, specialists (e.g., building inspectors or bomb technicians) or, most commonly, dismounted, forward-deployed soldiers. All of these groups operate in hazardous environments which may contain hostile, armed people, unstable structures, or environmental disasters. Although these personnel may operate a MAV, it is never their primary task. Rescue personnel are concerned with finding and saving victims, and soldiers may be on a patrol or searching for POIs. Operating the MAV is not an independent goal of these personnel. Instead, the MAV provides a means of effectively achieving their primary objectives. However, this additional aid comes at the cost of dividing the operator's attention and possibly diminishing his or her SA. The problem of divided attention currently makes MAVs effectively unusable by personnel who already have demanding tasks they cannot afford to ignore.



(a) Ascending Technologies AutoPilot Software, used with permission of Ascending Technologies, GmbH.



(b) DraganFlyer's ground control software, used with permission of DraganFly Innovations, Inc.

Figure 1-4: Examples of software interfaces for commercial MAVs.

1.1.2 MAV GCSs

Currently, MAVs are controlled via computer interfaces, as seen in Fig. 1-4. Some MAVs are controlled with more specialized GCSs. Typically a ruggedized laptop display, GCSs may incorporate specialized controls such as miniature joysticks or pen styli (Fig. 1-5) and range from a hand-held device to a large briefcase in size. Many of these GCSs take several minutes to assemble and establish a connection with the MAV every time they are used. In addition, the bulk of a GCS adds significant weight to the pack of an already fully-loaded soldier. Some interfaces, such as the AeroVironment's MAV GCS (Fig. 1-6) and DraganFlyer, offer goggles which allow the operator to view a video feed from the MAV. As seen in Figure 1-7, all interfaces require an operator to use both hands to interact with the MAV.



(a) Honeywell's RQ-16 Ground Control Station, which is similar to a rudimentary tablet PC, courtesy of Honeywell.



(b) DraganFly's X6 manual flight controller, used with permission of DraganFly Innovations, Inc.

Figure 1-5: Examples of MAV Ground Control Stations.

A majority of interfaces and GCSs require the full attention of the operator. These systems require extensive training before an operator can safely and effectively operate the MAV. GCSs which allow the operator to manually position and orient the MAV rely on a classical 1st order feedback control loop, which allows operators to directly control the thrust, pitch, and roll/yaw of the MAV. This complex

feedback loop demands full attention of the operator.

All current reconnaissance MAV interfaces are rooted in the constrained case that the operator's primary task is to operate the MAV, which includes both flying the vehicle and searching images from the vehicle concerning targets of interest. These design choices appear to be the extension of larger UAV ground stations (e.g, the Predator GCS). Other design choices have confusing rationale when considering the needs and divided attention of a field operator in a hostile environment. For example, video goggles (Fig. 1-6) blind operators to their surroundings, depriving them of critical perceptual cues, both foveal and peripheral. As a consequence, current GCSs and interfaces have a number of design decisions which preclude them from being used effectively by field operators, who almost universally have other, more urgent primary tasks to accomplish.

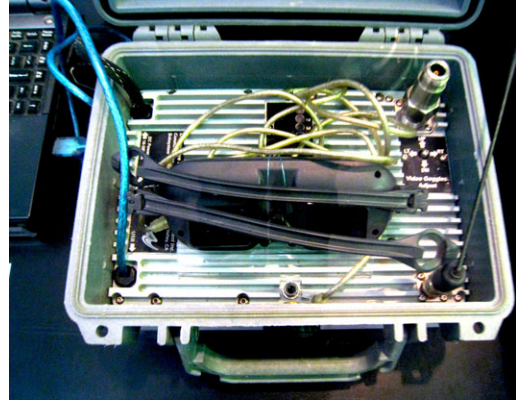


Figure 1-6: Video goggles for AeroVironment's GCS, photo by author.



Figure 1-7: An Army operator controlling a RQ-16 with a stylus, courtesy of U.S. Army.

While these types of interfaces have been successfully used in many conventional UAV interfaces, they ignore the unique capabilities afforded by a MAV. Given their short flight time, MAVs on ISR missions are best suited to act in collaboration with personnel on the ground in the same area the MAV is surveying. From a human-centered view, MAVs performing local ISR missions could report directly to personnel in the field, such as soldiers, police, or first-responders, and even collaborate together to discover an unexplored environment. Creating a high-level interface on a truly mobile device will mitigate many of the existing flaws in present-day MAV interfaces. This interface must appropriately balance the need to support intermittent interaction from a user and having safe, intuitive flight controls when the user needs fine-grained control over the MAV's position and orientation (such as peering

in a window during an ISR mission). Additionally, existing interfaces ignore the context of the situation in which they are used. With soldiers already carrying packs weighing 90 lb or more [5], the additional weight and size of any new equipment is an extreme burden. Likewise, in these dangerous environments, interaction methods which introduce more specialized equipment, such as a stylus or video goggles, make the equipment more prone to failure or being lost. Combined, these considerations clearly necessitate

a compact, simple device if soldiers and other operators are realistically expected to use a MAV in the field.

1.2 Mobile Devices

Like MAVs, handheld mobile devices have recently emerged as a smaller counterpart to traditional laptop/desktop computers. Although the concept has existed since the mid-nineties in the form of Personal Digital Assistants (PDAs), these devices previously had a minimum amount of computing power with poor displays, and were viewed as a digital extension of a notepad or contact list. The PDA industry lost ground after 2000, as mobile phones became more powerful and incorporated new features such as cameras and basic internet access. Only in the last few years have mobile devices finally achieved a state where they are able to run powerful applications and support meaningful interactions without the traditional keyboard and mouse. Known as smart phones, several platforms have emerged for these mobile devices, most notably the Apple iPhone[®] and Google Android[™]. Both of these platforms are based on a touch-screen, internet-enabled device which is aware of its location and can sense the device's tilt and orientation through on-board sensors. Although not specifically stated, these devices also have screen resolutions of 150-225 dots per inch (dpi), which allows them to display detailed imagery in a compact physical format compared to normal computer displays. Given their portability and user-centered focus, handheld devices provide a possible platform for a new type of MAV GCS.

1.3 Problem Statement

As robots become more integrated into ISR missions, effective interaction between humans and robotics will become more critical to the success of the mission and the safety of both the operator and robot. Previously, Human-Robot Interaction (HRI) has focused largely on working with robots through teleoperation, or as independent agents in their environment. Teleoperation interactions ignore the problem of the divided attention of the field personnel, the importance of environment, and the issue that their primary goal is to perform ISR, not to drive a Unmanned Vehicle (UV). For successful collaborative exploration of an unknown environment, an interface must first be developed which allows a user to work with an autonomous robot operating within the same environment. This interface should not restrict the operator from their primary task; it must be a truly mobile device and not require the operator's continual attention. This necessitates an interface that can allow an operator to easily control a robot at a high-level supervisory mode of interaction for general commands, as well as a fine-grained, lower level of control when more nuanced actions are required. A mobile device which successfully allows operators to supervise and occasionally directly operate a robot will dramatically change the way robots are used in high-risk environments.

1.4 Research Objectives

In order to address the problem statement, the primary goal of this research is to develop a mobile interface for interacting with an autonomous MAV. This goal is addressed through the following research objectives:

- Objective 1. **Determine the function and information requirements for a MAV interface.** To achieve this objective, a Cognitive Task Analysis (CTA) was performed with personnel likely to benefit from using a MAV to explore an unknown environment. These personnel were interviewed to identify needs and usage scenarios, as described in Chapter 3. Current practices for operating MAVs and designing relevant mobile interfaces were researched in support of this objective.
- Objective 2. **Develop a mobile interface which allows an operator to explore an unknown environment with a MAV.** Based upon the research performed for Objective 1, a mobile interface for a MAV was designed and implemented (Chapter 3).
- Objective 3. **Evaluate the usability of the interface.** An experiment with human participants was conducted (Chapters 4 and 5) to determine how well the interface supported an operator in exploring an unknown environment.

1.5 Thesis Organization

This thesis is organized into the following chapters:

- Chapter 1, *Introduction*, provides an overview of the research, the motivations, and the objectives of this thesis.
- Chapter 2, *Background*, examines related work in HRI and hand-held device communities.
- Chapter 3, *Interface Design*, describes the rationale and formulation of the application designed to control a MAV in an outdoor environment.
- Chapter 4, *Usability Evaluation*, describes the design of the usability study conducted involving the interface.
- Chapter 5, *Results*, analyzes data gathered from the Usability Evaluation and discusses important relationships across the data.
- Chapter 6, *Conclusions and Future Work*, compares the results obtained with the hypotheses postulated, discusses how well the research objectives were met, and proposes directions in which further research could be conducted.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 2

Background

2.1 Human-Robot Interaction

2.1.1 Human Supervisory Control

MAV interfaces embody a form of Human Supervisory Control (HSC), which as depicted in Fig. 2-1, is when a human supervisor executes control of a complex system by acting through an intermediate agent, such as a computer. This interaction is performed on an intermittent basis, which may be periodic or in response to changing conditions of the system. While engaged in supervisory control, a human will develop a plan with the assistance of the agent, then instruct the agent to perform the plan. As the plan is executed, the human supervises the agent and intervenes as mistakes are made, events change, or the agent requires assistance, then learns from the experience to improve future plans [6]. In the case of controlling a MAV using HSC, an operator creates a set of waypoints for a MAV to visit, evaluates the path planned by the MAV to visit the waypoints, and then changes waypoints. The intermediate

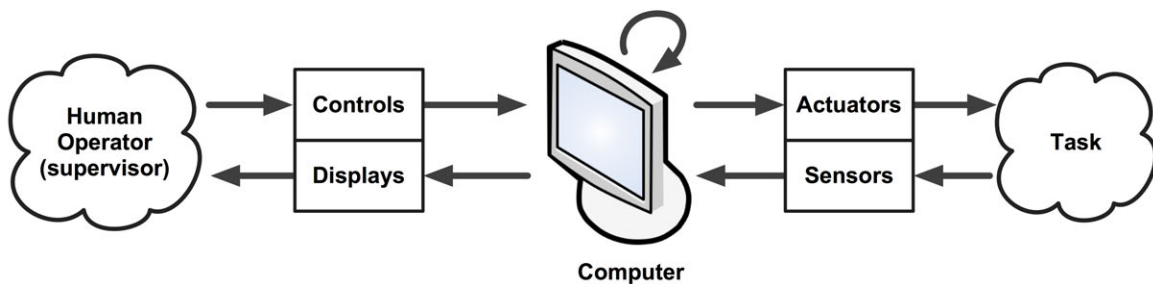


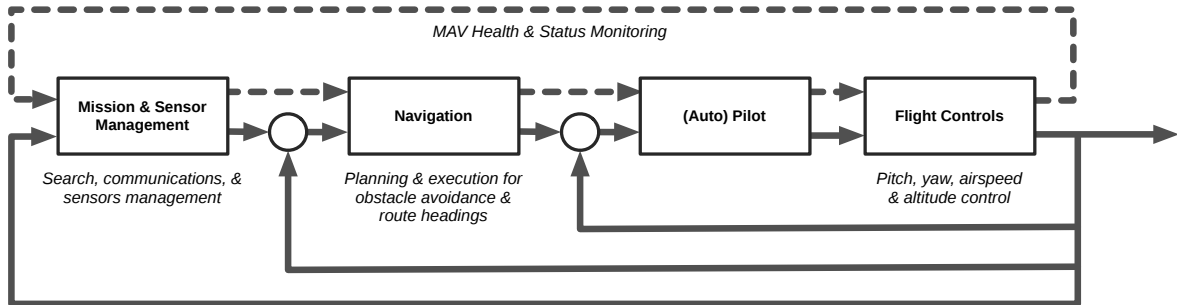
Figure 2-1: General Human Supervisory Control, adapted from Sheridan [7].

agent between the human and system may have some Level of Automation (LOA) to aid in managing

Table 2.1: Sheridan and Verplank Levels of Automation [8]

Automation Level	Automation Description
1	The computer offers no assistance: human must take all decision and actions.
2	The computer offers a complete set of decision/action alternatives, or
3	narrows the selection down to a few, or
4	suggests one alternative, and
5	executes that suggestion if the human approves, or
6	allows the human a restricted time to veto before automatic execution, or
7	executes automatically, then necessarily informs humans, and
8	informs the human only if asked, or
9	informs the human only if it, the computer, decides to.
10	The computer decides everything and acts autonomously, ignoring the human.

the complexity of the system and help reduce cognitive workload. These levels can range from low, where recommendations are provided by the automated agent and the human retains complete control of decision-making, to the highest levels, where an agent may independently make decisions without informing the human. The range of LOA (Table 2.1) were originally proposed by Sheridan and Verplank [8]. One might expect the human’s cognitive workload to decrease as the level of automation increases. However, increasing automation beyond what is needed may lead to loss of SA, complacency, and skill degradation.[9, 10] As Cummings et al. show in Fig. 2-2, HSC of a UAV relies upon a set of hierarchical control loops [11]. However, accomplishing higher level tasks, such as ISR, depends directly upon shifting some of the lower-level tasks, like piloting the vehicle and navigation, to automation. If an operator is required to manually perform the inner control loops, his attention becomes divided between the original task and lower level functions. Introducing automation into these inner control loops allows an operator to effectively execute HSC and devote most of their attention to their primary task of mission and payload management.

**Figure 2-2:** Hierarchical Control Loops for a Single MAV performing an ISR mission, adapted from Cummings et al. [11]

2.1.2 Human-Robot Interaction Roles

HRI is a subset of human supervisory control, which generally focuses on mediating interaction between a human and robot. In 2003, Scholtz defined various roles which a human interacting with a robot could fulfill (Table 2.2), such as Supervisor, Operator or Mechanic [12]. Although these roles have existed in the Human Supervisory Control community for some time, Scholz postulated the specific HRI needs and requirements (Table 2.2) that would be necessary for humans in these roles to perform their tasks, while still maintaining adequate SA.

Table 2.2: Scholtz’s HRI roles, with Goodrich and Schulz’s additions, adapted from [12, 13]

Role	Description
Supervisor	monitoring and controlling the situation
Operator	modify internal software or models when the robot’s behavior is not acceptable
Mechanic	handles physical interventions and modifying the robot’s hardware
Peer	teammates of the robot can give commands with larger goals/intentions
Bystander	may influence the robot’s actions by their own actions, but has no direct control
Mentor	the robot is in a teaching or leadership role for the human
Information Consumer	the human does not control the robot, but the human uses information coming from the robot in, for example, a reconnaissance task.

Most recently, Goodrich and Schulz performed a survey of HRI research efforts [13]. They present a case for classifying interaction as remote (where the human is physically separated or distant from the robot) and proximate (where the human is near or within the same environment as the robot). Providing this distinction allows for Goodrich and Schulz to more precisely identify which interaction techniques are best for each role. They also expand on Scholtz’s roles, extending the list by including Mentor and Information Consumer (Table 2.2). Relevant to this effort, the Information Consumer is defined as a human who seeks to use information provided by an autonomous robot, such as in an ISR mission. A soldier operating a MAV will alternate between acting as a Supervisor, as he develops plans for the MAV’s exploration, and Information Consumer when he is examining imagery returned by the MAV. Eventually, highly autonomous MAVs may support the role of a field operator as a Peer.

These broader research efforts in the HRI community provide a concept and framework for how to

view specific efforts and endeavors when creating a hand-held interface for controlling a small autonomous vehicle. By and large, all of the existing HRI research in the supervisor/operator/information consumer domains has tried to frame high-level HRI as Human-Computer Interaction (HCI) which as a by-product, manipulate robots operating in the real world.

2.1.3 Urban Search and Rescue (USAR)

One HRI domain that has received significant attention for human operators as Information Consumers and Supervisors has been the USAR field. The field of USAR seeks to develop robots and interfaces which allow first responders to remotely explore dangerous environments. Numerous studies have been conducted on the effectiveness of USAR interfaces (e.g., [14, 15]). Researchers have also studied USAR competitions to better understand why many of the teams failed to successfully finish the competitions [16, 17]. While these studies concentrated on using Unmanned Ground Vehicles (UGVs) to explore the environment, the problems in conducting unmanned USAR and MAV ISR missions are very similar. USAR teams must navigate unfamiliar environments with little pre-existing intelligence and dynamically build SA through the sensors available on the robot. In addition, navigation is particularly challenging in USAR due to the challenges of the terrain (e.g. rubble, low-visibility, confined spaces). Both Scholtz et al. [18] and Yanco and Drury [17] found that many teams in USAR competitions were hindered by their ability to effectively navigate their robots using interfaces which they had designed and extensively trained with in order to operate the vehicles. Operators were plagued by problems involving poor SA and fundamental usability issues, such as window occlusion and requiring operators to memorize key bindings [17]. In all cases, the teams were not constrained in how they could design their interface, and a wide variety of interfaces were observed, from multiple displays to joysticks to Graphical User Interfaces (GUIs) and command-line inputs. Both studies recommended USAR researchers focus on reducing the cognitive workload of the operator. Often, it was noted that the interface was designed first for developing the robot, with controlling it during a mission as an afterthought.

Recently, Micire et al. [19] created a USAR multitouch table-top interface. Multitouch gestures control camera orientation, and single finger gestures control the speed and direction of the robot. The touch gesture for controlling the speed and direction of the robot were directly mapped from a conventional joystick control into GUI with directional buttons. In a usability evaluation, Micire et al. found that users fell into two distinct groups when controlling the speed and direction of the robot. One group of users assumed the magnitude of their gestures was mapped to the speed the robot would move. The other group treated the magnitude of the gesture as unimportant, with their touch simply controlling the direction of the robot at a constant speed. No information about the users' performance was detailed in this study.

2.2 Research Relevant to MAV Interaction

The development of MAVs has occurred so recently that there is little published research examining how humans can best interact with them. Although commercial systems such as Ascending Technology’s quad rotor helicopters and DraganFly’s DraganFlyer are available (Fig. 1-3), these use proprietary interfaces and no information is available about their development. In the context of using MAVs in outdoor environments, two main areas are of interest for any MAV operated by a dismounted soldier or operator: level of control and the need for mobility.

The term teleoperation was first introduced by Sheridan in his work on levels of automation and human supervisory control [7]. Teleoperation refers to the concept of a human operator controlling a robot (or autonomous vehicle) without being present. Teleoperation is often performed via manual control (i.e., increase forward velocity by 1 m/s) through the use of a joystick or other interface which requires the constant attention of the operator. This drastically increases the cognitive workload of the operators, and in turn leaves less time for them to perform other tasks. As such, teleoperation is viewed as a difficult problem, especially when compounded with the practical constraints encountered in practice (i.e. time delays in communications, low bandwidth for information). After introducing the concept, Sheridan immediately followed with an exhaustive study of teleoperation which demonstrated how time delays between the robot and operator have a detrimental effect on performance [20]. If an operator is required to manually handle the inner loops of Fig. 2-2 using teleoperation, he or she will have less time and cognitive effort to spend on mission and sensor management, which is the primary task.

A large body of literature exists on teleoperation. Chen et al. distilled existing research into a set of constraints which were common to many teleoperation interactions (i.e., Field of View (FOV), orientation & attitude of the robot, frame rate, and time delays) [21]. Many of these constraints are still relevant to the case of an autonomous MAV which is delivering live imagery to the operator. Fong et al. proposed teleoperation with a semi-autonomous robot, which may reinterpret or ignore teleoperation commands from the operator [22]. While Fong et al.’s research is presented in the larger context of a human and robot having an intelligent dialogue, it is worth noting for the idea of introducing a full layer of autonomy between an operator’s teleoperation commands and what actions are executed by the robot.

Jones et al. performed a study examining an operator’s ability to teleoperate a robot through a rectangular opening in a wall [23]. They showed operators were not apt in judging their own ability to manoeuvre a robot through the opening and consequently performed poorly at driving the robot through openings. Surprisingly, operators were accurate at judging the dimensions of the opening in relation to the robot. This led the researchers to conclude that operators would require some form of assistance teleoperating a robot in confined spaces. Though this research only examined a two dimensional opening (a hole in a wall), it is even more applicable to the problem of teleoperating a MAV in three dimensions

in an urban setting. Although there has been little research on teleoperating MAVs, past research in teleoperating robots, large UAVs, and spacecraft indicates that controlling a MAV in three dimensions will be a difficult problem. Operators will have additional cognitive workload from the additional dimension, and the interface will still be subject to the constraints identified by Sheridan [7] and Chen et al. [21], such as time delay, frame rate, sensor FOV, and orientation of the MAV.

Several researchers [24, 25, 26, 27, 28] have investigated using an interface to control a robot from a hand-held device. However, no interface has been yet implemented which makes full use of the potential of the new class of hand-held devices that have emerged in recent years. Many of these interfaces simply use classical What-You-See-Is-What-You-Get (WYSIWYG) controls and widgets (i.e., sliders, buttons, scroll bars) with little regard that they are implemented on a hand-held device, which has a significantly different interaction paradigm from a computer desktop. While at a computer desktop, a user is often focused exclusively on interacting with an application using an information-rich display and a traditional GUI with a keyboard and mouse. In contrast, hand-held devices assume infrequent interaction with the user and display relatively little information, requiring an imprecise pointing device (e.g., a stylus or finger). None of the interfaces in these studies involved higher levels of human supervisory control, and instead required continuous attention from the operator to operate the robot. All of these interfaces followed a similar pattern of having separate imagery, teleoperation, and sensor displays. Many used four buttons along the cardinal directions for teleoperation. Few of these interfaces were evaluated with quantitative user studies, making it difficult to identify specific interaction issues which could be improved. Adams et al.'s interface [26] used a PDA to control a ground-based robot, displaying a camera image from the robot with overlaid cardinal direction buttons (i.e., forward, backward, left, right) for teleoperation. Their user study found that interfaces which incorporated sensor displays (either by themselves or overlaid on top of a video display) induced higher workload for users who were unable to directly view the robot or environment. However, they also found sensor-only displays resulted in a lower workload when the participant could directly view the robot or environment [26].

Using a multitouch hand-held device with a high-fidelity display for HRI, such as an iPod Touch[®], has been designed by Gutierrez and Craighead, and O'Brien et al., although neither group conducted user studies [27, 28]. O'Brien et al. implemented a multi-touch interface with thumb joysticks for teleoperation of a UGV. However, they note that these controls are small and difficult to use, with the additional problem of the user's thumbs covering the display during operation. Both of these interfaces are for the ground-based PackBot[®] and do not accommodate for changes in altitude. In a primitive form of interaction for UAV control, a mission-planning interface designed by Hedrick et al. (<http://www.youtube.com/watch?v=CRcld5aAN2E>) for an iPhone was simply a set of webpages which required the user to input detailed latitude/longitude coordinates using the on-screen keyboard.

Very little research exists specifically on interaction with MAVs. Durlach et al. completed a study

in 2008 which examined training MAV operators to perform ISR missions in a simulated environment [29]. Operators were taught to fly the simulated Honeywell RQ-16 MAV with either a mouse or game controller. Although Durlach et al. state that they limited the simulated MAV to a maximum velocity of six kilometers/second (km/s), the vehicle was fly-by-wire, with stabilized yaw/pitch/roll axes to maintain balanced flight, which participants could only crash by colliding with other objects in the simulation. Durlach et al. do not mention if their simulation incorporated video/communication delay. The study specifically looked at whether discrete or continuous input teleoperation controls yielded better performance using the two interfaces shown in Fig. 2-3a and Fig. 2-3b. To test these displays and controls, Durlach et al. trained and tested 72 participants. During these flights, the operators manually flew the helicopter, with no higher-level automation such as waypoint guidance. For training, participants flew seven practice missions navigating slalom and oblong race tracks and were allowed five attempts per mission. No information was provided on why participants needed seven practice missions and five attempts per mission. If the participants successfully completed the practice missions, they were given two ISR missions to perform (with additional practice missions in between the two ISR missions). Both missions involved identifying POIs and Objects of Interest (OOIs) in a simulated outdoor urban environment. The MAV was oriented to take reconnaissance photos of the POIs/OOIs with the fixed cameras. Twenty four participants were excluded from the first mission’s post-hoc analysis by the researchers due to their inability to identify all POIs.

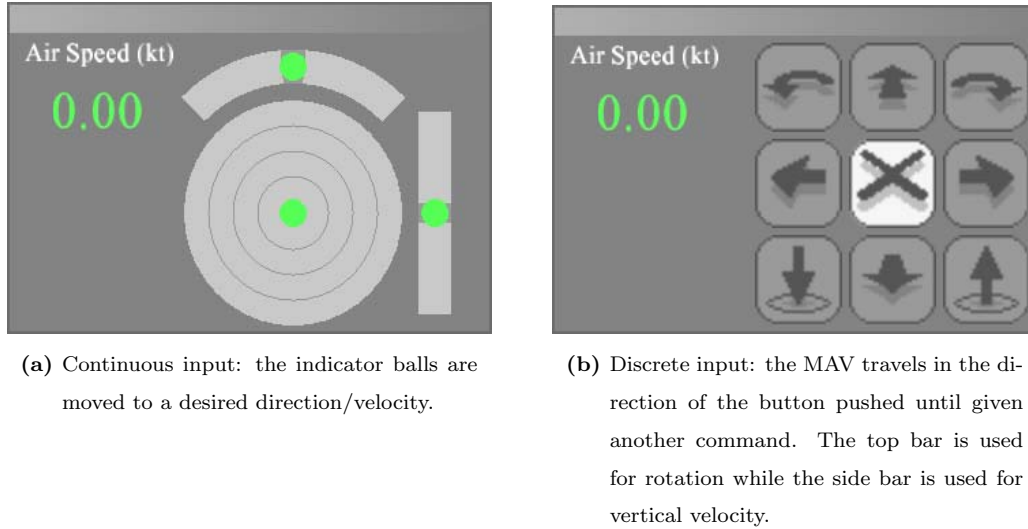


Figure 2-3: MAV control interfaces created by Durlach et al. [29]

By the end of the experiment, each participant received approximately two hours of training in addition to the primary missions. The first primary mission had no time limit, while the second had a seven minute time limit. While there were significant interaction effects between the controller and input methods (discrete vs. continuous) in some circumstances, participants using a game controller with a

continuous input teleoperation control performed statistically significantly better overall. Durlach et al. also identified a common strategy of participants using gross control movements to approach a target, then hovering and switching to fine-grained teleoperations controls to obtain the necessary ISR imagery. With both of these interfaces, over half of the participants collided with an obstacle at least once during the primary ISR missions. Contrary to one of the study’s original hypotheses, participants performed worse when given a dual-camera view instead of a single-camera view during the mission. The relevance of Durlach et al.’s results is limited because their controls and displays are simulated, with no lag and delay, which is inherent in real-world interactions. As shown by Sheridan, a delay and lag over 0.5 second (sec) within a teleoperation interface significantly affects the operator’s performance [20] , so these results are at best preliminary.

2.3 Research in Hand-held Devices

Hand-held devices present many additional interface challenges in addition to standard Human-Computer Interaction (HCI) concerns. Given their small form factor, display screen sizes are often very limited (typically 300-400 pixels wide). Keyboards are often not included, and input is via a stylus or touch-based interface. Hinckley and Sinclair significantly advanced the quality of interaction with their invention of capacitive touch devices [30], which is the technology behind most touch-enabled devices today. Previously, users had to interact with the interface via a stylus, which, as Adams et al. noted from interviews, essentially excludes the interfaces from being operational in a military domain because the stylus would likely be lost [26]. Pascoe et al. performed the first study of using a PDA for fieldwork, with users surveying animals in Africa [31]. Following this study, they proposed the idea of a Minimal Attention User Interface (MAUI), which emphasized high-speed interaction and supporting users with limited time to attend to the PDA.

Tilt-based hand held interfaces were first invented by Rekimoto in 1996, who provided an example application of map navigation via tilting the device (using accelerometer sensors) [32]. The idea was largely ignored until Jang and Park implemented a tilt-based interface with low- and high-pass filtering of the raw tilt data to generate a clean signal [33]. Recently, Rahman et al. performed an ergonomic study to determine what fidelity a user had in tilting a hand-held device, and how to best discretize the tilting motion [34]. Tilt-based interfaces offer an intuitive interaction for many people. However, while the general gestures may be easy to comprehend, there are many difficulties associated with interpreting the “resting” pose for the tilt gesture, and the natural range of the gesture may vary from person to person [35].

In large part, the push for innovation in hand-held devices has been driven by several key companies. Palm™ is famous for introducing the first popular mass-consumer hand-held device and the graffiti al-

phabet. Microsoft has steadily driven the development of hand-held computers, although these typically represent a smaller form-factor computer with keyboard and stylus rather than a true hand-held device. Apple’s introduction of the iPhone/iPod Touch has spurred a new wave of development in hand-held devices. Some of Apple’s notable contributions include multi-touch functionality, high fidelity displays, and incorporating accelerometers. A variety of companies have since followed in Apple’s steps and built their own devices with similar functionality [36].

A common trend seen in many hand-held interfaces designed for controlling robots is repackaging a traditional computer interface into a hand-held format. Very few research efforts examine using additional or different modalities such as multi-touch or tilt-based interaction to collaborate with a robot. Research on MAV interfaces is even more scarce, likely due to the fact that MAVs have largely been developed by companies with proprietary research.

2.4 Summary

There are several research gaps in HRI research regarding human interaction and collaboration with sophisticated autonomous robots, like MAVs, particularly for users who require control via mobile devices. Currently, MAV operator interfaces do not consider the real-world needs of field personnel. They require a laptop or other additional hardware and assume the operator’s primary task is controlling the MAV. Likewise, USAR researchers have long been occupied with interfaces which allow them simply to teleoperate a robot, rather than collaboratively explore the environment. Although some of these USAR interfaces have been developed for hand-held devices, they simply repackage the typical USAR interface into a smaller format, ignoring the need to support divided attention. Researchers who have taken advantage of more sophisticated hand-held devices with accelerometer or multitouch capabilities have also simply followed the same WYSIWYG paradigm. Finally, no HRI research exists on interacting with an actual MAV in a real-world setting.

MAVs are well-suited for performing many types of tasks for personnel in an unfamiliar or dangerous environment. To best fulfill this role, however, a MAV must truly act as a supporting agent collaborating with the user, rather than being closely supervised or operated. This demands an interface and underlying automation architecture, which allows users to be focused on completing their primary task and intermittently attending to their interaction with a MAV. Designing an interface for this role requires a departure from the previous avenues which the HRI community has developed for robot and MAV interfaces.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 3

Interface Design

3.1 Introduction

This chapter presents the design of a hand-held device application, Micro Aerial Vehicle Exploration of an Unknown Environment (MAV-VUE), for collaboratively exploring an unknown environment with a MAV. This application was the result of a Cognitive Task Analysis (CTA) which examined how a MAV could be used to help field personnel in an outdoor environment. The interface and displays are outlined along with a discussion of the theory and rationale behind the design.

3.2 Cognitive Task Analysis

A CTA was performed to gain a better understanding of how potential field operators would use hand-held devices to operate a MAV during an ISR mission. For the purposes of this thesis, CTA is defined by Chipman, Schraagen and Shalin as “the extension of traditional task analysis techniques to yield information about the knowledge, thought processes, and goal structures that underly observable task performance.”[37, p. 3] Three potential users of a MAV ISR system were interviewed in person or via phone. Each semi-structured interview consisted of open-ended questions (Appendix A) and lasted approximately forty-five minutes to one hour. These personnel individually had combat experience in Iraq, Afghanistan, or police Special Weapons and Tactics (SWAT). Combined, they represented a wide range of potential field operators, with over 40 years of shared experience. These personnel are typically required to operate in places with little advance knowledge of the situation or environment. This problem is compounded in urban areas where intelligence is often outdated or not detailed enough to give these personnel the SA they need to properly perform their primary task. In interviews with these personnel, it was often mentioned that they typically had a poor understanding of the operational environment.

Even when supplemented with satellite imagery or floor plans (which are rarely available), they found the intelligence to be unreliable due to a variety of factors relating to a changing urban environment, shadows in satellite imagery, and general lack of information which prevented them from building an accurate mental model (in three dimensions) of their environment. As an example, one soldier showed a satellite map used for patrol missions in Iraq and identified several areas where garages look like shadows from an adjacent building, or an exposed pit which appeared as a rooftop. This information was corroborated by civilian police operating on SWAT missions who often found building plans to be unreliable, with furniture and unmarked renovations significantly changing the interior space. Building a mental model of the space is so crucial that SWAT teams will typically first perform a mock run of their mission on an adjacent floor in the building to properly construct their mental model of the environment.

Soldiers, SWAT police teams, and other field personnel must make critical decisions given extremely limited information about their environment. Allowing a MAV to explore these environments will help these groups, but only if it will not add to their existing high workload. This thesis focuses on the design of an interface for controlling a MAV used for outdoor ISR missions which will meet these needs.

3.3 MAV Interface Requirements

3.3.1 Desired MAV Roles & Requirements

Based upon the results of the CTA, a potential set of roles was identified that a MAV could fulfill in assisting field personnel in outdoor settings (Table 3.1). It is expected that the MAV operator is engaged in a primary task in the environment (such as search and rescue) and only intermittently interacting with the MAV to receive status updates or formulate new plans.

Table 3.1: Potential roles for an operator and a MAV collaboratively exploring an outdoor environment.

Operator Role	MAV Role
Attending to MAV as secondary task	Autonomously exploring the environment
Controlling MAV as a primary task	Limited autonomy, exploration primarily under the control of the operator
Consuming information from MAV	Delivering sensor information about the environment

However, there may be points during the mission when the operator would need to take a more active role and teleoperate the MAV to explore in more detail, such as obtaining a particular view of the environment. Finally, at other times the operator may not be actively interacting, but fully focused on

consuming information delivered by the MAV's sensors. In discussing various roles a MAV could fulfill, interviewees focused almost exclusively on the ISR capabilities. Corresponding to this focus, personnel expressed little interest in weaponizing a MAV or using a MAV in a payload capacity beyond the ability to carry sensors. All personnel interviewed expressed a desire to have a MAV which weighed less than five lbs (including the GCS) and a flight time of 2-3 hours, but also felt twenty minutes of flight time was the minimum requirement for a MAV to be useful.

Table 3.2: Potential MAV outdoor ISR missions

Name	Description
Building Surveillance	The MAV scouts the exterior of a building, identifying relevant entrances, exits and other features of interest.
Environment Modeling	The MAV constructs a Three Dimensional (3D) model of the local environment.
Perch and Stare	The MAV flies to a vantage point with good visibility and lands, passively observing the environment in view.
Identification	The MAV is navigated through an urban environment, and imagery is used to identify Persons of Interest (POIs) or Objects of Interest (OOIs).
NBC Detection	Similar to the Perch and Stare role, the MAV would either be used to sweep or passively monitor an environment for Nuclear, Biological and Chemical (NBC) agents.
Rearguard	The MAV follows a dismounted soldier and monitors the environment for any enemies.

The ISR nature of these functions yields a proposed set of system requirements for a MAV, listed in Table 3.3. Although some of these requirements are not currently feasible, they are all technically possible given future development. Given the roles and potential missions, there is a clear need for an interface which can allow an operator to interact with a MAV in hostile, unprepared settings that range from limited attention to focused task manipulation.

Table 3.3: System requirements for a MAV performing an outdoor ISR mission.

System	Requirement
Take-off and Landing	VTOL capabilities are required to function in an urban environment
Weight & Size	The MAV and operator's interface must weigh less than 5 lb and fit within a pack
Flight Time	20 minutes minimum, 2-3 hours desired
Payload	Sufficient for sensor packages
Sensors	Video, NBC detectors, microphone, electro-optical, Light Detection and Ranging (LIDAR) for 3D dimensional mapping, alone or in combination
Maneuverability	Full 6 Degree of Freedom (DOF) movement
Communication	Ability to communicate with MAV a majority of the time. Communication loss during certain tasks (e.g. moving between waypoints) is acceptable
Position Awareness	Accuracy to within one meter is desired, but accuracy within 3 meters may be acceptable

3.3.2 Interface Requirements

Following the CTA and generation of the roles, mission, and system requirements of a MAV, functional (Table 3.4) and information (Table 3.5) requirements were generated for a HRI interface that would allow field personnel to use a MAV to collaboratively explore an unknown environment. These functional requirements define actions and system behaviors an operator needs in order to effectively interact with the MAV. Similarly, information requirements represent knowledge needed to aid the operator's cognitive processes, such as constructing a mental model, predicting future situations, and projecting the consequences of any plans or commands.

Table 3.4: Functional requirements for a HRI interface to support collaborative exploration of an unknown environment.

Function	Subfunctions
Semi-Autonomous Navigation	• Add, Edit and Delete Waypoints • Specify waypoint location, altitude and heading • Specify order of waypoints • Clear all waypoints • Prevent collisions with obstacles • Perform VTOL • Loiter
Fine-Tune Control	• Manually change the position, altitude or orientation • Engage/disengage manual flight controls • Prevent collisions with obstacles
Identification of POI and/or OOI	• View/hide sensor display • Rotate, pan and zoom sensors • Save information for review later
Communication	• Disconnect from MAV • Connect to MAV prior to flight • Reconnect with MAV mid-flight
Health & Status	• Disable/enable sensors

Table 3.5: Information requirements to support collaborative exploration of an unknown environment.

Function	Information
Autonomous Navigation	• Waypoints • Display location, orientation and altitude of waypoints • Display order of waypoints • Display if a waypoint has been visited • Map • Display explored areas of environment • Display where the MAV has been • Display features of environment (i.e., buildings, terrain, etc.) • MAV • Display location on map • Display direction of travel • Display sensor orientation • Display current altitude • Display sensor information
Fine-Tune Control	• Display current sensor information • Display current altitude • Display vertical velocity • Display if flight controls are engaged • Flight controls • Display position change commands • Display altitude change commands • Display orientation change commands
Identification of POI and/or OOI	• Display sensor information (e.g., a video feed)
Health & Status	• Display battery life of MAV • Display state of sensors • Display quality of location reckoning (both MAV and hand-held devices) • Display connection state • Display connection quality

3.4 MAV-VUE Displays and Interaction

MAV-VUE is an application for an iPhone/iPod Touch which satisfies many of the functional and information requirements for an interface to be used by an operator collaboratively exploring an environment with a MAV. While MAV-VUE is implemented on the iPhone OS, the interface is platform agnostic and could be implemented on many other hand-held devices. MAV-VUE is organized to allow the operator to interact with the MAV in two different modes, appropriate to different tasks. The Navigation mode allows the operator to have the MAV autonomously fly between specified waypoints. In flight mode, also known as Nudge Control, operators can fly the MAV to perform fine-tuned adjustments for adjusting the position and orientation of the MAV. These features are discussed in detail below.

3.4.1 Navigation Mode: Map & Waypoints

A map (Fig. 3-1) of the environment occupies the entire iPhone display, which is 320x480 pixels (px). The map displays relevant features of the environment, as well as the location of the MAV and waypoints.

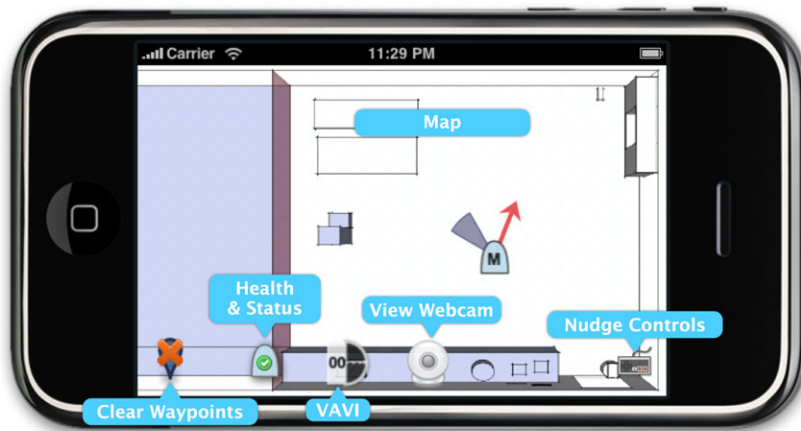


Figure 3-1: The map display and general interface of MAV-VUE.

Given the small display size of the iPhone, the user may zoom in and out of the map by using pinching and stretching gestures, as well as scroll the map display along the x or y axis by dragging the display with a single touch. Both actions are established User Interaction (UI) conventions for the iPhone interface. The MAV is represented by the air track friendly symbol from MIL-STD-2525B [38].

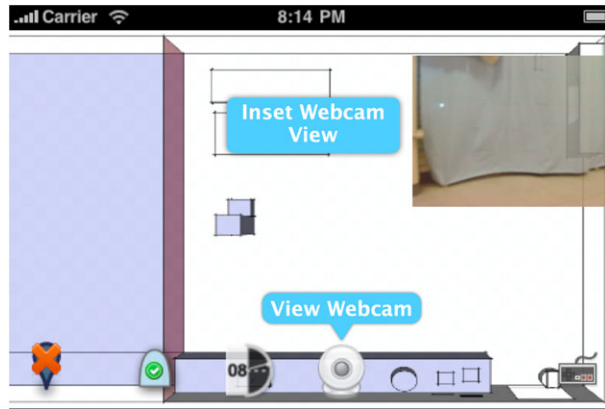


Figure 3-2: The inset webcam view within the map display.

As seen in Fig. 3-3, the MAV's direction and velocity are represented by a blue vector originating from the center of the MAV. The length of the vector indicates the speed of the MAV. Likewise, a blue arc shows the current orientation of the MAV's camera. The spread of this arc is an accurate representation of the FOV of the on-board camera.



Figure 3-3: Details of the MAV-VUE map display.

The map is intended mainly for gross location movements in the flight plan, with the Nudge Control mode (Section 3.4.4) intended for more precise movements while viewing the camera. As such, the map allows the user to construct a high-level flight plan using waypoints. The MAV autonomously flies between each waypoint with no action by the user. Each waypoint is represented by a tear-drop icon. The icon changes color depending on its current state, shown in Fig. 3-4. The flight plan is displayed as a solid line connecting waypoints in the order they will be visited.



(a) The waypoint has been created by the user, but has not been received by the MAV



(b) The waypoint is in the MAV's current flight plan.



(c) The waypoint has been visited by the MAV.

Figure 3-4: Waypoint states in the map display

Users double-tap on the map display to create a waypoint at the location of their taps (Fig. 3-3). This waypoint is then added to the queue of waypoints and transmitted to the MAV. Acting autonomously, the MAV plans a path between all of the given waypoints with no human intervention, avoiding known obstacles. In this capacity, the MAV is acting at the 7th LOA (Table 2.1), although the human dictates the higher-level objectives of the plan through waypoints, which helps to mitigate many of the concerns associated with higher LOA (e.g. loss of SA). At any point, the human may delete all waypoints and force the MAV to hover in place by tapping the Clear Waypoints icon.

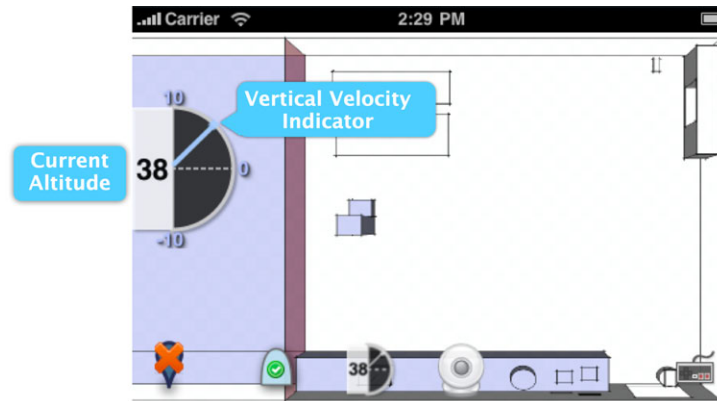


Figure 3-5: The Mini-VAVI shown on the map display.

3.4.2 Mini Vertical Altitude and Velocity Indicator

Embedded within the Map display is a miniaturized version of the Vertical Altitude and Velocity Indicator (VAVI) (Fig. 3-5), originally developed by Smith et al. [39]. In the map display, the VAVI shows the current altitude of the MAV as well as its vertical rate of change. Users may hide or show the VAVI by tapping on the VAVI button (Fig. 3-1). This button also displays the same real-time information as the full-sized VAVI, albeit in an extremely small format. A VAVI is designed to co-locate altitude and vertical velocity, in order to easily ascertain the altitude and vertical velocity of a VTOL air vehicle.

Smith et al., showed the VAVI significantly reduced cognitive workload [39]. Given the compact display size for MAV-VUE, the miniaturized VAVI only shows the right velocity arm while maintaining the same functionality. In post-experiment interviews, Smith found only half of the participants used the left arm [40].

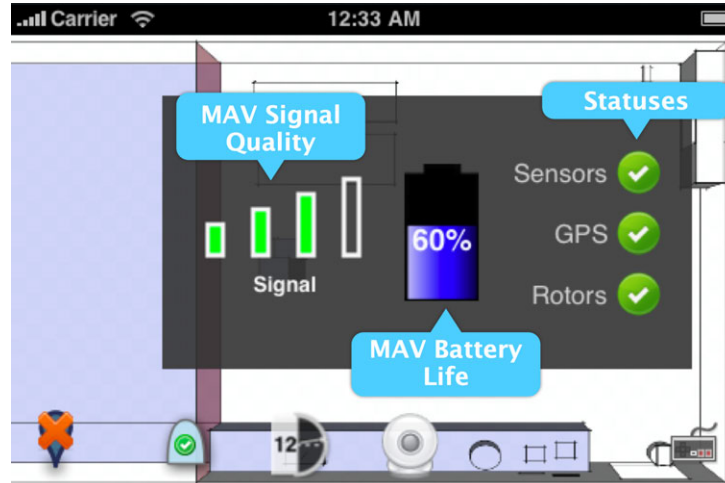


Figure 3-6: Health and Status monitoring display in MAV-VUE.

3.4.3 Health & Status Monitoring

A separate display was created for health and status monitoring of the MAV, shown in Fig. 3-6. By tapping on the MAV status button at the bottom of the map display, the user may show or hide the health and status monitoring pane. The display shows the current battery level of the MAV, signal quality, and status of various sensors on the MAV. The status of all rotors is shown, as a quad-rotor MAV can still fly with a faulty rotor, although it will have diminished performance.

3.4.4 Nudge Control Flight Mode

Nudge Control allows an operator fine-grained control over the MAV, which is not possible with the more general Navigation mode (Sec. 3.4.1). Nudge Control allows an operator the ability to more precisely position the camera (and thus the MAV) both longitudinally and vertically, in order to better see some object of person of interest. Nudge Control is accessed by tapping on an icon at the bottom of the map display (Fig. 3-1). Within the Nudge Control display, users are shown feedback from the MAV's webcam. Nudge Control in MAV-VUE (Fig. 3-7) can be operated in one of two modes on a hand-held device: Natural Gesture (NG) mode and Conventional Touch (CT) mode. For the NG mode, the device should have accelerometers, an Inertial Measurement Unit (IMU), or equivalent technology to provide information on its orientation in three dimensions. The NG mode is intended to be used with such an

orientation-aware device, while the CT mode only requires conventional multitouch technology.

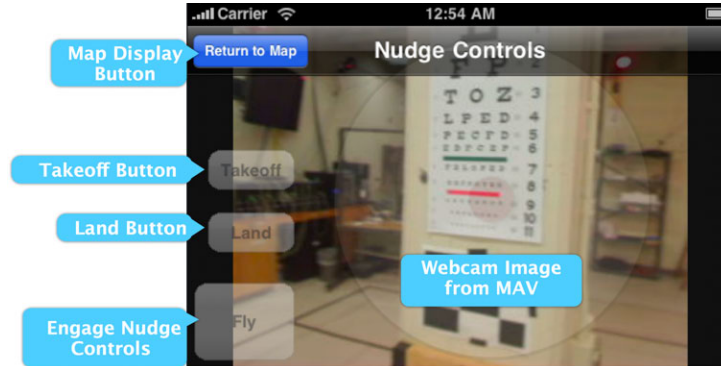


Figure 3-7: Overview of Nudge Control interface.

To activate Nudge Control, the user presses and holds down the Fly button. As long as the button is pressed, Nudge Control move commands are issued once per second, which was an empirically determined value from user testing and prototyping. The Fly button acts as a “dead man’s switch” to prevent the user from inadvertently making the MAV move, (i.e., due to distraction or dropping the device). When users press and hold the Fly button, the opacity of the directional controls increases to provide visual feedback that the user can now move the MAV.

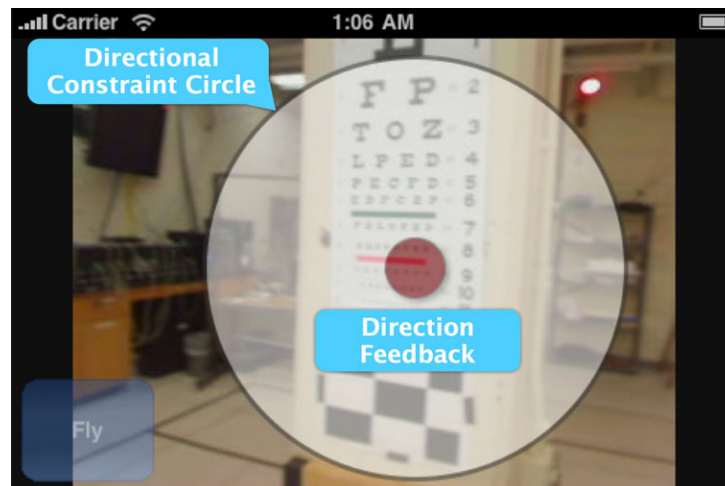


Figure 3-8: Details of Nudge Control directional interface.

The opacity of the directional controls was purposely chosen to partially obfuscate the webcam view to prevent users from trying to analyze imagery in detail while still controlling the MAV. Although this forces the user to choose between either analyzing imagery from the webcam or flying, this trade-off prevents the user from experiencing mode confusion, or becoming cognitively over-tasked and trying to examine a small, imperfect imagery feed.

3.5 Design Principles

Numerous studies have shown that users perform better when displays incorporate direct perception-action visual representations (e.g. [41], [42], [43]), which involve users' more efficient perceptual processes rather than more cognitively intensive processes of interference and recall. Several of these techniques, along with other design principles, were used in the design of the interface to improve the user's performance, and are discussed in more detail in the subsequent sections.

3.5.1 Direct Manipulation

The principle of direct manipulation asserts that a user should directly interact with objects and representations on a display, rather than issue commands which affect the objects [44]. Direct manipulation aims to decrease a user's workload by allowing them to continuously manipulate objects in a quick, easily reversible manner. It also transfers mappings and interactions users have learned in the real world, thus avoiding having to learn new mappings and interactions. Direct manipulation decreases users' cognitive workload by reducing the number of steps (including memory recall) a user must perform to accomplish a task. In the MAV-VUE interface, direct manipulation is used in the waypoint interface to allow users to quickly add and delete waypoints. Nudge Control also relies on direct manipulation, reducing a user's need for training by leveraging their existing real-world mappings of moving a flying vehicle by tipping the device in the direction of travel in NG mode, or making a dragging gesture in CT mode, described in detail in Section 3.5.3.

3.5.2 Multimodal Interaction

Multimodal interaction uses two or more distinct mediums of interactions to increase the usability of an interface [45]. Multimodal interaction was first explored in 1980 with the Put-that-there interface [46], but largely existed only in a research setting for the next twenty years. Advances in computing power and cheap electronics have made multimodal interfaces much more accessible to the general public, with several smart phones, cameras, and other electronics incorporating a combination of voice, touch and tilt modalities. Multimodal interfaces are prone to several common misconceptions and myths [47], but used properly, they allow users to more efficiently work with a computer in a manner which is most natural for the task at hand. By mixing modalities, designers are free to concentrate on implementing interfaces which are best suited for the task at hand, rather than trying to design a compromised interface. Nudge Control relies on multimodal interaction (multitouch gestures and tilting) to create a rich control interaction without cluttering the display. MAV-VUE can be implemented on both conventional and natural gesture-enabled devices.

3.5.3 Order Reduction

Human control of systems which incorporate one or more feedback loops is defined as a N^{th} order system, where N refers to the derivative of the feedback loop used in the controls. A 1st order feedback loop responds to changes in the first derivative of the system (i.e., velocity derived from position). Error, the difference between the output and the desired state, is fed back to the input in an attempt to bring the output closer to the desired state. For example, changing the heading of a MAV via a first-order feedback loop requires constantly changing the MAV's rate of yaw until the desired heading is reached. Typically this is executed by humans as a pulse input which requires at least two distinct actions: first starting the turn, then ending the turn, as seen in Fig. 3-9. In contrast, with a 0th order control loop, an operator simply gives a command with the desired heading (i.e., South) and the vehicle autonomously turns to this heading. A 1st order system requires more attention by the operator as compared to a 0th order system since he or she must continually oversee the turn in order to stop the vehicle at the right time.

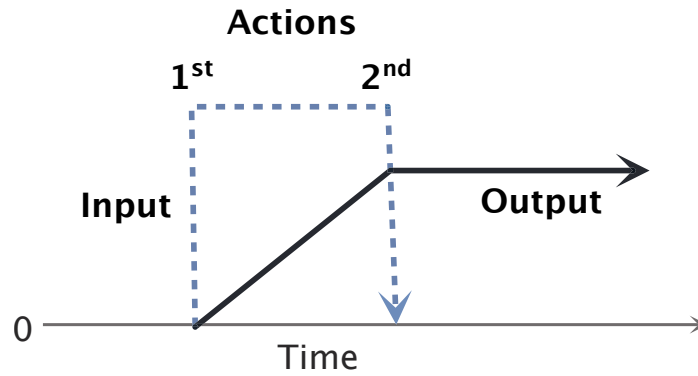


Figure 3-9: 1st order feedback control with a pulse input.

A 2nd order control loop relies on changing the acceleration of the system. It is generally recognized that humans have significant difficulty controlling 2nd order and higher systems [48]. Due to the increased complexity of the feedback loops and number of actions required to successfully complete a maneuver, an operator's cognitive workload is significantly higher for 2nd order systems than when operating 0th or 1st order controls, leading to lower performance as shown by Sheridan et al. [20, 8, 7]. Teleoperation only exacerbates these problems because additional latencies are introduced into the system, in addition to the lack of sensory perception on the part of the operator, who is not physically present. As a result, all UAVs have some form of flight control stabilization (i.e., fly-by-wire) since they use teleoperated 2nd order, or higher, control loops [49, 50]. While human pilots are thought to be effective 1st order controllers [48], it is doubtful whether UAV pilots can effectively execute 11th order control of UAVs. One-third of all US Air Force Predator UAV accidents have occurred in the landing phase of flight, when human

pilots have 1st order control of the vehicles. As a result, the US Air Force will be upgrading their fleet of UAVs to include autoland capability [51], effectively reducing the pilot's control to 0th order. System communication delays, the lack of critical perceptual cues, and the need for extensive training, which result in pilot-induced oscillations and inappropriate control responses, suggest 1st order control is a poor approach to any type of UAV control. This problem would likely be more serious for MAV operators who are not, by the nature of their field presence, able to devote the necessary cognitive resources needed to fully attend to the MAV's control dynamics.

Even though it is well-established that humans are not good at controlling second order and higher systems [48], people can effectively steer a car (a 2nd order system [52]) because they can see the road ahead. This is a form of preview which effectively reduces the order of the system, in the case of driving, to a 1st order system. So humans can control higher order systems with some preview display, however, there is some mental workload cost. This example demonstrates that visualizing predictive information reduces some complexity. However, the underlying control system is still a higher order control loop, which is more difficult to control as compared to a lesser order interface. While vehicles obtain higher performance with more complex control loops, humans' performance degrades rapidly for 2nd order control loops and higher, especially when they are robbed of critical visual preview displays. By comparison, a 0th order control loop significantly reduces the workload because the operator does not need to continually monitor the movement of the vehicle (e.g., as it turns to a new heading), however, there is some cost in vehicle maneuverability. For operating a MAV, 0th order interfaces represent the highest degree of safety because users are not prone to errors as they try to calculate the position of the vehicle.

In the context of operating an UV, such as a MAV, in a hostile setting which requires the operator's divided attention, it is imperative that a robot be operated by 0th order control whenever possible. Operators are under high workload, and their primary task is obtaining imagery (i.e., ISR missions), not flying the vehicle. Classic solutions to this problem such as using teleoperation have relied on trying to mitigate cognitive complexity by simply reducing a system's order through introducing additional feedback loops (Fig. 2-2). First order control systems require significantly more training before an operator can successfully operate an aircraft as compared to 0th order systems. However, 1st order controls are typically used for the precision orientation and positioning of a UV to obtain effective imagery required in an ISR mission. For these precise movements, 0th order control systems can be cumbersome and difficult to use, hence the low success rates of many teams at USAR competitions [17] and the findings in Durlach et al.'s study [29]. Unfortunately, providing a 1st order control interface to a MAV operator can cause human control instabilities, also as demonstrated by the Durlach et al. study. In addition, environmental pressures of a hostile setting, the need for formal and extensive training, and the need for divided attention of the operator suggest that 1st order systems are not appropriate for MAV control. As such, some balance between using 0th and 1st order control is warranted.

3.5.4 Perceived First Order Control

This thesis asserts that Perceived First Order (PFO) control can provide a stable and safe 0th order system control, while allowing operators to perceive 1st order control so as to achieve effective control of an ISR MAV with minimal training. The intention is to provide a design compromise which increases performance and safety by using a level of feedback which is appropriate for each aspect of the system (including the human). While users perceive that they are operating the vehicle via a 1st order control interface, PFO control actually uses a 0th order control loop to prevent the user from putting the vehicle in jeopardy. This allows the user to accurately and easily predict the movement of the MAV, as well as easily formulate plans. Users are given visual feedback (Fig. 3-10) of their 1st order commands by showing a red dot on the display, which is overlaid on top of sensor imagery. These commands are limited

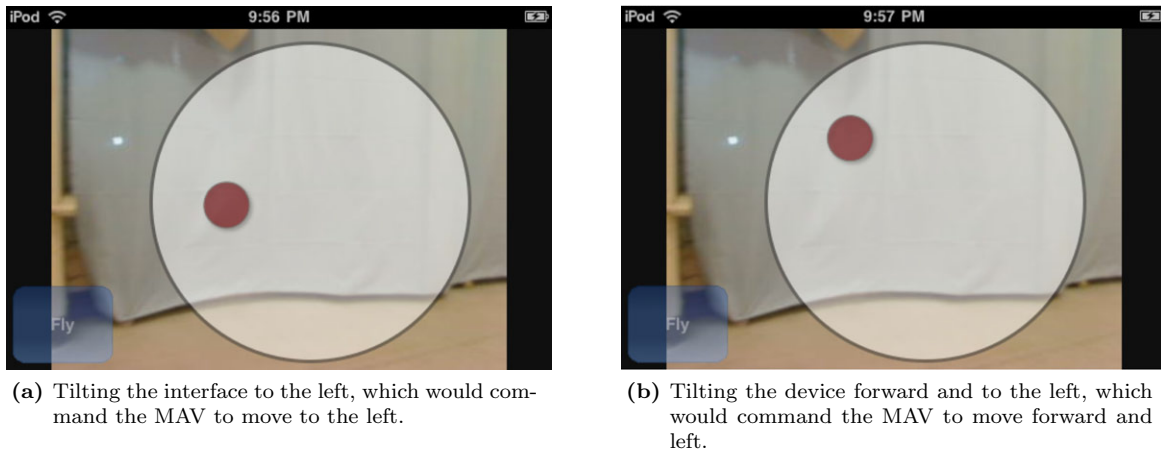


Figure 3-10: Interface feedback as a result of the operator performing a natural tilt gesture with the device.

by a constraint filter which prevents the user from over-shooting their target, represented by the white circle. Consequently, this interface allows users to feel like they have greater control over the vehicle's movements and orientation through what appears to be direct control. However, PFO control converts the user's 1st order commands, (relative velocity changes) into a 0th order control system, i.e., position updates (Fig. 3-11). By working in a 0th order control loop which uses absolute coordinates, commands are time invariant, unlike velocity or acceleration commands. This time invariance eliminates the problem of over/under-shooting a target inherent to 1st or 2nd order control systems with a "bang-bang" set of commands [48].

This approach also helps to mitigate known problems with time lag, caused by both human decision making and system latencies. This blend of 1st and 0th order control loops drastically decreases the training required to effectively use an interface for an ISR mission. PFO control is a way to get the best of both position and velocity control while giving the user enough control that they feel they can

effectively perform their mission without risking the vehicle's safety. PFO can be used in one of two different modes. NG mode uses a set of tilting gestures while CT mode uses multitouch gestures to interact with the vehicle. Both of these modes are described in more detail in the next section.

Natural Gesture Mode

In NG mode, an operator updates the x and y location of the MAV by tilting the entire device in the direction operators intend the UV to travel (Fig. 3-8). The user may also control the heading (ψ) and altitude (z) of the MAV. The Two Dimensional (2D) tilt vector defines the relative distance along the x/y axes from the MAV's existing location (which is considered the origin). A discrete-time high-pass filter is used to clean the incoming acceleration data. The angle and direction of tilt is detected by the orientation sensors of the device.

Conventional Touch Mode

In CT mode, the operator can also control the x/y direction of travel by touching and dragging in the direction intended from the center of the display (Fig. 3-12). The length of the drag corresponds to the relative distance to travel while the angle of the drag corresponds to the direction in which the MAV should travel.

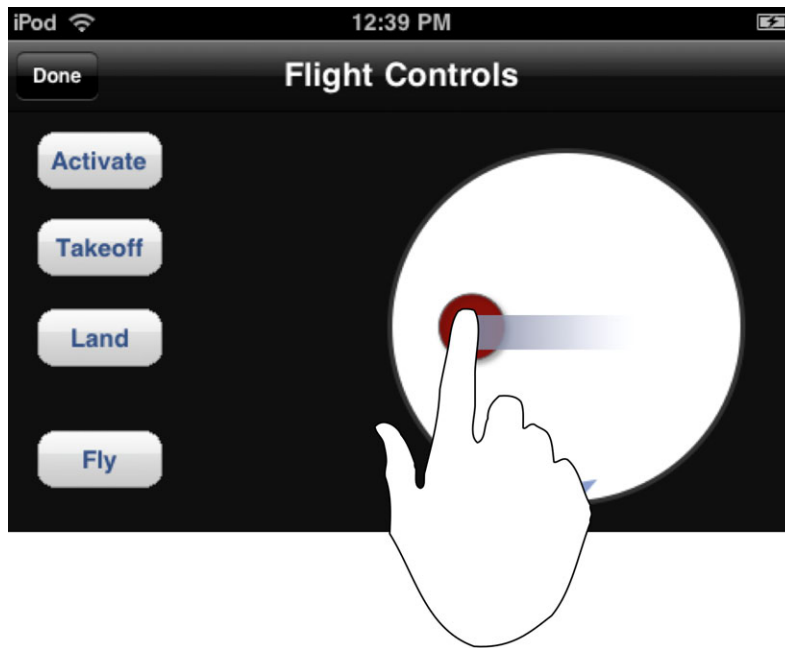


Figure 3-12: Using a swipe touch gesture to move the MAV left

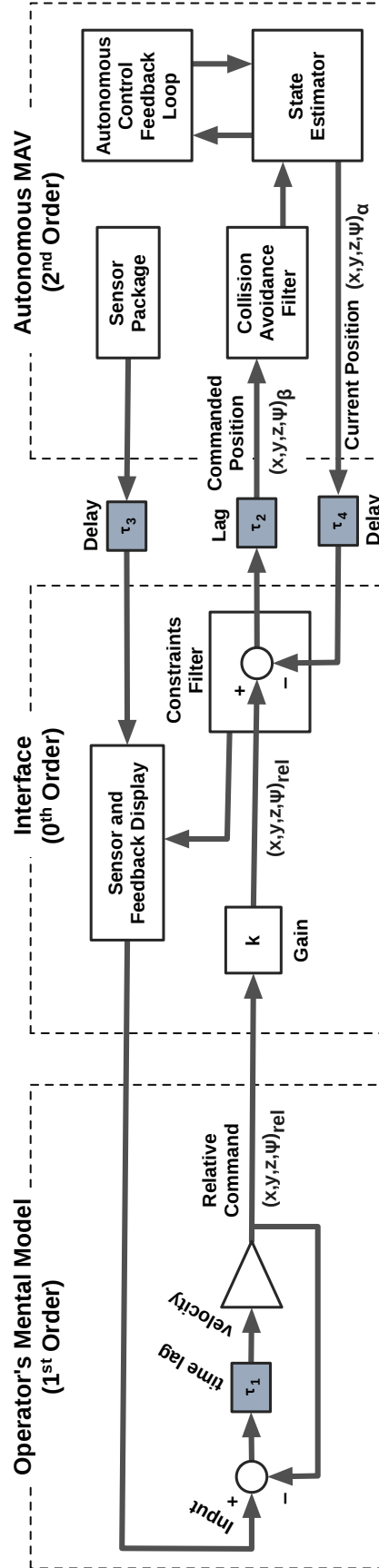


Figure 3-11: Diagram of Perceived First Order Control. Delays and lags are represented by τ , gains represented by k .

Constraint Filter

While the NG tilt and CT drag actions convey a sense of position and velocity to the operator, the actual action is first translated into a set of relative distance (*rel*) coordinates, $(x\ y\ z)_{rel}$ (e.g., in meters, $x\ y\ z$, and an angle, θ), from the MAV's current location, using the gain function, k (Fig. 3-11). A constraint filter limits the magnitude of user commands, as well as modifying commands which could send the MAV into an inaccessible region (i.e., No Fly Zone (NFZ)). The constraint filter, using Equation 3.1, translates the relative distance coordinates (*rel*) into an absolute set (θ) of coordinates, $(x\ y\ z)$ (e.g., latitude/longitude/altitude/heading). The MAV can move to these coordinates (θ) by incorporating feedback from the autonomous MAV's current (θ) coordinates, $(x\ y\ z)$, provided by its state estimate filter.

$$(x\ y\ z) = k((x\ y\ z)_{rel}) + (x\ y\ z) \quad (3.1)$$

The coordinates, θ , generated by the constraint filter are bounded by obstacle collision algorithms, which evaluate whether the coordinates exist within a space accessible to the MAV. This evaluation can be based on input from sensors (i.e, LIDAR), representations of the environment (i.e, Simultaneous Localization and Mapping (SLAM)), or user-defined parameters (i.e, NFZ). If the MAV cannot move to the given coordinates, a set of accessible coordinates close to the desired location are used instead. If no accessible coordinates can be determined, the device reverts to coordinates predetermined to be ultimate safe zones, which are vehicle and handheld-device dependent (i.e., a 2 foot (ft) radius could be determined *a priori* to be a safe zone without any external input).

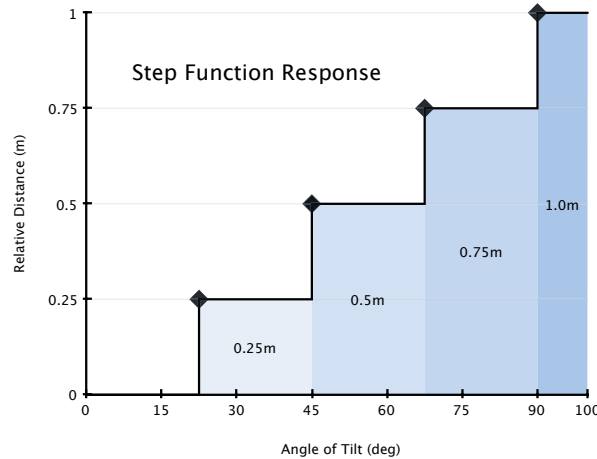


Figure 3-13: Discrete step function initially used for gain, k .

In both NG and CT mode, the constraint filter is represented visually by the semi-opaque directional circle overlaying imagery (Fig. 3-8). The maximum $x\ y$ movement the operator is allowed to take is represented by the outer circle of the display. This maximum distance can be an absolute operational

limit, such as a small distance for novice users, or based on the MAV's sensor's FOV, which effectively do not allow the operator to move the MAV beyond the area which can be sensed. This allows the device to translate the user's relative movement (*rel*) into an absolute set of coordinates, () based on the MAV's current position (). The constraint filter limits the user's commands as well as preventing them from issuing a command to move to a set of coordinates in a region known to be inaccessible.

Initially, the gain function, k in Fig. 3-11, was implemented as a discrete step function, as shown in Fig. 3-13. However, pilot testing found this gain function was impractical due to the lack of feedback when the function moved to a different step value. Consequently, a linear function was implemented, as shown in Fig. 3-14.

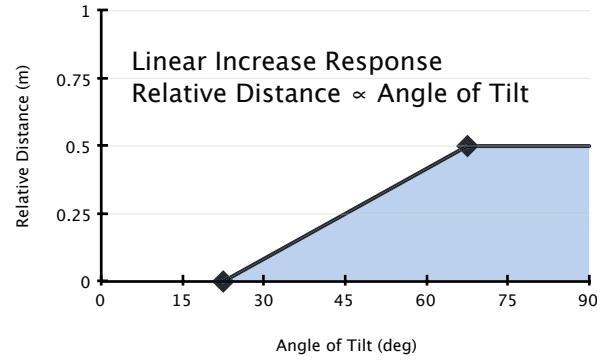


Figure 3-14: Linear step function used for gain, k .

The absolute coordinates are transmitted as a move command to the MAV with an expected time lag, τ_2 (Fig. 3-11). The MAV then moves to the commanded coordinates using its internal autonomous software and hardware. For an ISR mission, the MAV should periodically transmit information from its sensor package, with presumed delay τ_3 (Fig. 3-11), for the interface to update its sensor and feedback display. Given the nature of the interface, it would be preferable to update any visual images at or above 20 Hertz (Hz) [53].

Altitude Mode

An operator can change the z value of the new coordinates through two types of interactions. Performing a pinch or stretch gesture on the right control interface will cause the device to issue a new position command with a change in the z -axis. A stretch gesture results in a relative increment of the z coordinate, while a pinch gesture causes a relative decrement (Figs. 3-15a, 3-15b).

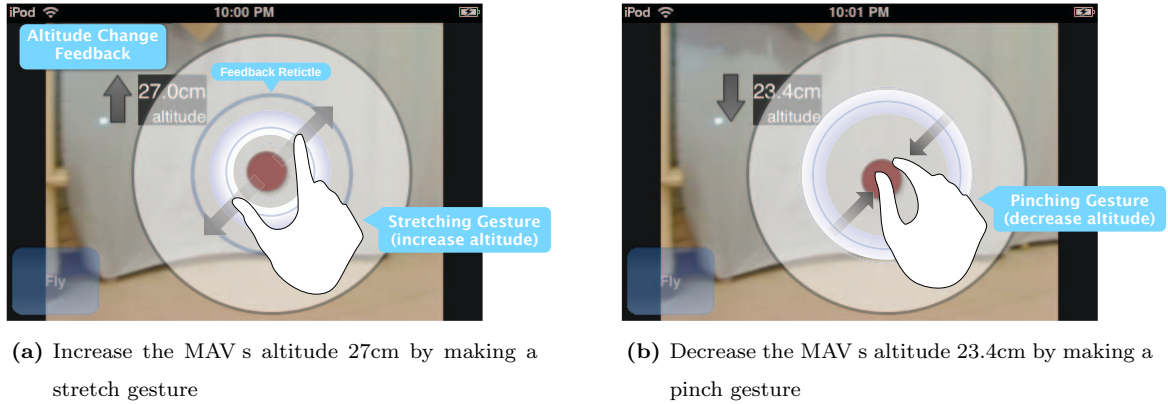


Figure 3-15: Gestures to change the z coordinate (altitude) of the MAV

As the operator performs these gestures, a set of circular rings provides feedback on the direction and magnitude of the gesture. As the user performs the gesture, the proposed altitude change is shown on-screen along with an arrow indicating the direction of travel. Likewise, an operator could change the MAV's height by dragging the right tape of the mini-VAVI up or down to a new height (Figs. 3-16a, 3-16b). Normally shown in the Nudge Control Mode when the Fly button is engaged, the VAVI is not implemented on the iPhone due to limited processing power, but users can still change altitude by performing a pinching or stretching gesture on the display.

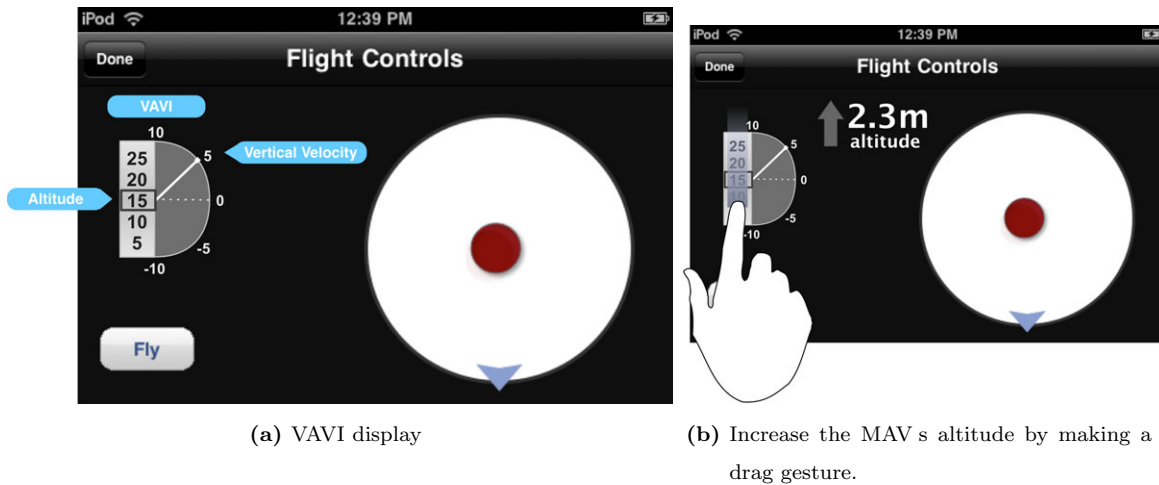


Figure 3-16: Using the VAVI to change the z coordinate of the MAV.

Automatic Nudge Control Deactivation

Should the operator drop the device or otherwise become incapacitated, several methods could be used to prevent inadvertent commands as the device tumbles through the air. Currently, the Fly button, dead

man’s switch,” on the display only allows commands to be sent in Nudge Control mode while it is pressed and held down. Alternatively, more sophisticated algorithms could detect the device in free fall, or detect no activity by the operator, and stop sending commands.

Heading Control

Operators indirectly control the yaw and pitch of the MAV’s sensors through natural gestures or conventional touch interactions. In NG mode, the sensor’s orientation is determined by performing a swiping gesture across the screen (Fig. 3-17).

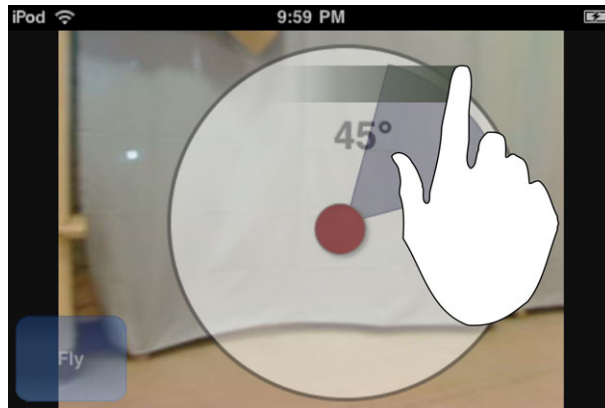


Figure 3-17: Swiping a finger across the screen causes the device to rotate (yaw) right or left.

The magnitude and direction (left or right) of the swipe corresponds to the magnitude and direction of the relative yaw command. An operator may also change the yaw orientation of the sensor by using the CT mode to tap on the circumference of the constraint circle (since the swiping gesture is used for directional control), which corresponds to an angle, θ , in polar coordinates which is used to change the yaw. Internally, the device performs the appropriate calculations to use either the sensor’s independent abilities to rotate, or, if necessary, the vehicle’s propulsion system to rotate the entire MAV, moving the sensor to the desired orientation. This device therefore leverages existing automated flight control algorithms to adjust yaw, pitch, and roll given the position updates that are translated via the user’s interactions.

3.6 Architecture

The MAV-VUE application is implemented using the iPhone SDK and open-source frameworks in Objective-C. The application uses a Model-Viewer-Controller (MVC) paradigm to organize code and define abstractions. A majority of the displays are natively drawn using the Core Animation graphics API. This application relies on a server application, MAVServer, which acts as a middleware layer interfacing be-

tween the iPhone application and the MAV's software. The MAVServer exists as a means to off-load computationally intensive tasks from the iPhone, log experimental data, and for ease of implementation, as the OS X environment is much easier to develop in than the iPhone environment. However, in the future as the computational power and developer environment mature on hand-held devices, this server could be eliminated entirely.

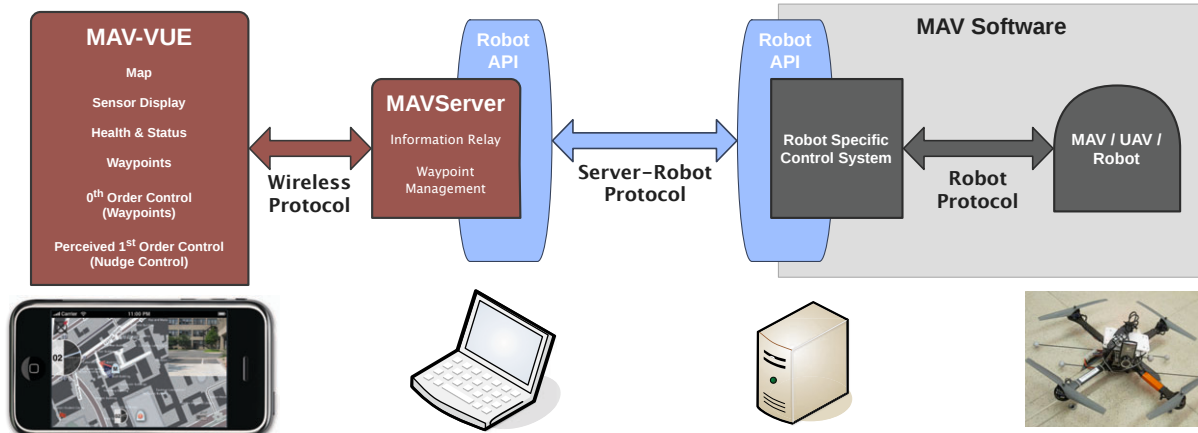


Figure 3-18: Communication architecture between MAV-VUE, MAVServer, and MAV.

Communication between these components is shown in Fig. 3-18. Although using a MAV is the focus of this thesis, the architecture is agnostic and treats the MAV as a subclass of a generic robot. Communication between the iPhone and the MAVServer occurs over wireless (802.11) using Transmission Control Protocol (TCP)/Internet Protocol (IP). The TCP/IP payload is a BLIP message, which encodes binary data (such as images) or JavaScript Object Notation (JSON)-formatted text (e.g. location updates, commands). Above a pre-determined payload size, the message is compressed using gzip [54]. Webcam images are transmitted in JPEG format, while map images are transmitted in PNG-24 format.

3.7 Summary

A CTA with potential field operators revealed many of the shortcomings in current ISR tactics and intelligence. These interviews provided a set of system requirements for a MAV to be used for ISR, as well as the roles and submissions that MAV could perform. Based on these roles, functional and information requirements were designed which provided a framework defining the features and uses of a hand-held interface.

MAV-VUE was designed in response to the need for a hand-held device interface which allows a human to perform an ISR mission with a MAV safely and effectively, even as his or her attention is divided between the interface and other critical tasks. This interface, MAV-VUE, allows users to easily control a MAV performing an ISR mission in one of two modes, Navigation and Nudge Control. Although

a large tablet device may provide more screen real estate, and a joystick a sense of “controlling” a MAV, MAV-VUE reflects the real-world constraints of the hostile environments in which these field personnel are likely to operate. In many instances while designing the interface, the question was not how to add more functionality, but how automation could augment functionality to better support operators with divided attention.

PFO control represents a novel way to take the best of both worlds in human control loops. While a 1st order interface may be more satisfying to use than a cumbersome 0th order interface in terms of precise control, it also requires more cognitive work on the part of the operator, requires significantly more training, and introduces many risks. In contrast, 0th order interfaces may be simpler and easier, but this simplicity could irritate users, potentially leaving them frustrated with the interface. With sufficient automation, there is no longer a reason to dictate that users must adhere to the same order of control as the interface or the vehicle. Instead, MAV-VUE chooses an order of control which best fits the user’s mental model (1st order) but provides the safer confines of a 0th order control loop. An important hypothesis that will be addressed in the next chapter is that PFO control should be easy to learn and safely use in a manner of minutes.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 4

Usability Evaluation

4.1 Study Objectives

A usability study was conducted to assess the MAV-VUE interface with untrained users, who completed a short MAV ISR task requiring navigation in an artificial urban environment. Performance was compared with a model of an “ideal” human, who performed this task perfectly, to understand how well the interface aided users with no specialized training in gaining SA and performing supervisory control of a MAV. The objective of this study was to ascertain the usability of hand-held interfaces for supervisory control and remote quasi-teleoperation of an autonomous MAV.

4.2 Research Questions

To achieve these objectives, the following research questions were investigated:

1. Does the interface allow a casual user to effectively control a MAV?
 - (a) Do users find the interface intuitive and supportive of their assigned tasks?
 - (b) Can the user effectively manipulate the position and orientation of the MAV to obtain information about the environment?
 - i. While given gross control of the MAV?
 - ii. While using fine control of the MAV?
 - (c) How well does a casual user perform the navigation and identification tasks compared to the model of a “perfect” participant?
2. Is the user’s SA of the environment improved by the information provided in the interface?

- (a) Is the user able to construct a mental model of the environment as evidenced by the time to complete objectives and the accuracy of their identifications?
 - i. Through the provided map?
 - ii. Through the provided webcam imagery?
- (b) Can the user find and accurately identify an OOI and/or a POI using the interface?

4.3 Participants

Fourteen participants (8 men and 6 women) were recruited from the MIT community using email. All participants were between the ages of 18 and 29, with an average age of 22 years (standard deviation (sd) 2.93 years). All self-reported corrected vision within 20/25, and no color blindness. The participant population ranged in experience with an iPhone and Remote Control (RC) vehicles from none to self-reported experts. Appendix B details more information on the participant demographics.

4.4 Test Bed

4.4.1 Apparatus

The study was conducted using one of two second generation iPod Touches (named Alpha and Bravo) running MAV-VUE with only NG Nudge Control enabled. Each had a screen resolution of 320x480px and 16-bit color-depth. Both iPods were fitted with an anti-glare film over the screens. The MAVServer was run on an Apple MacBook, using OS X 10.5 with a 2 Gigahertz (GHz) Intel Core 2 Duo and 4 Gigabytes (GB) of memory. Wireless communication occurred over one of two 802.11g (set at 54 Megabits (Mb)) Linksys 54G access point/routers, running either DDWRT firmware or Linksys firmware. The MacBook communicated with the Real-time indoor Autonomous Vehicle test ENvironment (RAVEN) motion-capture network over a 100Mb ethernet connection. The RAVEN facility [4] was used to control the MAV and simulate a GPS environment. Custom gains were implemented to control the MAV based upon the final vehicle weight (Appendix C)

An Ascending Technologies Hummingbird AutoPilot (v2) quad rotor was used for the MAV. This Hummingbird (Fig. 4-1) was customized with foam bumpers and Vicon dots to function in the RAVEN facility and the GPS module was removed. 3-Cell Thunderpower™ lithium polymer batteries (1,350 milli-amperes (mA) and 2,100 mA capacity) were used to power the MAV. Communication with the MAV was conducted over 72 Megahertz (MHz), channels (ch) 41,42,45 using a Futurba™ transmitter and a DSM2 transmitter using a Spektrum™ transmitter to enable the Hummingbird serial interface. The computer-command interface occurred over the XBee protocol operating at 2.4 GHz, ch 1. The MAV was

controlled at all times through its serial computer-command interface and the RAVEN control software which autonomously flew the MAV between given waypoints.



Figure 4-1: Modified Ascending Technologies Hummingbird AutoPilot used in study.

A Gumstix™ Overo Fire COM (4GB, 600MHz ARM Cortex-A8 CPU, 802.11g wireless adapter, Gumstix OE OS) with Summit Expansion Board was mounted on top of the MAV in a custom-built enclosure. Mounted on top of the MAV was a Logitech C95 webcam, with a maximum resolution of 1024x768px and a 60° FOV. The webcam was configured with auto-white balance disabled, focus at infinity, resolution at 480x360px, and connected to the Summit Expansion board via Universal Serial Bus (USB) 1.0. Webcam images were captured and transmitted in JPEG format, quality 90, via wireless using User Datagram Protocol (UDP) and a custom script based on the uvccapture software from Logitech™ limited to a maximum rate of 15 frames per second (fps), although the frame rate experienced by the user was lower due to network conditions and the speed of the network stack and processor on the iPod. The Gumstix and webcam were powered using 4 AAA 1,000 mA batteries. The total weight of the webcam, Gumstix, batteries and mounting hardware was 215 grams. Testing before and during the experiment indicated there was approximately a 1-3 second delay (which varied due to network conditions) from when an image was captured by the webcam to when it appeared in MAV-VUE. Position updates and sending commands between MAV-VUE and the MAV (i.e., creating a waypoint or a nudge control movement) typically took between a few milliseconds (essentially instantaneously) and 300-500 ms, dependent on the calibration of the VICON system and the quality of the XBee radio link.

4.5 Experiment Metrics

4.5.1 Task Performance Time

Participants' performance at searching for and identifying an eye chart and POI was compared to hypothetically perfect performance, which represented flying the test course with the minimum number of actions necessary. This performance differential assesses the role training and experience play in the ability to successfully operate the MAV. A small difference in performance supports the assertion that the interface requires little training before it can be effectively used by an operator.

4.5.2 Spatial Reasoning

Subjects' spatial reasoning abilities may be critical in their ability to use the interface for an ISR mission in an unknown environment. To account for this variability, participants were given two written tests to assess their spatial reasoning capabilities:

Vandenberg and Kuse Mental Rotation Test (MRT) Score [55]: The MRT is a pencil and paper test used to establish a subject's aptitude for spatial visualization by asking him or her to compare drawings of objects from different perspectives. The original test has largely been lost and a reconstructed version from 2004 was used [56, 57, 58].

Perspective Taking and Spatial Orientation Test (PTSOT) [59, 60]: PTSOT is a pencil and paper perspective-taking test shown to predict a subject's ability for spatial orientation and re-orientation.

4.5.3 Additional Mission Performance Metrics

To assess the usability of the MAV-VUE interface, an additional set of quantitative metrics were used. These metrics provide insight into how well users performed with the interface for the search and identify tasks:

1. Imagery Analysis
 - (a) POI identification
 - (b) Eye chart identification
 - i. The line identified
 - ii. The accuracy of the line identified
2. Navigation
 - (a) Number of waypoints

(b) Length of path

3. Nudge Control

(a) Descriptive statistics of commands issued for changing $x/y/z$ and yaw.

4.5.4 Qualitative Metrics

Finally, a set of qualitative metrics was used to identify the subject's familiarity with Remote Control (RC) vehicles, iPhones, and other relevant demographic information. A post-experiment usability survey was given to judge participants' perceptions of their performance during the flights and of the interface (Appendix D). Participants were also interviewed after the experiment about their experience to gain further feedback about the interface (Appendix A)

4.6 Procedure

Each participant performed the experiment individually. Participants signed an informed consent/video consent form (Appendix E), and completed a background questionnaire (Appendix B) which asked about experiences with computers, the military, iPhones, and video games. After finishing the demographic survey, the PTSOT and MRT tests were administered.

Following these tests, the experiment and interfaces were explained in detail to the participant. The experiment administrator demonstrated taking off, navigating via waypoints, flying using nudge controls to find a POI (represented as a headshot on a 8"x11" sheet of paper) and landing the MAV once (on average, flying for two to three minutes). All flights were performed with the participant standing upright and holding the mobile device with two hands in front of them. Participants were allowed to ask questions about the interface during this demonstration flight. Participants then completed a short training task to become acquainted with the interface and MAV. During this training task, participants were asked to create four waypoints and use nudge controls to identify the same headshot which was shown during the demonstration flight. The participant was allowed to ask questions about the interface, and was coached by the demonstrator if they became confused or incorrectly used the interface. Aside from the demonstration and three minute training flight, participants were given no other opportunities to practice with or ask questions about the interface before starting the scored task.

Once a participant completed the training task, he or she was given the supplementary map (Appendix F) and began the scored task, which was to search and perform identification tasks in an urban environment for five to six minutes. During this time, the experiment administrator provided no coaching to the participant and only reminded them of their objectives. Participants flew in the same area as the training exercise, with a new headshot and eye chart placed at different locations and heights in the room

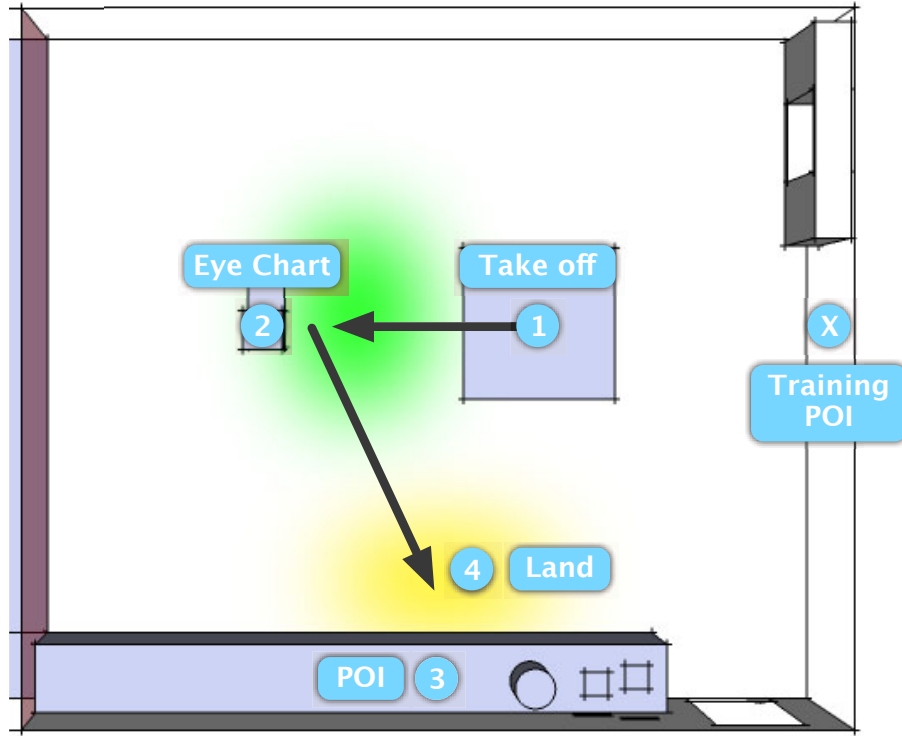


Figure 4-2: Annotated version of the map used in the study showing the layout of the environment.

(Fig. 4-2), with neither at the location used for the training POI.

Participants were first instructed to fly to the green area (Fig. 4-2, No. 2) indicated on the supplemental map using waypoints, and once there, to search for a Snellen eye chart (Appendix G) in the vicinity, which was placed at a different height (1.67m) than the default height the MAV reached after takeoff (0.5m). After identifying the eye chart, participants read aloud the smallest line of letters they could accurately recognize. Upon completing this goal, participants were asked to fly to the yellow area (Fig. 4-2, No. 4) of the supplementary map and to search the vicinity for a POI headshot (No. 3 in Fig. H-1) which was recessed into a box at location No. 3 in Fig. 4-2, placed at a height of 1.47m. After participants felt they could accurately identify the POI from a set of potential headshots, they were asked to land the MAV in place. Due to limited battery life, if the participant reached the five minute mark without reaching the POI, the MAV was forced to land by the experiment station. If the participant reached the POI with less than 30 seconds of flight time remaining, the station allowed the participant up to an extra minute of flight before landing. If the MAV crashed,¹ the experiment administrator would take over the MAV in its previous location before crashing and return control to the participant, with additional time allotted to compensate for the MAV taking over and stabilizing.

After finishing the task, the participant was asked to fill out a survey selecting the POI he or she

¹These crashes were never caused by subjects' actions, but instead were due to network anomalies and radio interference.

recognized during the flight from the photo contact sheet in Appendix H. Participants concluded the experiment by taking a usability survey (Appendix D) and answering questions for a debriefing interview conducted by the experiment administrator (Appendix I). Participants were paid \$15 and thanked for their participation. Each experiment took approximately 75 minutes.

4.7 Data Collection

Participants' navigation and flight commands were logged to a data file. The webcam imagery from each flight was also recorded, along with relevant parameters of the MAV's location, orientation, and velocity. Interface use was recorded on digital video. Field notes were taken during the experiment to record any emerging patterns or other matters of interest. Usability, mission performance, demographic, and experience data were collected by questionnaires, along with experiment debrief interviews.

4.8 Summary

This chapter presents the design of a usability study for evaluating the interface described in Chapter 3, and the metrics which were used to evaluate participants' performance. A commercial MAV was modified to function in an environment which provided simulated GPS coverage. Additionally, a webcam sensor package was developed using off-the-shelf hardware to provide streaming imagery to the interface. A usability study was designed to identify how well participants could use the interface to perform search and identification tasks in an urban environment. After a demonstration of the interface and hands-on training of up to three minutes, participants were required to find and identify an eye chart and POI headshot within five to six minutes. To evaluate participants' performance, a variety of qualitative and quantitative metrics were recorded for analysis in Chapter 5. Participants also took two written tests to evaluate their spatial abilities and to analyze if any correlation existed between their abilities and performance on the tasks. Participants also filled out a usability survey and were interviewed after the experiment to obtain subjective responses which could be used to improve the interface.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 5

Results and Discussion

5.1 Introduction

This chapter presents the results of the usability study described in Chapter 4. Subjects and the interface were evaluated using a combination of qualitative and quantitative metrics. One participant's times and Nudge Control command data was not used due to the MAV crashing, which occurred as a result of network interference and was not caused by the participant's actions. However, the participant's eye chart, POI, and demographic data was still used. Another participant's scored task was interrupted due to a faulty battery, forcing the MAV to land prematurely. The participant's overall time was adjusted to compensate for time lost to the landing, takeoff, and time needed to re-orient after take-off.

During the study, participants completed a scored task which had two main objectives: to find and read the smallest line of letters they could identify on an eye chart, and to find a POI which they were asked to identify after the eye chart task (see Section 4.6 for more details). For the scored task (Fig. 4-2), the participants flew a path, on average, 13.00 m long (sd 10.57 m) and created between one and six waypoints (median 3) in Navigation Mode. It took participants an average of 308 s (sd 52.76 s) to complete the scored task. Further descriptive statistics on participants' performance is shown in Appendix J. Given the small sample size, much of the focus of this chapter is on the qualitative evaluation of the interface. Non-parametric tests were used to analyze quantitative metrics when appropriate. An α of 0.05 was used for determining the significance of all statistical tests.

5.2 Score Task Performance Time

Participants, on average, took 308 s (sd 52.76 s) to complete the scored task (measured as the time from takeoff to the time a land command was issued). Subjects' times to complete the scored task were

compared to that of a hypothetical “perfect” human who performed the same task with no errors. Given the optimal course path of 4.77 m (Fig. 4-2), it was empirically determined that a perfect human subject would take approximately 83 s to complete the scored task. The time of 83 s was based on the speed of the MAV, the minimum number of inputs required to perfectly align the MAV to find and identify the eye chart and POI, and also incorporated the delay of receiving imagery from the quad. During the experiment, it was observed that this delay was typically between one and two seconds, with a maximum of three seconds. Therefore, the maximum time delay of three seconds was used in this calculation.

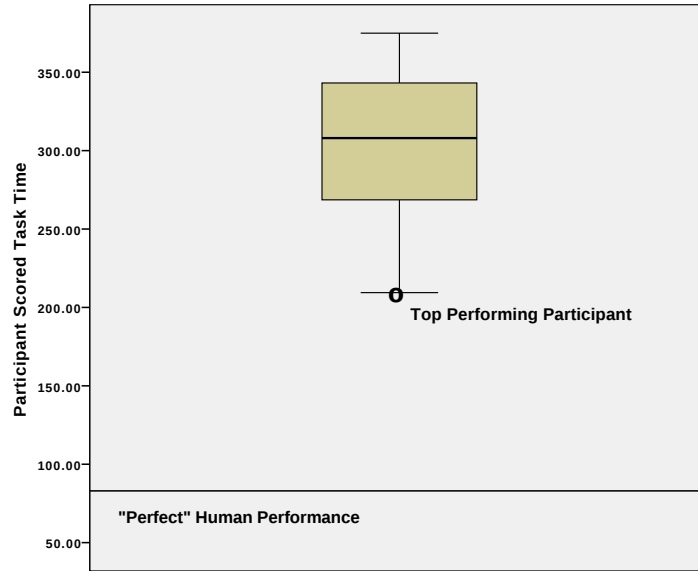


Figure 5-1: Box plot of participants' times to complete the scored task.

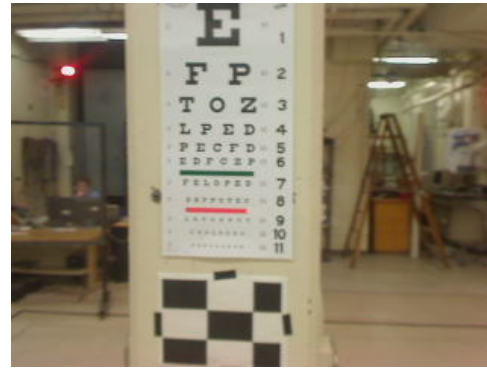
This ideal time was compared to the subject's flight time using a single point comparison (two-tailed, one sample student's t test), with $t(13) = 15.09$ and $p < 0.0001$. In comparison, the top performing participant, who completed the task the fastest and accurately identified the POI and all letters on the fourth line of the eye chart, completed the scored task in 209 s, approximately 1.87 standard deviations below the mean time (Fig. 5-1).

5.3 Eye Chart Identification

During the scored task, participants' first objective was to move to the green area near the eye chart (No. 2 in Fig. 4-2) using the Navigation Mode, then switch to Nudge Control to find the eye chart and identify the smallest line of letters they could read. All participants successfully found and identified the eye chart. Participants were able to read between lines 2 and 6 of the eye chart (Appendix G), with a median of line 4. *Post-hoc* analysis found participants' PTSOT scores were positively correlated with their time to find the eye chart using Nudge Control (Pearson, $r = .545$, $p = .044$, $N = 14$). A lower



(a) Image from a task in which the participant successfully identified line 4 of the eye chart.



(b) Image from a task in which the participant successfully identified line 6 of the eye chart.

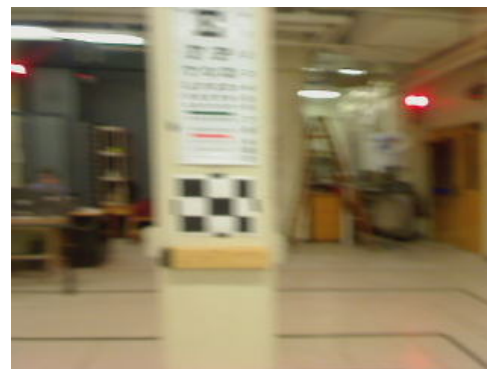
Figure 5-2: Images which participants saw while trying to read a line of the eye chart.

PTSOT score is better, so participants with superior spatial orientation abilities were able to find the eye chart faster. Example images from participants' flights are shown in Fig. 5-2. As a comparison, a person with 20/20 vision could read line 4 from 30 ft away, although this number is not directly applicable because the imagery shown to the participant was degraded by a variety of factors including the webcam lens, focus, image resolution, and jpeg compression.

All of the letters of each subject's lowest line were accurately identified by 64% of participants, with 31% correctly reading 75% of the letters on the line they identified. One participant (5%) only identified 50% of the letters correctly. On average, participants spent a total of 136.5 s (sd 40.3 s) using Nudge Control to find and identify the eye chart, with an average of 71.5 s (sd 32.5 s) spent to searching for the eye chart. After seeing the eye chart, participants spent, on average, 57.5 s (sd 31.4 s) positioning the MAV for the best view and trying to identify the lowest readable line of the eye chart.



(a) Slight blurring resulting from the motion of the MAV.



(b) Severe blurring as a result of the motion of the MAV.

Figure 5-3: Examples of blurred imagery seen by participants while trying to identify the eye chart.

Although participants were successful at identifying a line of the eye chart, it was not without difficulty. While hovering, the MAV is not perfectly still, but constantly compensating for drift and atmospheric instabilities. This motion caused the webcam image to blur at times (Fig. 5-3), which often prevented subjects from immediately obtaining clear imagery shown. The line of the eye chart that participants were able to read was negatively correlated with the number of yaw commands issued (Spearman Rho, $\rho = -.586$, $p = .035$, $N = 13$). This correlation indicates that participants who rotated the MAV less were more likely to identify a lower line of the eye chart. The two participants who were best at eye chart identification correctly identified line 6 of the eye chart, although both participants took much longer than other participants to examine the eye chart after it was found (58.5 s and 42.5 s longer than the mean, 1.86 and 1.35 sd above the mean, respectively).

5.4 Person of Interest Identification

Once participants finished reading a line of the eye chart, their next objective was to fly to the yellow area of the map (No. 3 in Fig. 4-2) using the Navigation Mode, then switch to Nudge Control to find the headshot of a POI (Fig. H-1, which was recessed into a box). They examined the POI until they felt they could identify the headshot again after finishing the task. Nearly all of the participants, 13 of 14, successfully found the POI. Of those 13 participants who found the POI, 12 correctly identified the POI as No. 3 from the photo contact sheet shown to them after the experiment (Fig. H-2), with one participant incorrectly choosing No. 2. Using Nudge Control, participants took, on average, 98.1 s (sd 41.2 s) to find and identify the POI. During this time, participants spent an average of 27.7 s (sd 18.2 s) searching for the POI and used, on average, 70.5 s (sd 38.2 s) moving the MAV to obtain better imagery or examine the POI. Example imagery from participants' flights can be seen in Fig. 5-4.

Participants' standard deviations in their yaw commands were positively correlated to their times

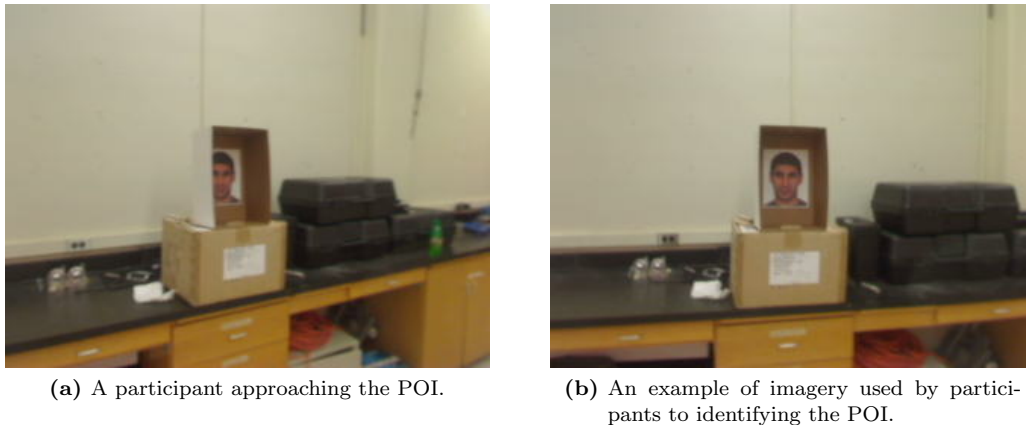


Figure 5-4: Examples of imagery seen by participants while finding and identifying the POI.

to find and identify the POI (Pearson, $r = .625$, $p = .030$, $N = 12$), indicating that participants who had less variance in their yaw movements were more likely to find and identify the POI faster. Three participants tied for being the fastest to find the POI in 10 s, which was 17.7 s faster than the mean time (0.96 sd below the mean), but they had no strategy in common nor did they find the POI from similar locations. The time participants spent finding and identifying the eye chart was negatively correlated with the time spent finding and identifying the POI (Pearson, $r = -.593$, $p = .033$, $N = 13$), indicating a learning effect, i.e., participants who took longer to initially find the eye chart then took less time to find the POI.

5.5 Nudge Control Analysis

All commands issued by a participant while using Nudge Control were logged during the course of the scored task. X and Y commands correspond to moving the MAV left/right and forward/backward (by tilting the iPod) relative to the view from the on-board webcam, while Z commands changed the altitude of the MAV. Participants could also rotate the MAV by issuing a yaw command, which changed the heading of the MAV and the affixed webcam (Fig. 4-1). The mean absolute value of these commands, as well as their standard deviations, were also calculated as descriptive statistics (Table J.3).

To move the MAV sideways, participants issued X commands to the MAV by tilting the iPod left or right. Video game experience was negatively correlated with the absolute mean value of X commands (Spearman Rho, $\rho = -.632$, $p = .021$, $N = 13$), indicating that participants with less video game experience were more likely to move the MAV further sideways (i.e., tilting the iPod farther on average). Participants moved the MAV forwards and backwards, relative to the webcam's view, by issuing Y commands resulting from tilting the iPod forward or backwards. A negative correlation was found between participants' MRT scores and the absolute mean values of Y commands (Pearson, $r = -.605$, $p = .029$, $N = 13$), suggesting that participants who scored better on the MRT were more likely to move the MAV less, either forwards or backwards than average. A negative correlation was also found between participants' scores on the PTSOT and the standard deviations of their Y commands (Pearson, $r = -.558$, $p = .029$, $N = 13$). Participants who scored better on the PTSOT controlled forward and backward movements with more fidelity.

Pinching or stretching gestures on the display of the interface resulted in a Z command, which allowed participants to decrease or increase the altitude of the MAV. A negative correlation was found between the standard deviation of these Z commands and participants' self-reported iPhone experience (Spearman Rho, $\rho = -.649$, $p = .016$, $N = 13$). Users who had iPhone experience were likely to have a smaller variance in the commands, and, therefore, the pinching and stretching gestures they made were more consistent. Finally, the number of Z commands issued by participants was negatively correlated with

the percentage of letters participants correctly identified on the line they read on the eye chart (Pearson, $r = -.590$, $p = .009$, $N = 13$). This shows participants who issued fewer altitude change commands to the MAV were also more likely to correctly read a line on the eye chart.

Participants changed the MAV's heading by swiping a finger left or right across the display, resulting in a yaw command (Fig. 3-17). The absolute mean values of yaw commands positively correlated with the PTSOT scores (Pearson, $r = .597$, $p = .031$, $N = 13$). A lower (better) PTSOT score indicates that a user is more likely to issue smaller, more controlled yaw commands. Additionally, participants' PTSOT scores were positively correlated with the number of yaw commands they issued (Pearson, $r = .686$, $p = .01$, $N = 13$). Participants who issued fewer yaw commands had better performance, which also corresponded to a lower (better) PTSOT score. Therefore a user's PTSOT score predicts that he or she will perform better since fewer yaw commands are issued. Given this correlation and that larger yaw commands, on average, will likely result in overshooting a target, especially when combined with delay from webcam imagery, this study suggests that PTSOT predicts how adept users will be at orienting the MAV.

5.5.1 Webcam Imagery and Frame Rate

The average frame rate of the webcam imagery shown to participants while using Nudge Control was recorded for analysis. On average, participants experienced an average webcam frame rate of 8.38 fps (sd 1.93 fps), with a minimum average 6.98 fps and a maximum average of 13.41 fps. Video analysis of participants' usage of the interface also showed all participants randomly receiving corrupted images every 1-3 seconds due to degraded network conditions. Frame rate was positively correlated with the number of commands issued by the participant while using Nudge Control in the x (Pearson, $r = .636$, $p = .019$, $N = 13$), y (Pearson, $r = .626$, $p = 0.19$, $N = 13$), and z (Pearson, $r = .816$, $p = .001$, $N = 13$) axes. No significance was found between the number of yaw commands issued and the frame rate. As the frame rate increased, which gave participants better imagery, they were also more likely to use Nudge Control to move the MAV more in the x, y, and z axes. While it is generally accepted that video must occur above 20 fps for imagery to not be disorienting [53], participants performed well overall at a frame rate which was less than half of the recommended number for perceptual fusion. All participants identified the eye chart, and twelve of fourteen participants also successfully identified the correct POI with this degraded, low resolution imagery.

5.6 Participants' Navigation Strategies

Participants' waypoint and Nudge Control commands were reconstructed from logged data. This provides insight into strategies used by participants during the scored task. Three participants are examined in

depth: A, who performed the worst, B, who represents an average strategy used by many participants during the experiment, and C, who performed the best. The paths shown in Figs. 5-5, 5-6, and 5-7 outline the participants' flight paths when they used waypoints and Nudge Control. Each participant's path is shown in gray. Navigation mode waypoints are shown as large numbered yellow circles, and Nudge Control movements are shown as smaller red dots. The orientation of the MAV's webcam is shown as a blue arc which, to prevent visual clutter, does not represent the full 60° width of the FOV. The takeoff location is shown as a large black circle in the center of the figures. The location of the scored task POI and eye chart are shown as labeled gray boxes.

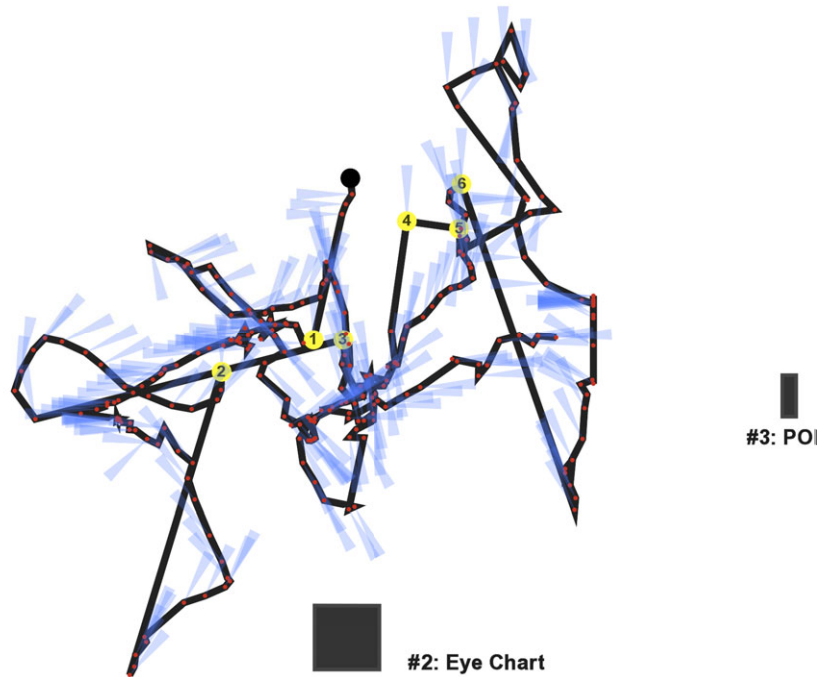


Figure 5-5: Participant A's flight path.

5.6.1 Participant A

Participant A had the worst performance in the experiment, with a time of 373 s (1.34 sd above the mean), six Navigation waypoints, 241 Nudge Control commands, incorrect identification of the POI, and only correctly identified 50% of the letters on line 4 of the eye chart. The participant had a MRT score of 6/20 (1.2 sd below mean) and a PTSOT score 33.1° (1.1 sd below the mean). The flight path of Participant A is shown in Fig. 5-5. Compared to others, the participant did not place waypoints close to the eye chart or the POI, leaving Participant A at a disadvantage to quickly identify either objective.

The participant's poor performance was primarily due to a meandering path and lack of surveying the environment with yaw commands. Video review of Participant A's Nudge Control usage revealed that

the participant's path was largely due to unintentionally tilting the iPod towards her, which continually sent commands to the MAV to move backwards.

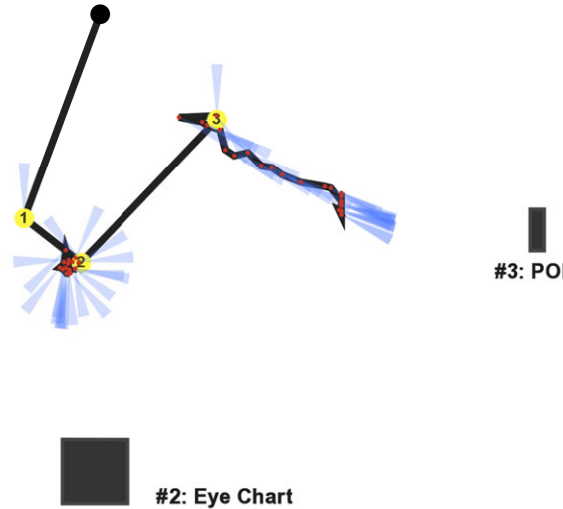


Figure 5-6: Participant B's flight path.

5.6.2 Participant B

Participant B, who represents participants with average performance, took 268.6 s (0.67 sd below the mean) to complete the scored task, using three Navigation waypoints and 45 Nudge Control commands. The participant correctly identified the POI and all letters of line 4 on the eye chart. The participant had a perfect MRT score of 20/20 (1.4 sd above the mean), and a worse than average PTSOT score of 29.0° (0.77 sd below the mean). The participant's flight path is shown in Fig. 5-6.

After reaching the vicinity of the eye chart using Navigation mode, the participant rotated the MAV in an effort to find the eye chart. While surveying for the eye chart, the participant found the POI, which allowed him to easily find the POI in the second half of the task. After increasing the altitude of the MAV and identifying the eye chart, the participant used Navigation mode to send the MAV to the vicinity of the eye chart. Once in the vicinity, the participant quickly rotated the MAV to point directly at the POI, and then used Nudge Control to move closer to better identify the POI. Many other participants who were average performers in the study executed a similar strategy to find both the eye chart and POI.

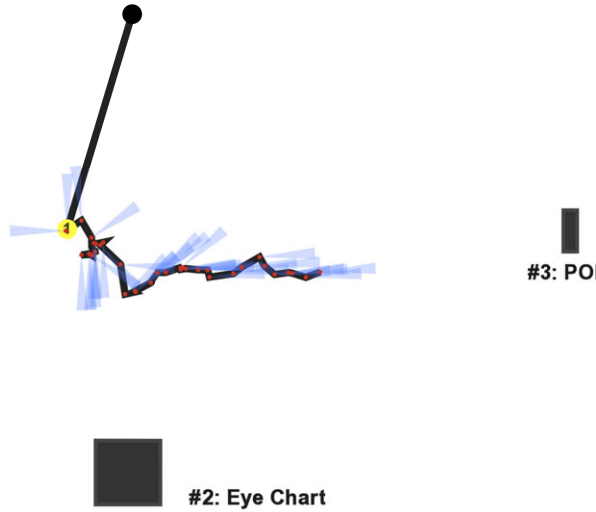


Figure 5-7: Participant C's flight path.

5.6.3 Participant C

Participant C performed the best overall by being the fastest participant to accurately complete the scored task in 209.44 s (1.79 sd below average). Participant C used one Navigation waypoint and 35 Nudge Control commands to complete the task, correctly identifying the POI and line 4 of the eye chart. Participant C had an average MRT score of 12/20 (0.1 sd below the mean) and an above average PTSOT score of 12.75° (0.58 sd above the mean). Participant C's flight path is shown in Fig. 5-7. Although the participant's flight started with an average strategy of using the Navigation mode to move to the vicinity of the eye chart, several key differences began to emerge after that point. While surveying the environment for the eye chart, Participant C only executed yaw commands after it was clear from the webcam imagery that the MAV had finished rotating. This patient approach was very different from many other participants who transmitted further yaw commands to the MAV while they were still receiving imagery of the MAV rotating. Also, unlike many other participants, Participant C was quick to build SA of the environment and infer that the MAV's altitude had to be increased in order to view the eye chart. While finding the eye chart, the participant noticed the POI, typical of average performers.

However, after identifying the eye chart, Participant C continued to use Nudge Control to re-orient the MAV and find the POI, which cleverly used the participant's existing SA to help find the POI quickly. Other participants, even if they noticed the eye chart, used Navigation mode to immediately move the MAV to the vicinity of the POI as instructed, which caused many of them to lose SA. Losing

SA required these other participants to begin surveying the environment again in order to find the POI. After finding the POI from the vicinity of the eye chart, Participant C did not follow the experimenter’s verbal instructions to use Navigation mode to move to the vicinity of the eye chart and solely used Nudge Control to approach and identify the POI, which turned out to be a more effective strategy. Participant C’s overall time was better than the 2nd fastest participant. Participant C’s strategy was not accidental, as the participant was the only one in the post-experiment survey to rank her confidence in her actions as “absolutely confident.”

5.7 Subjective Responses

After completing the tasks, participants answered a usability survey (Appendix D) and were interviewed to gain general feedback on the interface. The questions used to guide the interview are listed in Appendix I. Responses about the usability of the interface are discussed in detail here. Participants generally felt confident about their performance using MAV-VUE, with 43% reporting they were confident about the actions they took, and 50% felt very confident about their actions. The complete set of participant answers is listed in Appendix D.

5.7.1 Navigation Mode

Participants found the Navigation Mode, consisting of the map and waypoints display, easy to use. A third (36%) felt very comfortable using waypoints, and 43% were comfortable using waypoints. All participants felt they understood adding a waypoint and using the webcam view very well. In the map display, 92% of participants rated that they understood the location of the MAV very well, with 79% understanding the orientation of the MAV very well. The MAV’s direction of travel (the velocity vector in Fig. 3-3) was understood very well by 86% of participants. Twelve participants wrote comments on the survey indicating that they found the Navigation mode easy to use. When questioned about the usefulness of the Navigation Mode, one participant made an insightful statement about the map display’s inset webcam view (only eight participants opened the inset webcam view): “As soon as I brought up the webcam view in the map, I immediately thought ‘wait, I want to be in the manual [Nudge] Control anyways now.” This response is consistent with other participants’ actions during the experiment, which showed a preference for using Nudge Control while viewing imagery from the webcam.

5.7.2 Nudge Control

When asked about aspects of the interface they found confusing or hard to use and what they found easy to use, participants had conflicting responses on a variety of topics. Four participants stated they found Nudge Control difficult due to the time lag between issuing commands and receiving webcam

imagery back from the MAV. Other participants completely disassociated the delay in feedback, writing that they found Nudge Control easy to use, but felt that the MAV ignored their commands or did something different. Seven participants had positive feedback on Nudge Control, repeatedly expressing the same sentiments that Nudge Control was “easy,” “straight-forward,” or “very intuitive.” However, every subject mentioned the time lag in their feedback. When further questioned about the time delay, several participants felt the delay was more annoying than an actual impediment to them being able to interact with the MAV.

Five participants mentioned that they found changing the heading of the MAV while using Nudge Control to be difficult. To rotate the MAV, a user is supposed to swipe a finger either left or right across the entire display (Fig. 5-8a). This gesture was adopted after pilot subjects found tapping on the circumference of the circle in Nudge Controls to be unintuitive. Five participants tried to rotate the MAV by moving their finger along the circumference of the constraint circle (Fig. 5-8b), despite explicit instructions by the demonstrator that this was the not the way the interface was intended to be used.

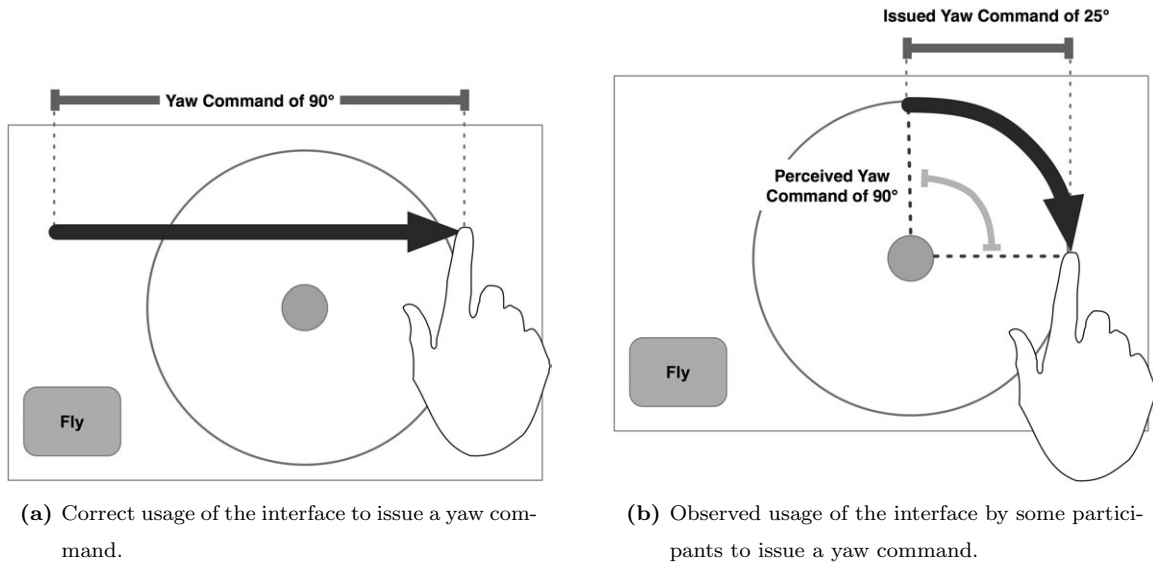


Figure 5-8: Intended and observed interaction to issue a Nudge Control yaw command.

Clearly, the visual feedback of the blue arc rotating within the circle as the user changed the heading was a very salient signal which prompted users to make a circular gesture on the constraint circle. Originally, an interaction similar to Fig. 5-8b was implemented, but was changed to Fig. 5-8a due to feedback from pilot subjects.

5.8 Experiment Observations

Upon reviewing video tape of participants during the study, several other trends in usage of the hand-held display and interface became apparent. Two of the most important findings which were not evident from other sources was the participant rest pose when using Nudge Control and usage of the Fly button. When using Nudge Control, it was observed that many participants' natural postures for holding the iPod was to have it tilted slightly towards them (Fig. 5-9b) instead of the intended horizontal orientation (Fig. 5-9a).

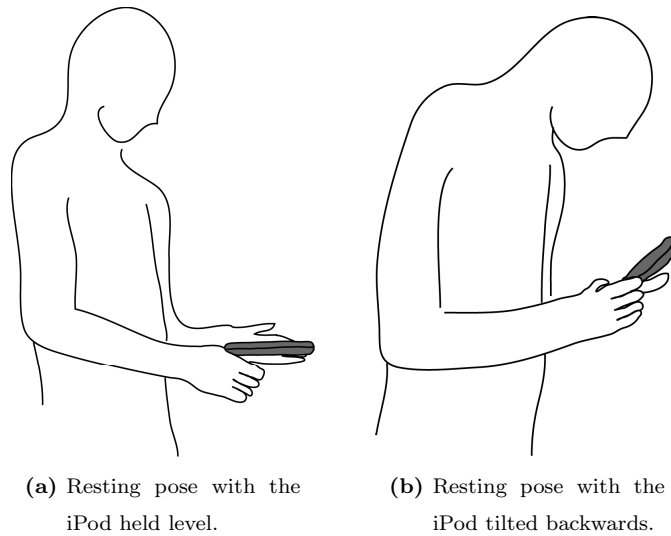


Figure 5-9: Resting poses observed while participants used the interface, courtesy of Nadya Peek.

This appeared to be partly due to the participants instinctively finding a viewing angle which minimized glare, as well as the need for an ergonomically comfortable pose. However, this tilted rest pose corresponds to a command to move the MAV backwards since the neutral position was to have the device almost level (small tilt values within a few degrees of zero were filtered out as being neutral). Unfortunately, for many participants the angle of their pose was subtle enough that they did not realize they were commanding the MAV to move backwards, and the MAV would slowly creep backwards as they focused on the identification tasks.

Video footage also revealed five subjects who tried to command the MAV in Nudge Control without the Fly button depressed. The interface ignores all Nudge Control commands unless the Fly button is continually pressed, in an effort to prevent inadvertent or accidental input. While nearly all of the participants immediately recovered from this error after one or two attempts to move or rotate the MAV, two participants continually forgot to engage the Fly button during the experiment.

Finally, an unexpected interaction occurred when participants became focused on the webcam imagery.

Rather than disengaging Nudge Control and raising the iPod up to their face to look closer at the display, participants would keep the iPod flat and level (even if Nudge Controls were disengaged) and hunch over the display. This action raises interesting questions about the potential loss of SA and peripheral cues, even when using a hand held device.

5.9 Summary

This chapter presented the results of the usability study described in Chapter 4. Participants' performance at achieving search and identification objectives, was examined, along with other qualitative and quantitative metrics. Participants' subjective responses were also presented, along with demographic and observational information. Together, these results describe a clear picture of participants' ability to use MAV-VUE for performing a search and identification task.

Overall, participants performed extremely well for the given objectives. Furthermore, the MAV never crashed or had a collision as a result of the actions taken by participants. While some participants had difficulty using the interface, others took near-ideal actions for exploring an unknown environment. Nearly all of the participants found and correctly identified the POI after only three minutes of hands-on training. Additionally, all of the participants found the eyechart and were able to successfully identify a line. These tasks were completed in a realistic environment, with an appreciable time delay between the MAV sending imagery and receiving commands, typical of environments which make traditional teleoperation very difficult to perform well.

Many areas for improvement were identified in the subjective responses. However, participants also indicated they generally found the interface intuitive and easy to use. In addition, participants were generally confident of their actions, even given their limited training. While the MRT proved to be of limited usefulness as a performance predictor, the PTSOT had better predictive power, especially in correlating Nudge Control yaw command (an indicator of performance). Several significant correlations were found between participants' usage of Nudge Control and various aspects of their task performance. Although there were no definitive trends, it is clear that participants' usage of yaw commands to change the heading of MAV were related to their ability to quickly find and identify the eye chart and POI.

Examining the flight paths of participants provided valuable information on participants' strategies and SA while performing the scored task. While some participants clearly suffered from usability issues in the interface, other participants were able to overcome the same usability issues, suggesting the issues were not critical. Average performers shared a common strategy of thoroughly surveying the environment with Nudge Control to build SA and efficiently find and identify the eye chart and POI. Top performers were especially effective at continually using Nudge Control to build and maintain SA of an environment.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 6

Conclusions

Even with the availability of satellite imagery, many shortcomings of this technology prevent it from being a complete solution in helping field personnel such as soldiers, SWAT teams and first responders to construct an accurate mental model of their environment. Collaboratively exploring a hostile environment with an autonomous MAV has many attractive advantages which can help solve this problem. Field personnel are potentially kept out of harm, while the MAV can navigate difficult terrain and environments which may otherwise be inaccessible. Autonomous MAVs are just becoming commercially available to the public. Unfortunately, current interfaces for MAVs ignore the needs of an operator in a hostile setting. These interfaces require the full, undivided attention of the operator, as well as physically requiring the operator to be completely engaged with a laptop or similar device. Solutions to aid the operator, such as video goggles, only further prevent the operator, who likely has many tasks beyond operating the MAV from maintaining an awareness of their surroundings.

At the same time that MAVs have begun to enter the commercial sector, hand held devices have made a resurgence in the form of powerful mobile computing platforms. These sophisticated devices have high fidelity displays, support multiple interaction modes (e.g., touch, tilting, voice), and can communicate with the outside world through a variety of wireless mechanisms. Unfortunately, HRI research on interacting with a robot, such as a MAV, has largely been limited to teleoperation interfaces and PDAs. While teleoperation has been studied for many years, the real-world constraints of time delays, and an operator's cognitive workload, controlling a robot at a distance through order reduction of the associated feedback control have required extensive training to be successfully used.

Combined, these factors demonstrate a clear need for a way to allow field personnel to collaboratively explore an unknown environment with a MAV, without requiring the operator's continual attention, additional bulky equipment, and specialized training. MAV-VUE is an interface which satisfies these

demands without sacrificing operator performance. Central to MAV-VUE is the invention of PFO Control, which allows an operator with minimal training to safely and precisely performed fine-tuned control of a MAV without the traditional human control problems found in teleoperation interfaces. Finally, to the best knowledge of the author, this is the first time a formal study has examined using an HRI interface to control and work with a MAV in the real world, and not a simulated environment and vehicle.

6.1 Research Objectives and Findings

The objective of this research was to develop a mobile interface for interacting with an autonomous MAV to explore outdoor environments. This goal was addressed through the following research objectives:

- Objective 1. Determine the functional and information requirements for a MAV interface (Chapter 3).
- Objective 2. Develop a mobile interface which allows an operator to explore an unknown environment with a MAV (Chapter 3).
- Objective 3. Evaluate the usability of the interface (Chapters 4 and 5).

The CTA performed in Chapter 3 provides insight into envisioned operators' needs, roles, and uses of the interface. These were integrated into functional and information requirements for the interface. The resulting design of the interface was discussed in Chapter 3, along with the technical theory of PFO Control and its possible implementations. Chapter 4 described the setup and procedure of a study to evaluate the usability of the interface in a real-world setting with realistic constraints. This study sought to determine whether if a casual user could effectively use MAV-VUE to control a MAV and if the user's SA of the environment was improved by information presented in MAV-VUE (Section 4.2).

The results of this study unambiguously demonstrate the feasibility of a casual user controlling a MAV with MAV-VUE to perform search and identify tasks an unknown environment. With only three minutes of training, all participants successfully found and were able to read a line from an eye chart. Participants could clearly manipulate the position and orientation of the MAV to obtain information about the environment. This demonstrates the suitability of using this type of interface for performing detailed surveying tasks, such as structural inspections. Participants were also able to construct an accurate mental model of the environment through both the provided map and webcam imagery. Twelve of fourteen participants found and accurately identified a headshot of a POI, showing that this interface has real-world applications for ISR missions performed by soldiers and police SWAT teams. Equally important to the participants' success, the MAV never crashed or had a collision due to participants' actions. Several statistically significant correlations were found between usage of Nudge Control and participants' performance. PTSOT scores also correlated to participant performance, suggesting that this test can be used as a predictor of participant's performance with the interface.

6.2 Future Work

Even though participants were very successful in achieving their objectives during the usability study, the study also revealed many areas of further investigation to improve the interface. The following are recommendations for future follow-on work based on the research presented in this thesis.

- Improve the interface to support displaying 3D models of an environment generated by a MAV through stereoscopic computer vision, LIDAR, or other means.
- Utilize more advanced autonomy to have a MAV support more of the functionality listed in Table 3.2.
- Investigate the gains used in PFO Control, specifically the function used and the potential of gains which adapt to the operator's expertise and usage over time.
- In PFO Control, if the transmission delay, (τ_4 in Fig. 3-11) is significant, or transmissions are infrequent, it may be advantageous to incorporate a time stamp of the position to allow the constraint filter to predict the current position of the robot when it issues a movement command.
- Investigate the effects on users' performance when using Nudge Control's NG mode as compared to using the CT mode.
- Address the usability issues found in Section 5.7, specifically how the user could better control the heading of the MAV, and how to resolve the ambiguity of a user's resting pose.
- Conduct field testing with envisioned users, such as soldiers, police SWAT teams, and first responders.

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix A

Cognitive Task Analysis Interview Questions

1. What information do you need when exploring an unknown environment?
2. How do you currently get the mapping information you need? (and is it accurate?)
3. How you build a map of a building you do not know?
4. How do you find and determine entries/exits to a building?
5. Describe how you conduct an ISR (or equivalent) mission?
6. How can you envision using a MAV?
7. What is a reasonable size for a MAV?
8. What sensor packages would be useful on a MAV?
9. How long would you expect the MAV to last? (battery life)

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix B

Demographic Survey and Statistics

MAV-VUE Usability Study Pre-experiment Survey

1. Subject ID: _____

2. Age: _____

3. Gender: M F

4. Occupation: _____

if student, (circle one): Undergrad Masters PhD
year of graduation: _____

5. Military experience (circle one): No Yes

If yes, which branch: _____
Years of service: _____

6. Give an overall rating of your past two nights of sleep. (circle one)

Poor Fair Good Great

7. How much experience do you have with video games? (circle one)

Never play games Play games once a month Weekly gamer Frequent gamer
Extreme gamer

Types of games played:

8. How much experience do you have with RC helicopters/airplanes or unmanned vehicles? (circle one)

Never Used Previously Used Used Monthly Used Weekly

9. Comfort level with Google Maps? (1 is little, 5 is most comfortable)

1 2 3 4 5

10. Comfort level with Google Earth? (1 is little, 5 is most comfortable)

1 2 3 4 5

11. Have you used an iPhone, iPod Touch, or other touch-based device before? (circle one) Yes
No

a. If Yes, what is your comfort level with using one? (1 is little, 5 is most comfortable)

1 2 3 4 5

12. Are you far-sighted/unable to see read text as it moves closer to you? (circle one) Yes
No

a. If Yes, are you currently wearing corrective contacts or glasses? (circle one) Yes
No

13. Are you red/green color blind? (circle one) Yes No

14. What applications do you think are appropriate for Unmanned Aerial Vehicles (UAVs)?

Table B.1: Responses for self-assessed sleep quality

	Poor	Fair	Good	Great
Give an overall rating of your past two nights of sleep	1	6	4	3

Table B.2: Responses for self-assessed video game experience

	Never play games	Play games once a month	Weekly gamer	Frequent gamer	Extreme gamer
How much experience do you have with video games?	4	6	1	3	0

Table B.3: Responses for self-assessed RC vehicle experience

	Never Used	Previously Used	Used Monthly	Used Weekly
How much experience do you have with RC helicopters/airplanes or unmanned vehicles?	5	9	0	0

Table B.4: Responses for self-assessed comfort with Google Maps™, Google Earth™, and iPhone

	Never Used	1	2	3	4	5
Comfort level with Google Maps? (1 is little, 5 is most comfortable)	0	0	0	2	7	5
Comfort level with Google Earth? (1 is little, 5 is most comfortable)	1	2	1	3	6	1
Comfort level with iPhone? (1 is little, 5 is most comfortable)	0	1	0	3	5	5

Appendix C

Configuration of the MAV

The following gains were used in the RAVEN flight control software:

$$Kp_{xyvel} = 0.16$$

$$Ki_{xyvel} = 0.06$$

$$Kp_{roll} = 1.0$$

$$Kp_{pitch} = 1.0$$

$$Kp_{yaw} = 1.1$$

$$Ki_{yaw} = 0.8$$

$$Kp_{throttle} = 0.7$$

$$Ki_{throttle} = 0.1$$

$$Kd_{throttle} = 0.4$$

$$ctl- > trim_{throttle} = -0.1$$

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix D

Usability Survey

MAV-VUE Usability Study Survey

1. Subject ID: _____
2. Eye Chart Results (filled out by PI)
 Line read: _____
 Letters: _____
3. Using the photo sheet provided, please identify which mugshot you saw during the final task?
 # _____
4. Using the application: (check one box per row)

	No Confidence	Minimal Confidence	Somewhat Confident	Mostly Confident	Absolutely Confident
How confident were you about the actions you took?					

5. How well did you feel you performed on the following aspects of the task? (check only one box per row)

	Very Poor	Poor	Satisfactory	Good	Excellent
Controlling the MAV by using waypoints					
Identifying lines in the eye chart					
Controlling the MAV using manual flight controls					
Identifying Persons of Interest					

6. Please indicate how well you understood each of the following parts of the application: (check only one box per row)

	Poorly Understood	Somewhat Understood	Well Understood	Did Not Use
Adding a Waypoint				
Webcam View				
Manual Flight Controls - Tilting				
Manual Flight Controls – Rotating Helicopter				
Manual Flight Controls – Changing Altitude				
MAV Orientation				
MAV Location				
MAV Direction of Travel				

7. Were there any aspects of the application which you found confusing or hard to use? If so, how were they confusing/hard to use and what did you expect?

8. What aspects of the interface did you find easy to use and why?

Table D.1: Responses for self-assessed confidence in actions during scored task.

	No Confi- dence	Minimal Confidence	Somewhat Confident	Mostly Confident	Absolutely Confident
How confident were you about the actions you took?	0	0	6	7	1

Table D.2: Responses for self-assessed performance during scored task.

	Very Poor	Poor	Satisfactory	Good	Excellent
Identifying lines in the eye chart	0	0	3	5	6
Controlling the MAV by using waypoints	0	4	8	2	0
Controlling the MAV using manual flight controls	0	1	9	4	0
Identifying Persons of Interest	1	0	4	4	5

Table D.3: Responses for self-assessed understanding of interface.

	Poorly Under- stood	Somewhat Understood	Well Under- stood	Did Not Use
Adding a Waypoint	0	0	14	0
Webcam View	0	0	13	1
Manual Flight Controls - Tilting	1	2	11	0
Manual Flight Controls - Rotating Helicopter	0	3	11	0
Manual Flight Controls - Changing Altitude	0	4	10	0
MAV Orientation	0	2	11	1
MAV Location	0	1	13	0
MAV Direction of Travel	0	1	12	1

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix E

Consent to Participate Form

CONSENT TO PARTICIPATE IN NON-BIOMEDICAL RESEARCH

MAV-VUE Usability Study

You are asked to participate in a research study conducted by Dr. Mary L. Cummings, from the Aeronautics/Astronautics Department at the Massachusetts Institute of Technology (M.I.T.) You were selected as a possible participant in this study because the population this research will influence is expected to contain men and women between the ages of 18 and 55 with an interest in using computers. You should read the information below, and ask questions about anything you do not understand, before deciding whether or not to participate.

• PARTICIPATION AND WITHDRAWAL

Your participation in this study is completely voluntary and you are free to choose whether to be in it or not. If you choose to be in this study, you may subsequently withdraw from it at any time without penalty or consequences of any kind. The investigator may withdraw you from this research if circumstances arise which warrant doing so.

You may be withdrawn from the research if your vision is worse than 20/25 when corrected with glasses or contacts

• PURPOSE OF THE STUDY

The purpose of this study is to evaluate how easy it is to use an application on a small mobile device, such as an iPhone®, for exploring an unknown environment with an autonomous vehicle as well as learning more about that environment in real-time.

• PROCEDURES

If you volunteer to participate in this study, we would ask you to do the following things:

Spatial Reasoning Testing

You will be asked to take two pen and paper tests to determine your ability to orient yourself on a map and rotate objects in space.

Application Training

Participate in a 10 min practice trial using the application. You will be taught how the application works and have an opportunity to try using it yourself.

Experiment

Participate in a 10 min mission of exploring an urban area with an autonomous micro aerial vehicle (MAV), using the application. You will be given a map of the area which

may or may not be accurate. Your main goal will be to send the MAV to a building and ID suspects inside the building from the MAV's webcam. You will also be asked to visit Areas of Interest (AOI) and identify any suspicious objects which may or may not be present. You will remotely control the MAV and at no time be in the same room with it.

You will be assigned a score for the mission based on several factors. How well you identify people and objects, as well as being penalized for misidentifying, will be one part of the score. Another part will be how fast you complete the mission and how efficient is the path you command the MAV take during the entire mission.

All testing will take place at MIT in room 35-220

Total Time: 1 hour and 15 mins

- **POTENTIAL RISKS AND DISCOMFORTS**

There are no anticipated physical or psychological risks

- **POTENTIAL BENEFITS**

There are no potential benefits you may receive from participating in this study

This study will assist in the design of better human/unmanned vehicle systems

- **PAYMENT FOR PARTICIPATION**

You will be paid \$15 for your participation in this study, which will be paid upon completion of your debrief. Should you elect to withdraw during the study, you will be compensated for your time spent in the study. The subject with the best performance will be given a reward of a \$100 Best Buy Gift Card

- **CONFIDENTIALITY**

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law. You will be assigned a subject ID which will be used on all related documents to include databases, summaries of results, etc.

You consent to be audio/videotaped during the experiment. You will have the right to review and edit the video data. Only the study personnel will have access to the tapes, which will be erased 90 days after the analysis of the study is completed.

- **IDENTIFICATION OF INVESTIGATORS**

If you have any questions or concerns about the research, please feel free to contact

Principle Investigator

Mary Cummings
77 Massachusetts Ave.
33-311
Cambridge, MA 02139
ph. 617-252-1512

Student Investigator

David Pitman
77 Massachusetts Ave.
35-220
Cambridge, MA 02139
ph. 617-253-0993

- **EMERGENCY CARE AND COMPENSATION FOR INJURY**

If you feel you have suffered an injury, which may include emotional trauma, as a result of participating in this study, please contact the person in charge of the study as soon as possible.

In the event you suffer such an injury, M.I.T. may provide itself, or arrange for the provision of, emergency transport or medical treatment, including emergency treatment and follow-up care, as needed, or reimbursement for such medical services. M.I.T. does not provide any other form of compensation for injury. In any case, neither the offer to provide medical assistance, nor the actual provision of medical services shall be considered an admission of fault or acceptance of liability. Questions regarding this policy may be directed to MIT's Insurance Office, (617) 253-2823. Your insurance carrier may be billed for the cost of emergency transport or medical treatment, if such services are determined not to be directly related to your participation in this study.

- **RIGHTS OF RESEARCH SUBJECTS**

You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you feel you have been treated unfairly, or you have questions regarding your rights as a research subject, you may contact the Chairman of the Committee on the Use of Humans as Experimental Subjects, M.I.T., Room E25-143B, 77 Massachusetts Ave, Cambridge, MA 02139, phone 1-617-253 6787.

SIGNATURE OF RESEARCH SUBJECT OR LEGAL REPRESENTATIVE
--

I understand the procedures described above. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been given a copy of this form.

Name of Subject

Name of Legal Representative (if applicable)

Signature of Subject or Legal Representative

Date

SIGNATURE OF INVESTIGATOR

In my judgment the subject is voluntarily and knowingly giving informed consent and possesses the legal capacity to give informed consent to participate in this research study.

Signature of Investigator

Date

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix F

Maps

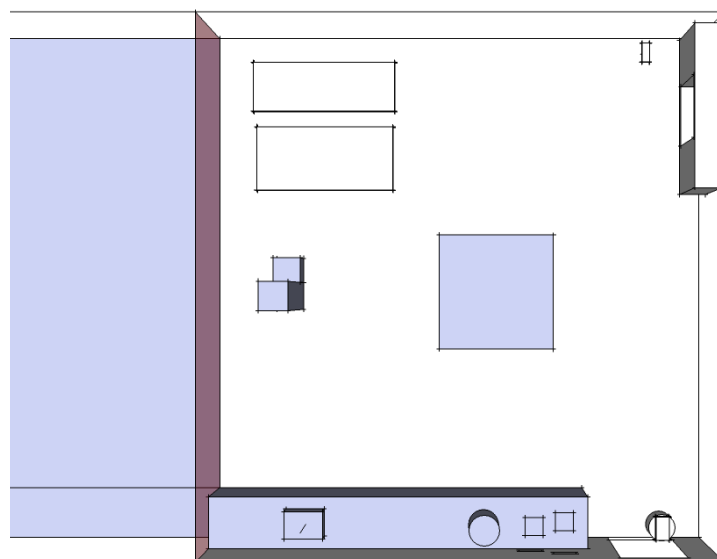


Figure F-1: The map displayed on the subject's iPod.

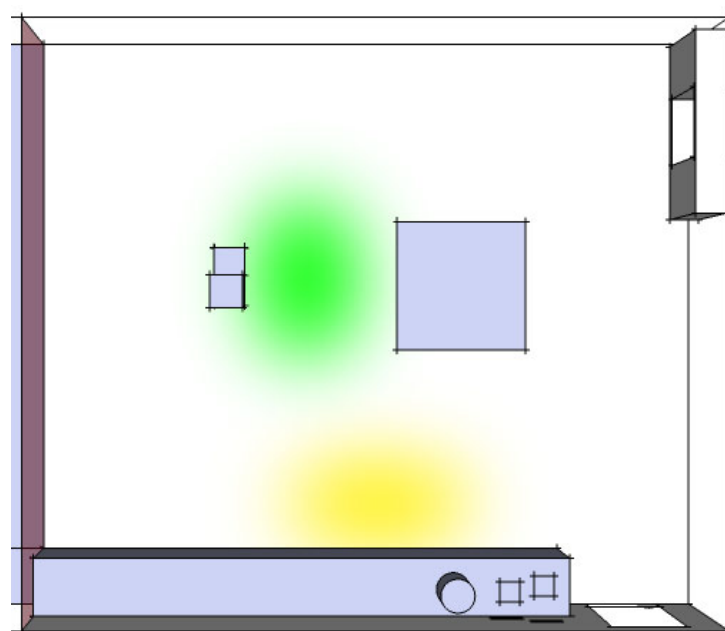


Figure F-2: Supplementary map given to subject for their scored task.

Appendix G

Snellen Eye Chart

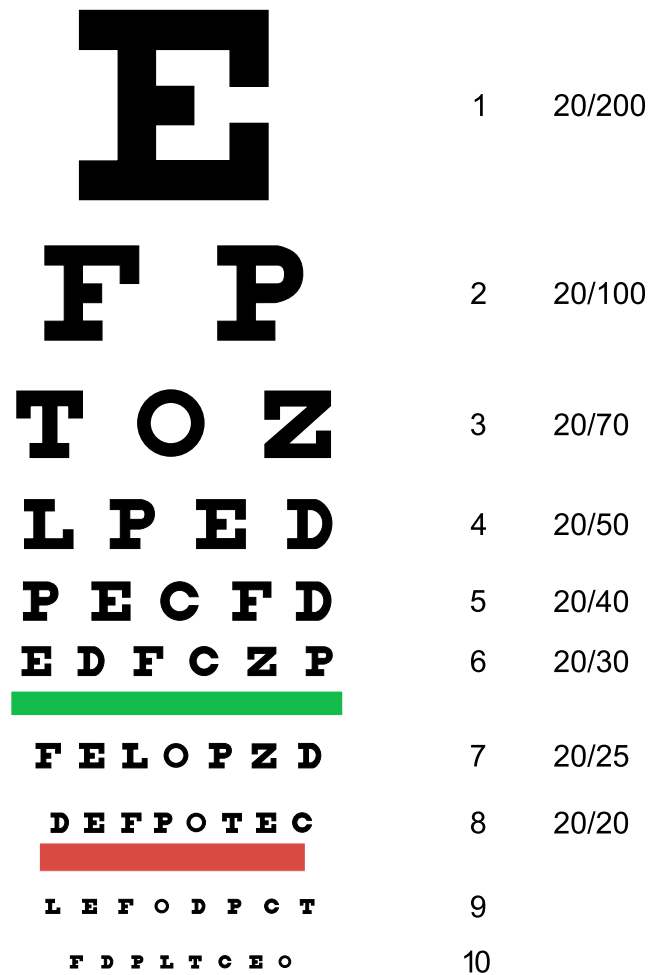


Figure G-1: Example of the standard Snellen eye chart used in the study, by Jeff Dahl. Licensed under the Creative Commons Attribution-Share Alike 3.0 Unported License.

Appendix H

Person of Interest Used and Identification Sheet



Figure H-1: POI used in the scored task of the usability study.



1



2



3



4



5



6



7



8

Figure H-2: POI identification sheet used in usability study.

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix I

Debriefing Interview Questions

The following questions were asked by the experimenter at the end of the experiment:

1. Do you feel like you could effectively control the MAV?
2. Were there any unexpected events (either with the interface or MAV) which occurred during the experiment?
3. Do you feel like you had good general control over the MAV?
4. Do you feel you could fly the MAV between two close walls (i.e., a narrow alleyway) using the nudge controls?
5. Did you find the interface easy or hard to use? In what ways?
6. Do you have any other feedback?

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix J

Scored Task Descriptive Statistics

Table J.1: Performance Descriptive Statistics.

	N	Mean	Median	Mode	Std. Dev.	Min.	Max.
MRT	14	12.57	12.50	20	5.27	3	20
PTSOT	14	19.74	13.54	11.50	12.03	9.00	42.67
Eye Chart Line	14	4	4	4	-	2	6
Eye Chart Line: % correct	14	89.29%	100%	100%	16.1	50%	100%
Framerate	13	8.38	7.74	6.98	1.93	6.98	13.41
Path Distance	13	13.00	10.57	5.26	10.73	5.26	47.17
Num. Waypoints	13	3.23	3	2	1.59	1	6
Num. Nudge Control Com- mands	13	62.92	45	45	56.47	28	241

Table J.2: Times Descriptive Statistics.

	N	Mean	Median	Mode	Std. Dev.	Min.	Max.
Total Time	13	303.80	308.00	209.44	52.76	209.44	374.91
Eye Chart: Total Time	13	141.5	136.5	71.00	40.3	71.0	213.0
Eye Chart: Time to Find	13	84.0	71.5	36.0	32.5	36.0	150.0
Eye Chart: Time Identifying	13	57.5	59.5	22.0	31.4	21.0	116.0
POI: Total Time	12	98.2	81.0	74.0	41.3	41.0	183.0
POI: Time to Find	12	27.7	27.0	10.0	18.2	10.0	73.0
POI: Time Identifying	12	70.5	58.0	71.0	38.2	33.0	150.0

Table J.3: Descriptive statistics of Nudge Control commands performed by participants during the scored task.

	N	Mean	Median	Mode	Std. Dev.	Min.	Max.
Num. X Commands	13	23.92	15	2	30.56	2	119
Num. Y Commands	13	15.53	5	0	27.11	0	97
Num. Z Commands	13	10.15	7	5	13.37	3	54
Num. Yaw Commands	13	18.62	16	8	10.60	8	44
Mean of X Commands	13	0.138	0.141	0.090	0.023	0.090	0.170
Mean of Y Commands	13	0.061	0.076	0.000	0.035	0.130	0.110
Mean of Z Commands	13	0.227	0.239	0.130	0.055	0.130	0.300
Mean of Yaw Commands	13	3.340	3.570	1.310	0.985	1.310	4.370
Std. Dev. of X Commands	13	0.047	0.053	0.010	0.021	0.010	0.090
Std. Dev. of Y Commands	13	0.035	0.037	0.000	0.028	0.000	0.090
Std. Dev. of Z Commands	13	0.086	0.750	0.040	0.032	0.040	0.160
Std. Dev. of Yaw Commands	13	2.393	2.412	1.920	0.297	1.920	2.780

References

- [1] J. M. McMichael and M. S. Francis, “Micro Air Vehicles - Toward a New Dimension in Flight,” tech. rep., DARPA, Nov 1997.
- [2] G. de Croon, K. de Clercq, and R. Ruijsink, “Design, aerodynamics, and vision-based control of the DelFly,” *International Journal on Micro Air Vehicles*, vol. 1, pp. 71–97, Jan 2009.
- [3] W. Keller and D. Jones, “Developing the Class I Unmanned Aerial System (UAS),” *Army AL&T*, pp. 30–33, 2008.
- [4] J. How, B. Bethke, A. Frank, D. Dale, and J. Vian, “Real-time indoor autonomous vehicle test environment,” *IEEE Control Systems*, vol. 28, pp. 51–64, Jan 2008.
- [5] National Resource Council, “Tactical Display for Soldiers Human Factors Considerations,” tech. rep., National Academy of Science, 1997.
- [6] T. Sheridan, *Humans and Automation: System Design and Research Issues*. New York, New York, USA: John Wiley & Sons, Inc., 2002.
- [7] T. Sheridan, *Telerobotics, Automation, and Human Supervisory Control*. Cambridge, Massachusetts USA: MIT Press, Jan 1992.
- [8] T. Sheridan and W. Verplank, “Human and Computer Control of Undersea Teleoperators,” *Office of Naval Research*, Jan 1978.
- [9] R. Parasuraman, T. Sheridan, and C. D. Wickens, “A Model for Types and Levels of Human Interaction with Automation,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 30, pp. 286–297, Jan 2000.
- [10] T. Sheridan and R. Parasuraman, “Human versus automation in responding to failures: An expected-value analysis,” *Human Factors*, vol. 42, pp. 403–407, Jan 2000.
- [11] M. Cummings, S. Bruni, S. Mercier, and P. Mitchell, “Automation Architecture for Single Operator, Multiple UAV Command and Control,” *The International C2 Journal*, vol. 1, Jan 2007.
- [12] J. Scholtz, “Theory and Evaluation of Human Robot Interactions,” in *HICSS ’03: Proceedings of the 36th Annual Hawaii International Conference on System Sciences (HICSS’03) - Track 5*, (Washington, DC, USA), p. 125.1, IEEE Computer Society, 2003.
- [13] M. Goodrich and A. Schultz, “Human-Robot Interaction: A Survey,” *Foundations and Trends in Human-Computer Interaction*, vol. 1, no. 3, pp. 203–275, 2007.
- [14] M. Kadous, R. Sheh, and C. Sammut, “Effective user interface design for rescue robotics,” in *HRI ’06: Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, (New York, NY, USA), pp. 250–257, ACM, 2006.
- [15] H. A. Yanco, B. Keyes, J. L. Drury, C. W. Nielsen, D. A. Few, and D. J. Bruemmer, “Evolving interface design for robot search tasks: Research Articles,” *Journal of Field Robotics*, vol. 24, no. 8–9, pp. 779–799, 2007.

- [16] M. Baker, R. Casey, B. Keyes, and H. Yanco, "Improved interfaces for human-robot interaction in urban search and rescue," in *in Proceedings of the IEEE Conference on Systems, Man and Cybernetics*, pp. 2960–2965, 2004.
- [17] H. Yanco and J. Drury, "Rescuing interfaces: A multi-year study of human-robot interaction at the AAAI Robot Rescue Competition," *Autonomous Robots*, vol. 22, pp. 333–352, May 2007.
- [18] J. Scholtz, J. Young, J. Drury, H. Yanco, and G. NIST, "Evaluation of human-robot interaction awareness in search and rescue," *Robotics and Automation*, vol. 3, pp. 2327–2332, 2004.
- [19] M. Micire, J. L. Drury, B. Keyes, and H. A. Yanco, "Multi-Touch Interaction for Robot Control," in *IUI '09: Proceedings of the 13th international conference on Intelligent user interfaces*, (New York, NY, USA), pp. 425–428, ACM, 2009.
- [20] T. Sheridan, "Space Teleoperation through Time Delay: Review and Prognosis," *IEEE Transactions on Robotics and Automation*, vol. 9, p. 592, Jan 1993.
- [21] J. Chen, E. Haas, and M. Barnes, "Human Performance Issues and User Interface Design for Teleoperated Robots," *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Review*, vol. 37, pp. 1231–1245, Jan 2007.
- [22] T. Fong, C. Thorpe, and C. Baur, "Collaborative Control: A Robot-Centric Model for Vehicle Teleoperation," *AAAI Technical Report*, vol. 99, Jan 1999.
- [23] K. Jones, E. Schmidlin, and B. Johnson, "Tele-operators' judgments of their ability to drive through apertures," in *HRI '09: Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, (New York, NY, USA), pp. 249–250, ACM, 2009.
- [24] T. Fong, N. Cabrol, C. Thorpe, and C. Baur, "A Personal User Interface for Collaborative Human-Robot Exploration," in *International Symposium on Artificial Intelligence, Robotics, and Automation in Space, Montreal, Canada*, June 2001.
- [25] H. Keskinpala, J. Adams, and K. Kawamura, "PDA-based human-robotic interface," in *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, vol. 4, pp. 3931– 3936, Oct 2003.
- [26] J. Adams and H. Kaymaz-Keskinpala, "Analysis of Perceived Workload when Using a PDA for Mobile Robot Teleoperation," in *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, vol. 4, pp. 4128–4133 Vol.4, 26-May 1, 2004.
- [27] R. Gutierrez and J. Craighead, "A native iPhone packbot OCU," in *HRI '09: Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, (New York, NY, USA), pp. 193–194, ACM, 2009.
- [28] B. J. O'Brien, C. Karan, and S. H. Young, "FOCU:S - Future Operator Control Unit: Soldier," in *Proceedings of SPIE: Unmanned Systems Technology XI* (G. R. Gerhart, D. W. Gage, and C. M. Shoemaker, eds.), vol. 7332, p. 73320P, SPIE, 2009.
- [29] P. Durlach, J. Neumann, and D. Billings, "Training to Operate a Simulated Micro-Unmanned Aerial Vehicle with Continuous or Discrete Manual Control," *U.S. Army Research Institute for the Behavioral and Social Sciences*, Jan 2008.
- [30] K. Hinckley and M. Sinclair, "Touch-sensing input devices," in *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, (New York, NY, USA), pp. 223–230, ACM, 1999.
- [31] J. Pascoe, N. Ryan, and D. Morse, "Using while moving: HCI issues in fieldwork environments," *ACM Transactions on Computer-Human Interaction*, vol. 7, no. 3, pp. 417–437, 2000.

- [32] J. Rekimoto, "Tilting Operations for Small Screen Interfaces," in *UIST '96: Proceedings of the 9th annual ACM symposium on User interface software and technology*, (New York, NY, USA), pp. 167–168, ACM, 1996.
- [33] I. Jang and W. Park, "Gesture-Based User Interfaces for Handheld Devices Using Accelerometer," *Advances in Multimedia Information Processing - PCM 2004*, vol. 3331, pp. 359–368, Jan 2004.
- [34] M. Rahman, S. Gustafson, P. Irani, and S. Subramanian, "Tilt Techniques: investigating the dexterity of wrist-based input," in *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, (New York, NY, USA), pp. 1943–1952, ACM, 2009.
- [35] K. H. Jo, Y. Kuno, and Y. Shirai, "Manipulative hand gesture recognition using task knowledge for human computer interaction," in *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, pp. 468–473, Apr 1998.
- [36] E. Oliver, "A survey of platforms for mobile networks research," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 12, no. 4, pp. 56–63, 2008.
- [37] J. Schraagen, S. Chipman, and V. Shalin, *Cognitive Task Analysis*. Mahwah, New Jersey USA: Lawrence Erlbaum Associates, Inc., Jan 2000.
- [38] DoD, "Interface Standard, Common Warfighting Symbolology MIL-STD-2525B," tech. rep., U.S. Department of Defense, 1999.
- [39] C. Smith, M. L. Cummings, L. M. Forest, and L. J. Kessler, "Utilizing Ecological Perception to Support Precision Lunar Landing," in *Proceedings of HFES 2006: 50th Annual Meeting of the Human Factors and Ergonomics Society*, pp. 91–95, Jun 2006.
- [40] C. Smith, "An Ecological Perceptual Aid for Precision Vertical Landings," Master's thesis, MIT, May 2006.
- [41] P. Sanderson, J. Flach, and A. Buttigieg, "Object Displays Do Not Always Support Better Integrated Task Performance," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 31, pp. 183–198, Jan 1989.
- [42] A. Buttigieg and P. Sanderson, "Emergent features in visual display design for two types of failure detection tasks," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 33, pp. 631–651, Dec 1991.
- [43] K. Bennett and J. Flach, "Graphical displays: Implications for divided attention, focused attention, and problem solving," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 34, pp. 513–533, Oct 1992.
- [44] B. Shneiderman, "Direct manipulation: A step beyond programming languages," *ACM SIGSOC Bulletin*, vol. 13, no. 2-3, p. 143, 1982.
- [45] J. Jacko and A. Sears, eds., *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, ch. Multimodal Interfaces. Mahwah, New Jersey USA: Lawrence Erlbaum Associates, Inc., Jan 2003.
- [46] R. Bolt, "'Put-that-there': Voice and gesture at the graphics interface," in *SIGGRAPH '80: Proceedings of the 7th annual conference on Computer graphics and interactive techniques*, (New York, NY, USA), pp. 262–270, ACM, 1980.
- [47] S. Oviatt, "Ten Myths of Multimodal Interaction," *Communications of the ACM*, vol. 42, pp. 74–81, Nov 1999.
- [48] R. J. Jagacinski and J. M. Flach, *Control Theory for Humans: Quantitative Approaches to Modeling Performance*. Mahwah, New Jersey USA: Lawrence Erlbaum Associates, Inc., 2003.

- [49] S. Bouabdallah, A. Noth, and R. Siegwart, "PID vs LQ Control Techniques Applied to an Indoor Micro Quadrotor," in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 3, pp. 2451–2456, Jun 2004.
- [50] M. Efe, "Robust Low Altitude Behavior Control of a Quadrotor Rotorcraft Through Sliding Modes," in *Proceedings of the 15th Mediterranean Conference on Control & Automation*, pp. 1–6, Jul 2007.
- [51] S. I. Erwin, "UAV Programs Illustrate DoD's Broken Procurement System," *National Defense*, vol. 2009, May 2009.
- [52] C. D. Wickens and J. G. Hollands, *Engineering Psychology and Human Performance*. Upper Saddle River, New Jersey USA: Prentice Hall, 3rd ed., 2000.
- [53] G. Salvendy, ed., *Handbook of Human Factors and Ergonomics*. New York, New York USA: Wiley-Interscience Publication, 2nd ed., 1997.
- [54] L. P. Deutsch, "RFC 1952: GZIP file format specification version 4.3," tech. rep., Internet Engineering Task Force, 1996.
- [55] S. Vandenberg and A. Kuse, "Mental rotations, a group test of three-dimensional spatial visualization," *Perceptual and Motor Skills*, vol. 47, pp. 599–604, Jan 1978.
- [56] M. Peters and C. Battista, "Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library," *Brain and Cognition*, vol. 66, pp. 260–264, Jan 2008.
- [57] M. Peters, B. Laeng, K. Latham, and M. Jackson, "A Redrawn Vandenberg and Kuse Mental Rotations Test: Different Versions and Factors That Affect Performance," *Brain and Cognition*, vol. 28, pp. 39–58, Jan 1995.
- [58] M. Peters, J. Manning, and S. Reimers, "The Effects of Sex, Sexual orientation, and Digit Ratio (2D: 4D) on Mental Rotation Performance," *Archives of Sexual Behavior*, vol. 36, pp. 251–260, April 2007.
- [59] M. Kozhevnikov and M. Hegarty, "A dissociation between object manipulation spatial ability and spatial orientation ability," *Memory and Cognition*, vol. 29, pp. 745–756, Jan 2001.
- [60] M. Hegarty and D. Waller, "A Dissociation between Mental Rotation and Perspective-Taking Spatial Abilities," *Intelligence*, vol. 32, pp. 175–191, Jan 2004.