

Research Article

Collaborative Exploration with a Micro Aerial Vehicle: A Novel Interaction Method for Controlling a MAV with a Hand-Held Device

David Pitman and Mary L. Cummings

Humans and Automation Lab, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Correspondence should be addressed to David Pitman, dpitman@mit.edu

Received 19 April 2012; Accepted 18 June 2012

Academic Editor: Kerstin S. Eklundh

Copyright © 2012 D. Pitman and M. L. Cummings. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to collaboratively explore an environment with a Micro Aerial Vehicle (MAV), an operator needs a mobile interface, which can support the operator's divided attention. To this end, we developed the Micro Aerial Vehicle Exploration of an Unknown Environment (MAV-VUE) interface, which allows operators with minimal training the ability to remotely explore their environment with a MAV. MAV-VUE employs a concept we term Perceived First-Order (PFO) control, which allows an operator to effectively "fly" a MAV with no risk to the vehicle. PFO control utilizes a position feedback control loop to fly the MAV while presenting rate feedback to the operator. A usability study was conducted to evaluate MAV-VUE. This interface was connected remotely to an actual MAV to explore a GPS-simulated urban environment.

1. Introduction

Field personnel, such as emergency first responders, police, specialists (e.g., building inspectors or bomb technicians), or dismounted, forward-deployed soldiers, often rely on satellite-based maps to gain information prior to or during field operations. All of these groups operate in hazardous environments, which may contain hostile, armed people, unstable structures, or environmental disasters. Satellite maps, currently the standard for performing Intelligence, Surveillance and Reconnaissance (ISR) of an outdoor environment, have many inherent flaws. As 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. Moreover, this imagery is often outdated or relevant only to the season in which the image was taken. Combined, these flaws often give field personnel a false mental model of their environment.

Many of these flaws and dangers could be remedied by having field personnel operate a robot to locally explore and map their environment. Given the need of these personnel to simultaneously perform another primary task, such as looking out for snipers, an autonomous robot (i.e., an Unmanned Vehicle (UV)) would allow these groups to better perform ISR and improve their Situational Awareness (SA) in real time by reducing attention needed from operating the robot. However, performing an ISR mission aided by a UV requires an interface, which allows the user to easily transition between a low workload, high-level control of the robot (e.g., moving to locations of interest) and 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 vehicle, 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 commercial and military applications. Rotorcraft MAVs may have two, four, or six rotors, are typically less than two feet across, and can carry payloads of up to a kilogram, which are typically digital cameras. Rotorcraft MAVs are capable of Vertical Take-Off and

Landing (VTOL), which 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. These MAVs are able to precisely hover and move to a fixed point in the air. 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 [1].

MAVs are currently controlled via computer interfaces known as Ground Control Stations (GCSs). Typically a ruggedized laptop display, GCSs, may incorporate specialized controls such as miniature joysticks or pen styli and range from a hand-held device to a large briefcase in size. If an operator is required to assume the role of a traditional pilot, that is, having command authority over velocity and yaw, roll, and pitch, this aid comes at the cost of increased training requirements, 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, such as navigating hostile environments.

Current operational MAV interfaces are constrained in that the operator’s primary task is to operate the MAV, which includes both flying the vehicle and searching imagery from the vehicle concerning targets of interest. These design choices appear to be the extension of larger Unmanned Aerial Vehicle (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. As a consequence, current GCSs and interfaces have a number of design decisions, which require extensive and costly training, and preclude them from being used effectively by field operators, who almost universally have other, more urgent primary tasks to accomplish.

From a human-centered view, MAVs performing local ISR missions could report directly to personnel in the field 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 that allow fine-grained control over the MAV’s position and orientation, such as peering in a window or over a high wall.

2. Background

2.1. Human Supervisory Control. MAV interfaces embody a form of Human Supervisory Control (HSC), where 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 [2].

HSC of a UAV relies upon a set of hierarchical control loops [3]. If an operator is required to manually perform the inner control loops within this hierarchy, such as piloting a MAV, his attention is divided between the original task (i.e., looking for victims) and lower level functions (i.e., keeping the MAV airborne and free from obstacles). Introducing automation into the inner control loops of basic flight control and navigation allows an operator to effectively execute higher level mission-related goals. To this end, in a later section we will describe a control architecture and user interface that allows a field operator the ability to use a MAV to explore an environment without having to spend critical cognitive resources on low level control and navigation tasks.

2.2. Related Work. Teleoperation was first introduced by Sheridan in his work on Levels of Automation (LOA) and HSC [4]. 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 operator, and in turn leaves less time to perform other tasks. As such, teleoperation is viewed as a difficult problem, especially when compounded with the practical constraints encountered such as time delays in communications and low bandwidth for information, among others.

A large body of literature exists on teleoperation. Chen et al. distilled existing research into a set of constraints common to many teleoperation interactions including Field of View (FOV), orientation and attitude of the robot, frame rate, and time delays [5]. Many of these constraints are still relevant in the case of an autonomous MAV, which delivers live imagery to the operator.

Several researchers [6–10] have investigated using an interface to control a robot from a hand-held device. Many of these interfaces use classical What-You-See-Is-What-You-Get (WYSIWYG) controls and widgets (i.e., sliders, buttons, scroll bars). Multitouch hand-held devices with high-fidelity displays for Human-Robot Interaction (HRI), such as an iPod Touch, have been designed by Gutierrez and Craighead, and O’Brien et al., although neither group conducted user studies [9, 10]. O’Brien et al. implemented a multitouch interface with on-screen joysticks for teleoperation of a Unmanned Ground Vehicle (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 PacBot and do not accommodate changes in altitude.

Murphy and Burke performed a qualitative survey of Unmanned Search and Rescue (USAR) operator’s interaction with robots in search and rescue missions, which led to a specific list of lessons learned [11]. Based on real-world emergency situations and several live exercises, they found the major hurdle to adoption and use of robots in USAR is not due to current robotic capabilities, but the interaction

between the robot and operator. Foremost among their findings were that operators often did not have enough SA to operate the robot or interpret information from the robot's sensors. They also state that the interaction between operator and robot in a USAR domain should be based on consuming information from the robot's sensors rather than operating the robot. Murphy and Burke make a convincing argument for an interface where the primary focus is to facilitate and enhance operator SA through consuming information rather than operating the robot.

Very little research exists specifically on operator interaction with MAV. Durlach et al. completed a study in 2008 which examined training MAVs operators to perform ISR missions in a simulated environment [12]. Operators were taught to fly a 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, and roll axes to maintain balanced flight, which participants could only crash by colliding with other objects in the simulation. No mention was made as to the incorporation of video/communication delay. The study specifically looked at whether discrete or continuous input teleoperation controls yielded better performance using two different two interfaces.

To evaluate these displays and controls, Durlach et al. trained and tested 72 participants. During these flights, the operators manually flew the MAV, 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 Persons of Interest (POIs) and Objects of Interest (OOIs) in a simulated outdoor urban environment. During the mission, the participant had to orient the MAV to take reconnaissance photos of the POIs/OOIs with the MAV's fixed cameras. Twenty-four participants were excluded from post hoc analysis of the first mission by the researchers due to their inability to identify all POIs.

By the end of the experiment, each participant had 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 versus 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.

The generalizability of Durlach et al.'s results is limited because their controls and displays were simulated, with no time delay, or lag, between imagery received by the MAV and displayed to the user and vice versa, inherent in real-world interactions. As shown by Sheridan, a time delay over 0.5 Second (sec) within a teleoperation interface significantly affects the operator's performance [13], so these results are preliminary but provide important lessons learned for user strategies and preferences.

3. Interface Design

A Cognitive Task Analysis (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. While the details of the CTA are provided elsewhere [14], it was found that operators would intermittently use a MAV during their mission. 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 sensors hard mounted on the MAV. The resulting interface, the Micro Aerial Vehicle Exploration of an Unknown Environment (MAV-VUE), is outlined in the following sections along with a discussion of the theory and rationale behind the design.

3.1. MAV-VUE Displays and Interaction. MAV-VUE is a hand-held application that supports 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. Although MAV-VUE is designed to interact with any ground-based or in-air UV, our implementation used a small quad-rotor helicopter, which is capable of VTOL and hovering at a fixed position and heading. MAV-VUE allows the operator to interact with the MAV in two different modes, appropriate to different tasks. The first, Navigation mode, allows the operator to direct the MAV to autonomously fly between specified waypoints. In the second 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 for imagery analysis.

3.2. Navigation Mode: Map and Waypoints. In the Navigation Mode, a map (Figure 1) of the environment occupies the entire iPhone display, which is 320×480 pixels (px). The map displays relevant features of the environment, as well as the location of the MAV and waypoints.

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 an icon typically used in command and control environments.

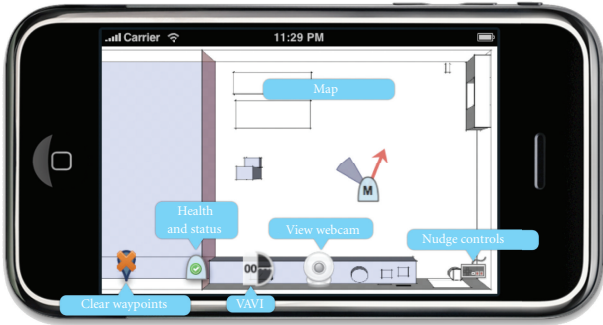


FIGURE 1: The map display and general interface of MAV-VUE.



FIGURE 2: Details of the MAV-VUE map display.

As seen in Figure 2, the MAV's direction and velocity are represented by a red 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. Additionally, users may toggle a small inset view of the MAV's camera. A tool bar along the bottom of the display provides the ability to switch to Nudge Control or show other interface components, such as Health and Status monitoring, or the MAV camera's view.

The map is intended mainly for gross location movements of the MAV, while the Nudge Control mode (Section 3.3) is intended for more precise movements while viewing imagery from the MAV's camera. As such, the map allows the user to construct a high-level flight plan using waypoints. Users double-tap on the map display to create a waypoint at the location of their taps (Figure 2). 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 or detected obstacles.

3.3. Nudge Control Flight Mode. Nudge Control (Figures 3 and 4) allows an operator fine-grained control over the MAV, which is not possible in the more general navigation mode (Section 3.2). A user has the ability to more precisely position the camera (and thus, the MAV) both longitudinally and vertically, in order to better see some object or person of interest. Within the Nudge Control display, users are shown

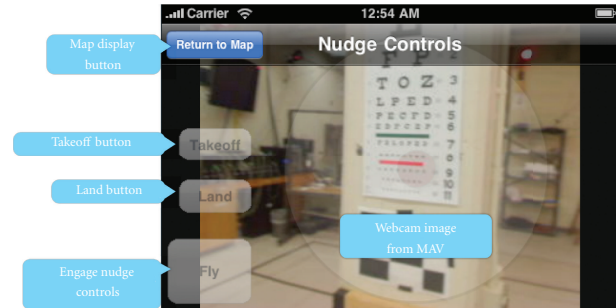


FIGURE 3: Overview of nudge control interface.

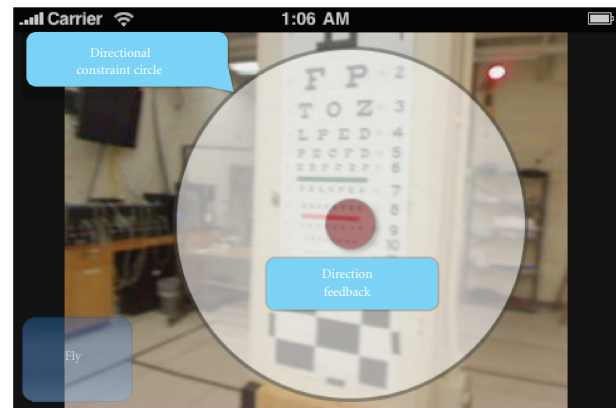


FIGURE 4: Details of nudge control directional interface.

feedback from the MAV's webcam, which is discussed in more detail in the next section.

3.4. Order Reduction of Operator Controls. Control of systems, which incorporate one or more closed feedback loops, is defined as a N th-order system, where N refers to the derivative of the differential equation which describes the feedback loop in the controls used by the human operator. For example, a first-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 of the controls and the desired state of the system, is fed back to the input in an attempt to bring the output closer to the desired state. First-order and higher control systems are commonly known as rate-based control due to the operator manipulating the rate of change of an aspect of the system. In contrast, zero-order control systems are often referred to as position based because operators only provide position coordinates as an input to the system [15]. As an example, changing the heading of a vehicle from 30° to 60° via a 1st order feedback loop requires constantly changing the robot's rate of yaw (how fast the vehicle is turning) until the desired heading is reached. For first-order order systems, operators typically perform a pulsed control input, which has, at least, two distinct actions: first starting the turn at 30° , then ending the turn as the vehicle approaches 60° . In contrast, with a zero-order control loop, the operator simply gives a command of 60° and the vehicle autonomously turns to

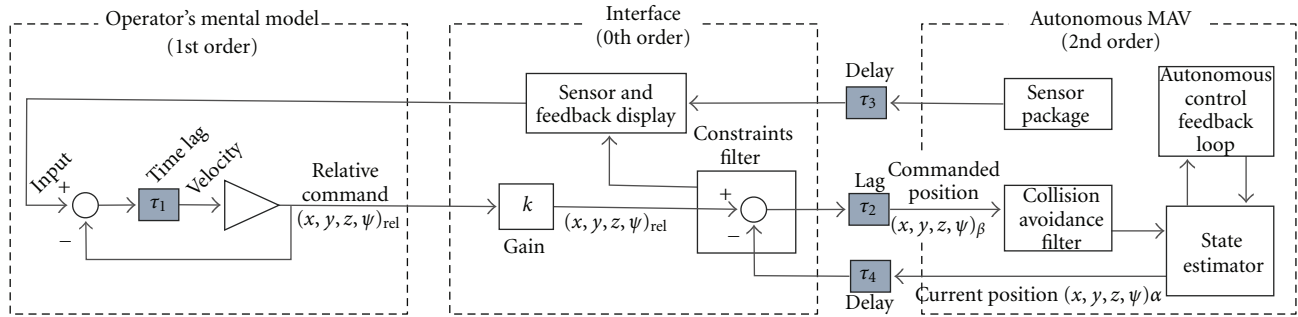


FIGURE 5: Diagram of perceived first-order control. Delays and lags are represented by τ , gains represented by k .

this heading. A 1st order system (changing velocity) requires more attention by the operator as compared to a zero-order system (changing position) since he or she must continually monitor the turn in order to stop the robot at the right time.

A second 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 as they typically use an incorrect cognitive model of a 1st order feedback loop for any higher-order rate-based controls [15]. 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 zero- or 1st order controls, leading to lower performance as shown in a number of studies [4, 13, 16].

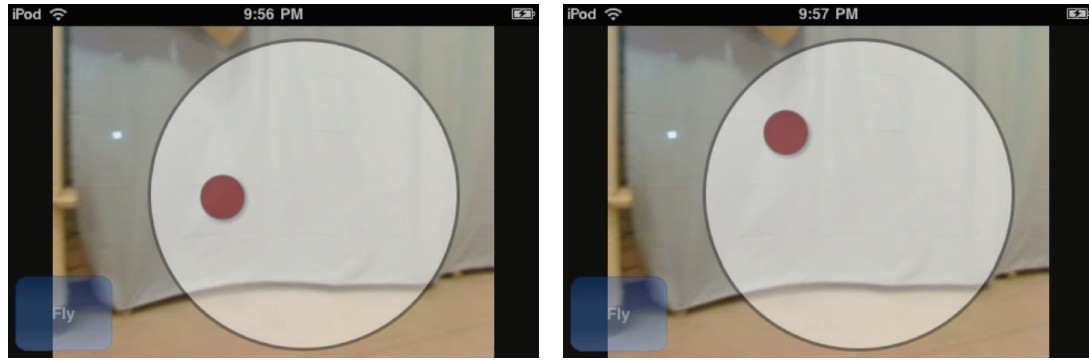
Teleoperation only exacerbates these problems because additional time latencies are introduced into the system, which increases the effect of error in the feedback loop and prevents immediate responses by the operator. In addition, the lack of sensory perception on the part of the operator, who is not physically present at the location of the vehicle, reduces SA which may otherwise allow the operator to compensate for these hindrances. All UAVs use teleoperated 2nd order, or higher, control loops and as a result, have some form of flight control stabilization (i.e., fly-by-wire) to autonomously augment the operator's controls [17, 18].

While human pilots are thought to be effective 1st order controllers, due to their capability to form a working cognitive model of 1st order feedback loops [15], it is doubtful whether UAV pilots can effectively use 1st order controls. 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 autonomously land [19], effectively reducing the pilot's control to zero-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 loops will result in poor operator performance for any type of UAV. This problem would likely be more serious for MAV operators who are not, by the nature of their presence in the field, able to devote the cognitive resources needed to fully attend to the MAV's control dynamics.

For field personnel, it is imperative to reduce the complexity of operating a robot, such as a MAV, which is used

primarily for the purpose of exploring an unknown environment. Operators are under high workload, with their attention divided between many tasks, and their goal is to obtain imagery (i.e., ISR missions), not to fly the vehicle. A solution is to make the control system simpler by reducing the order of the feedback loop to a position-based, zero-order control system, which require less attention and SA than higher-order systems, as well as significantly less training. However, for the precision positioning and orientation required to obtain effective imagery in an ISR mission, position-based, zero-order control systems can be cumbersome and difficult to use. While, in theory, they are safer and less prone to error, unwieldy zero-order control interfaces have impaired many teams at USAR competitions [20] and participants in Durlach et al.'s study [12]. Unfortunately, providing a velocity-based, 1st-order interface to a MAV operator can cause operator control instabilities (e.g., pilot-induced oscillations), as also demonstrated by the Durlach et al. study [12]. In addition, for field personnel controlling a MAV, the environmental pressures of a hostile setting, the need for formal and extensive training, and the issue of divided attention suggest that any type of rate-control systems are not appropriate [11]. As such, some balance between using position-based, zero-order and rate-based, higher-order control is warranted in these scenarios to optimize an operator's performance.

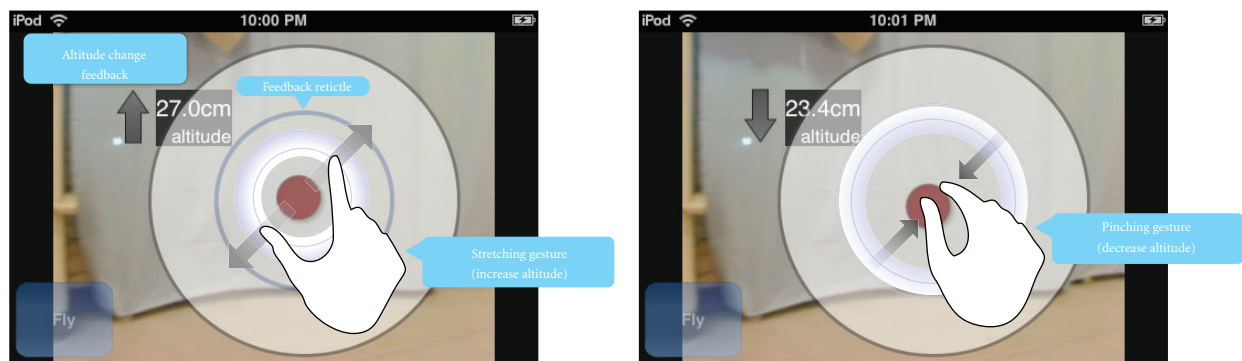
3.5. Perceived First-Order Control. Perceived First-Order (PFO) control can provide a stable and safe zero-order control system, while at the same time presenting 1st-order controls to improve the usability of the operator's interface. We propose that this approach will allow users to achieve effective control of an ISR MAV with minimal training. The intention is to provide a design compromise that increases performance and safety by using different levels of feedback loops which are appropriate to each aspect of the system (including the human). Users perceive that they are operating the robot via a velocity-based, 1st order control interface, which matches their mental model of rate-based controls. However, PFO control converts the user's rate-based 1st order commands (relative velocity changes), into a position-based, zero-order control system (Figure 5). By working in a zero-order control loop that uses absolute position coordinates, commands are time invariant, unlike velocity or acceleration commands. This time invariance eliminates the problem of over/undershooting a target inherent to 1st



(a) Tilting the interface to the left, which commands the MAV to move to the left

(b) Tilting the device forward and to the left, which commands the MAV to move forward and left

FIGURE 6: Interface feedback as a result of the operator performing a tilt gesture with the device.



(a) Increase the MAV's altitude 27 cm by making a stretch gesture

(b) Decrease the MAV's altitude 23.4 cm by making a pinch gesture

FIGURE 7: Gestures to change the z coordinate (altitude) of the MAV.

or 2nd order control systems when operators issue a “bang-bang” set of commands (e.g., a discrete forward command followed by a discrete stop/slow down command) [15]. This hybrid approach allows the user to more accurately and easily predict the movement of a remotely operated robot, such as a MAV, as well as easily formulate plans without sacrificing safety.

In MAV-VUE, users are given visual feedback (Figure 6) of their rate commands by a red dot on the display in the Nudge Control Flight Mode (Section 3.3), which is overlaid on top of sensor imagery. An operator changes the x and y location of the MAV by tilting the hand-held device in the relative direction he or she intends the UV to travel (Figure 4). A tilt gesture has the benefit of leaving the imagery display unobstructed while the user is maneuvering the MAV, unlike a corresponding touch gesture which will obstruct an operator's view of the displayed imagery. The Two-Dimensional (2D) tilt vector of the hand-held device defines the relative distance along the x - y axes from the MAV's existing location (which is considered the origin).

The angle and direction of tilt is calculated by the orientation sensors (e.g., accelerometers) of the device. A discrete-time high-pass filter is used to clean the device's orientation data in order to provide a stable tilt vector [21]. Additionally, a “dead zone” was implemented which ignored tilt gestures

that were within $\pm 14.5^\circ$ in the horizontal x and y plane. This value was empirically chosen based on the testing apparatus and the research of Rahman et al. [22]. The user may also control the heading (θ) and altitude (z) of the MAV.

This interface allows users to feel like they have greater control over the robot's movements and orientation through what appears to be direct control of the robot. Internally, PFO control translates a user's inputs into a position-based, zero-order control loop to prevent the user from putting the robot in jeopardy. This approach also helps to mitigate known problems with time lag, caused by both human decision-making and system latencies. This blend of rate and position control loops drastically decrease the training required to effectively use an interface for an ISR mission. PFO control achieves the best of both position and velocity control while giving users enough control such that they feel they can effectively perform their mission without risking the vehicle's safety.

3.5.1. Altitude Mode. Performing a pinch or stretch gesture on the flight 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 (Figures 7(a) and 7(b)).

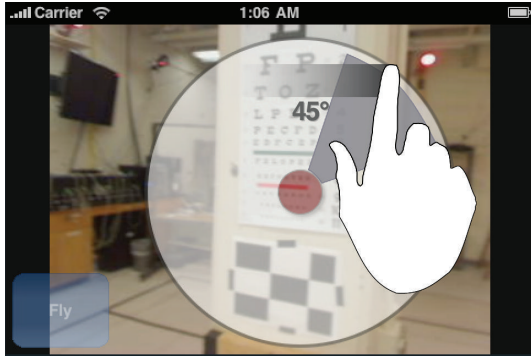


FIGURE 8: Swiping a finger across the screen causes the device to rotate (yaw) right or left.

As the operator performs these gestures, a set of circular rings provides feedback on the direction and magnitude of the gesture. Additionally, the proposed altitude change is shown on-screen along with an arrow indicating the direction of travel.

3.5.2. Heading Control. Operators indirectly control the yaw and pitch of the MAV's sensors through natural touch gestures. The sensor's orientation is determined by performing a swiping gesture across the screen (Figure 8).

The magnitude and direction (left or right) of the swipe corresponds to the magnitude and direction of the relative yaw command, 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.

4. Usability Evaluation

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 of an autonomous MAV.

To achieve these objectives, four research questions were investigated. First, do users find the interface intuitive and supportive of their assigned tasks? Second, can the user effectively manipulate the position and orientation of the MAV to obtain information about the environment? Third, how well does a casual user perform the navigation and identification tasks compared to the model of a "perfect" participant? Fourth, can the user find and accurately identify an OOI and/or a POI using the interface?

The study was conducted using one of two second-generation iPod Touches running MAV-VUE. Each had a screen resolution of 320×480 px and 16-bit color-depth. Both iPod Touches were fitted with an antiglare 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 802.11 g (set at 54 Megabytes (Mb)) Linksys 54 G access point/routers, running either DDWRT firmware or Linksys firmware. The MacBook communicated with the Real-Time indoor Automation Vehicle test Environment (RAVEN) motion-capture network over a 100 Mb ethernet connection. The RAVEN facility [1] 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.

An Ascending Technologies Hummingbird AutoPilot (v2) quad rotor was used for the MAV. This Hummingbird 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.

A Gumstix Overo Fire COM (4 GB, 600 MHz ARM Cortex-A8 CPU, 802.11 g wireless adapter, Gumstix OE OS) with a Summit Expansion Board was mounted on top of the MAV in a custom-built enclosure along with a Logitech C95 webcam, with a maximum resolution of 1024×768 px and a 60° FOV. The webcam was configured with auto-white balance disabled, focus at infinity, resolution at 480×360 px, and connected to the Summit Expansion board via a 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 and 300–500 ms, dependent on the calibration of the RAVEN system and the quality of the XBee radio link.

A preexperiment survey identified each participant's familiarity with Remote Control (RC) vehicles, iPhones, and other relevant demographic information. A postexperiment usability survey was given to judge participants' perceptions of their performance during the flights and of the interface. Participants were also interviewed after the experiment about their experience to gain further feedback.

Since participants' spatial reasoning abilities may be critical in their ability to use the MAV-VUE interface for exploring an unknown environment, participants were given two written tests to assess their spatial reasoning capabilities. The first was the Vandenberg and Kuse Mental Rotation Test (MRT) [23], which is a pencil and paper test used to establish a participant's aptitude for spatial visualization by comparing drawings of objects from different perspectives. The original test has largely been lost and a reconstructed version from 2004 was used [24–26].

The second was the Perspective Taking and Spatial Orientation Test (PTSOT) [27, 28], which is a pencil and paper perspective-taking test shown to predict a participant's ability for spatial orientation and reorientation. Both of these tests were chosen because they have been shown to be a statistically valid predictor of a participant's spatial reasoning skills [26, 28].

4.1. Participants and Procedure. Fourteen participants (8 men and 6 women) were recruited from the MIT community. All participants were between the ages of 18 and 29, with an average age of 22 years (standard deviation (sd) 2.93 years). All had self-reported corrected vision within 20/25, and no color blindness. Nine participants were undergraduate students, three were graduate students, and two were working professionals. Each participant performed the experiment individually. Participants signed an informed consent/video consent form and completed a background questionnaire, which asked about experiences with computers, the military, iPhones, and video games. After finishing the demographic survey, the two spatial reasoning tests were administered.

Following these tests, the experiment and interfaces were explained in detail to the participant. Participants were in a separate room from the MAV and never saw the MAV or environment until the experiment concluded. The experiment administrator demonstrated taking off, navigating via waypoints, flying using nudge controls to find a POI (represented as a headshot on a 8'' × 11'' 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. Participants were allowed to ask questions about the interface and were assisted by the demonstrator if they

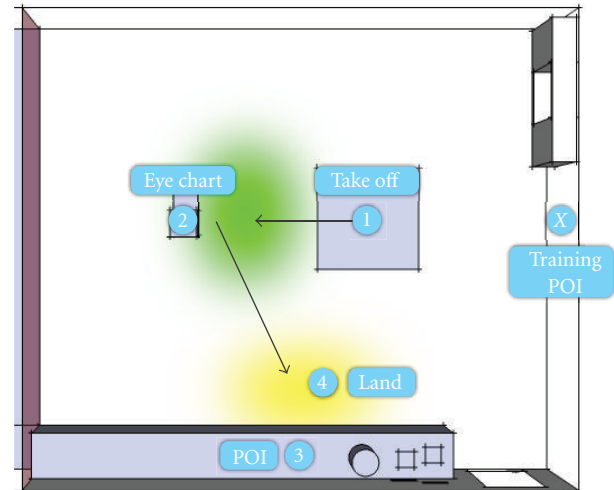


FIGURE 9: Annotated version of the map used in the study showing the layout of the environment.

became confused or incorrectly used the interface. Aside from the demonstration and a three-minute training flight, participants were given no other opportunities to practice with or ask questions about the interface before starting the primary experimental scored task.

Once a participant completed the training task, he or she was given an unannotated version of the supplementary map (Figure 9) on paper for the purposes of receiving instructions about their tasks 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 participants 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 9), with neither at the location used in training.

Participants were first instructed to fly to the green area (Figure 9, no. 2) indicated on the supplemental map using waypoints, and once there, to search for a Snellen eye chart in the vicinity, which was placed at a different height (1.67 m) than the default height the MAV reached after takeoff (0.5 m). 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 (Figure 9, no. 4) of the supplementary map and to search the vicinity for a POI headshot, which was recessed into a box at location no. 3 in Figure 9, placed at a height of 1.47 m. 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 staff. If the participant reached the POI with less than 30 seconds of flight time remaining, the staff allowed the participant up to an extra minute of flight before landing.

After finishing the task, each participant was asked to fill out a survey selecting the POI he or she recognized during the flight from a photo contact sheet. Participants concluded the experiment by taking a usability survey and answering questions for a debriefing interview conducted by the experiment administrator. Each experiment took approximately 75 minutes.

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. The results are presented in the next section.

5. Results and Discussion

Participants 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 were 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 reorient after take-off.

During the study, participants completed a scored task, which had two main objectives: (1) to find and read the smallest line of letters they could identify on an eye chart and (2) to find a POI which they were asked to identify after the eye chart task. Given the small sample size, much of the focus of this section is on the qualitative evaluation of the interface. Nonparametric tests were used to analyze quantitative metrics when appropriate. An α of 0.05 was used for determining the significance of all statistical tests.

5.1. Overall Performance. 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). For the scored task (Figure 9), 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 the Navigation Mode. Further descriptive statistics on participants' performance is shown in the appendix. Participants' 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 (Figure 9), it was empirically determined that a perfect human participant 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.

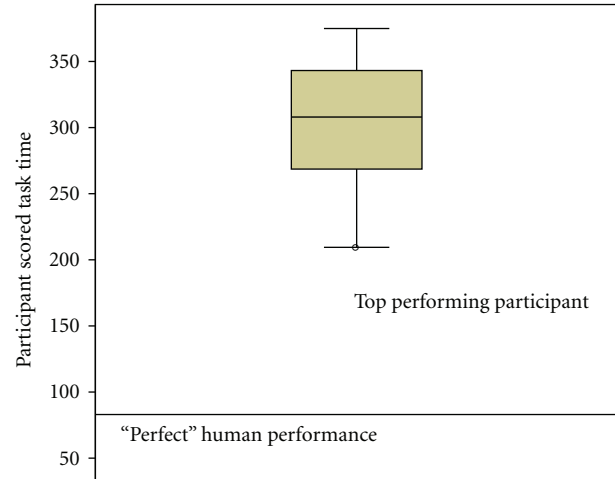


FIGURE 10: Box plot of participants' times to complete the scored task.

This ideal time was compared to the mean of the participants' 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 (Figure 10).

5.2. Eye Chart Identification. During the scored task, participants' first objective was to move to the green area near the eye chart (no. 2 in Figure 9) 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 the eye chart. Participants were able to read between lines 2 and 6 of the eye chart, with a median of line 4. Participants' PTSOT scores were positively correlated with their time to find the eye chart using Nudge Control (Pearson, $r = 0.545$, $P = 0.044$, $N = 14$). A lower PTSOT score is better, so participants with superior spatial orientation abilities found the eye chart faster. Example images from participants' flights are shown in Figure 11. As a comparison, a person with 20/20 vision could read line 4 from 30 foot (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.

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, which often prevented participants from immediately obtaining clear imagery. The line of the eye chart that participants were able to read was negatively correlated with the number of yaw commands issued (Spearman Rho, $\rho = -0.586$, $P = 0.035$, $N = 13$). This correlation indicates that participants who rotated the MAV less were more likely to identify a lower line of the



(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 11: Images obtained by the MAV's camera that participants viewed while trying to read a line of the eye chart.

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, resp.).

5.3. 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 Figure 9) using the Navigation Mode, then switch to Nudge Control to find the headshot of a POI. 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 from the photo contact sheet shown to them after the experiment. 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. Once they initially found the POI, participants used, on average, 70.5 s (sd 38.2 s) repositioning the MAV to obtain better imagery or examine the POI. Example imagery from participants' flights can be seen in Figure 12.

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, $r = -0.593$, $P = 0.033$, $N = 13$), indicating a learning effect, that is, participants who took longer to initially find the eye chart then took less time to find the POI.

5.4. Participants' Navigation Strategies. Three participants' waypoint and Nudge Control commands were reconstructed from logged data, which represent the worst, average, and

best performance. This provides insight into strategies used by participants during the scored task. The paths shown in Figures 13(a), 13(b), and 13(c) 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.

Participant A had the worst performance in the experiment, with a time of 373 s (1.34 sd above the mean), six Navigation waypoints, and 241 Nudge Control commands. 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. 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. Participant A, B, and C's flight paths are shown in Figures 13(a), 13(b), and 13(c), respectively. As shown by these flight paths, participants who issued the fewest commands, that is, more precisely controlled of the MAV to accomplish the same task at hand, had better performance.

5.5. Subjective Responses. After completing the tasks, participants answered a usability survey and were interviewed to gain general feedback on the interface. Participants generally felt confident about their performance using MAV-VUE, with 43% reporting that they were confident about the actions they took and 50% reporting that they felt very confident about their actions.

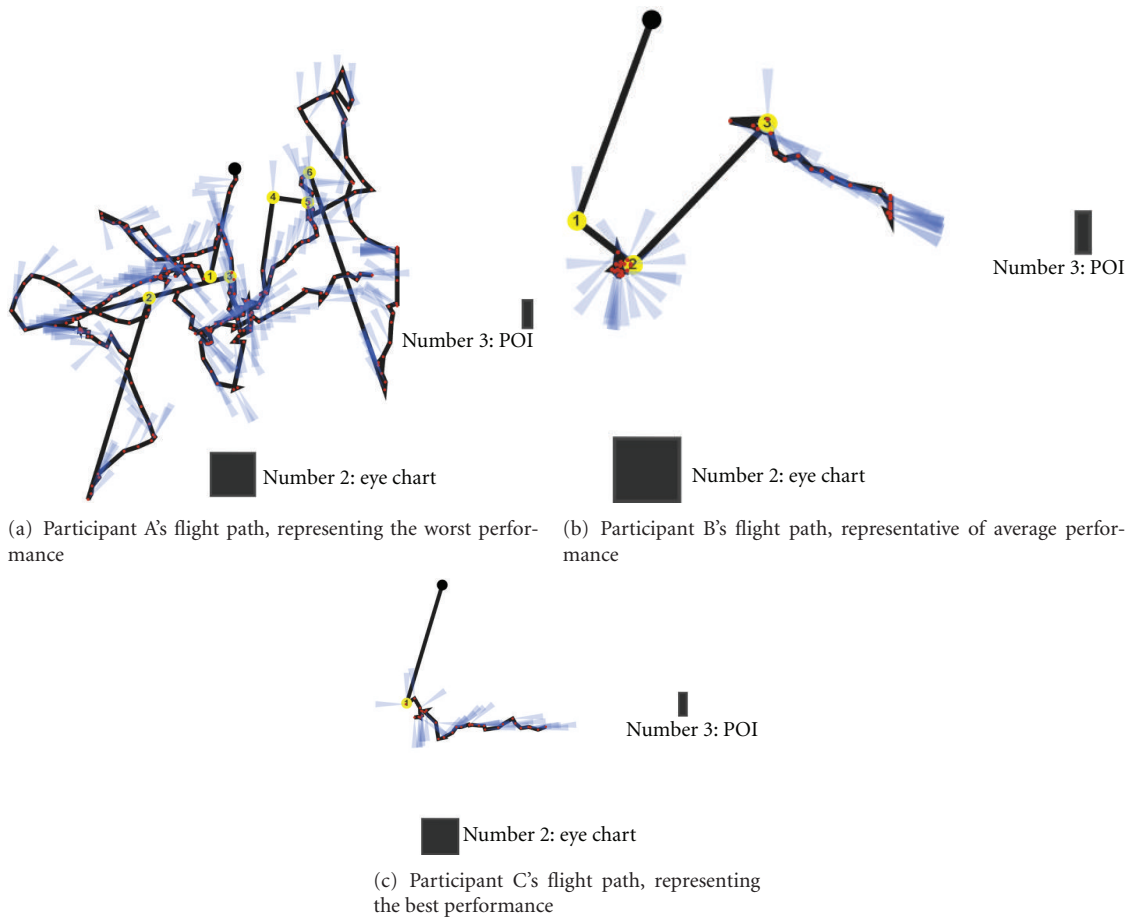
Participants found the Navigation Mode, consisting of the map and waypoints display, easy to use. A third (36%) felt



(a) A participant approaching the POI

(b) An example of imagery used by participants to identifying the POI

FIGURE 12: Examples of imagery seen by participants while finding and identifying the POI.



(a) Participant A's flight path, representing the worst performance

(b) Participant B's flight path, representative of average performance

(c) Participant C's flight path, representing the best performance

FIGURE 13: Examples of participant flight paths.

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 Figure 2) was understood very well by 86% of participants. Twelve participants wrote comments on the survey indicating they found the Navigation mode easy to use.

When asked about aspects of the interface they found confusing or easy to use, participants had conflicting

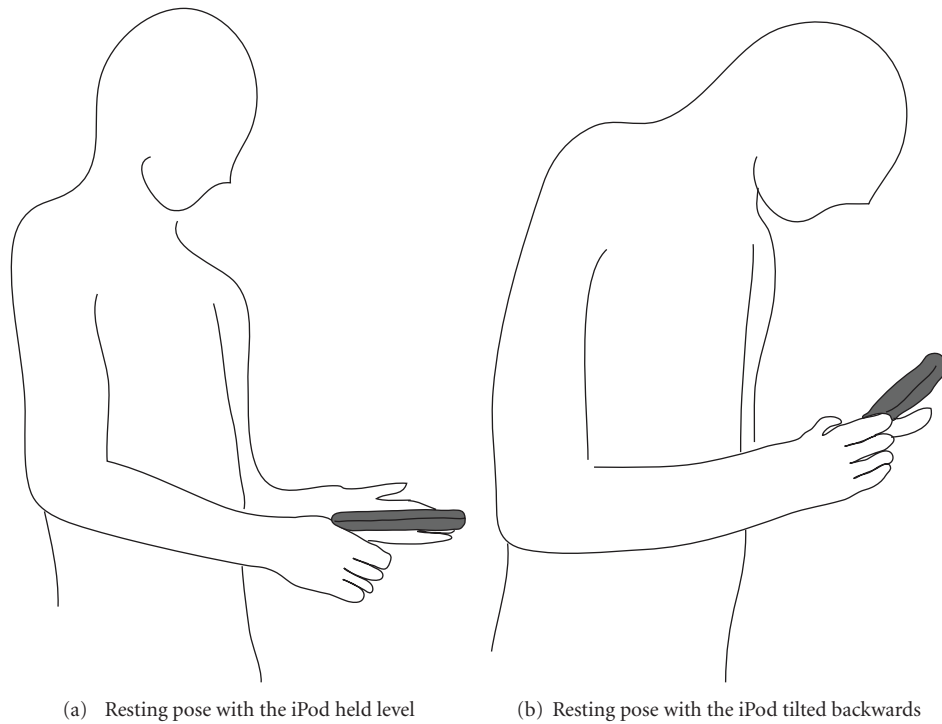


FIGURE 14: Resting poses observed while participants used the interface.

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 concerning Nudge Control, repeatedly expressing the same sentiments that Nudge Control was “easy,” “straight-forward,” or “very intuitive.” However, every participant 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 interacting with the MAV.

5.6. 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 that were not evident from other sources were 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 were to have it tilted slightly towards them (Figure 14(b)) instead of the intended horizontal orientation (Figure 14(a)).

This appeared to be partly due to 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 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. While detected during pilot testing, which led to the development of the dead zone around the neutral point, the full experiment demonstrated the need for either a larger zone or individualized calibration.

6. Conclusions

Even with the availability of satellite imagery, many shortcomings 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 immediate harm, while the MAV can navigate difficult terrain and environments, which may otherwise be inaccessible. Unfortunately, current interfaces for MAVs ignore the needs of an operator in such 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. The U.S. Army has stated that they intend to begin issuing smart phones to recruits for use in the field, so leveraging such ubiquitous tools for MAV operation could reduce both equipment and training costs [29, 30].

TABLE 1: Performance descriptive statistics.

	<i>N</i>	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 (m)	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 commands	13	62.92	45	45	56.47	28	241

TABLE 2: Times descriptive statistics (in seconds).

	<i>N</i>	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 3: Descriptive statistics of nudge control commands performed by participants during the scored task.

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

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 that satisfies these demands while allowing novice users with minimal training to successfully control a MAV in a surveillance setting. 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 authors, this is the first time a formal study has examined using an HRI interface to control and work with a MAV in a real-world setting, and not a simulated environment and vehicle.

The results of this study unambiguously demonstrate the feasibility of a casual user controlling a MAV with a hand-held device to perform search and identify tasks in 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 easily 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 possibly performing detailed surveying and inspection tasks, such as structural inspections. 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. PTSOT scores were also

correlated with participant performance metrics, suggesting that this test can be used as a predictor of participants' performance with the interface. MAV-VUE extends the perception of an operator exploring an unknown environment. Unlike traditional teleoperated UVs though, MAV-VUE does not require that the operator devote their full attention to controlling the UV. Given the cooperative nature between the MAV and operator with MAV-VUE, where the UV intelligently traverses to an AOI and the operator uses Nudge Controls to perform fine-grained reconnaissance of an area, we view this interaction as a collaborative effort which utilizes the strengths of both autonomous robots and human intellect to better explore unknown environments.

Appendix

A. Scored Task Descriptive Statistics

For more details see Tables 1, 2, and 3.

Disclosure

This paper is based on the M. Eng thesis of David Pitman. This research was funded by the Office of Naval Research Grant N00014-07-1-0230 and The Boeing Company.

Conflict of Interests

The authors do not have any conflict of interests with the contents of this paper.

References

- [1] J. P. How, B. Bethke, A. Frank, D. Dale, and J. Vian, "Real-time indoor autonomous vehicle test environment," *IEEE Control Systems Magazine*, vol. 28, no. 2, pp. 51–64, 2008.
- [2] T. Sheridan, *Humans and Automation: System Design and Research Issues*, John Wiley & Sons, New York, NY, USA, 2002.
- [3] M. L. Cummings, S. Bruni, S. Mercier, and P. J. Mitchell, "Automation architecture for single operator, multiple UAV command and control," *The International C2 Journal*, vol. 1, no. 2, pp. 1–24, 2007.
- [4] T. Sheridan, *Telerobotics, Automation, and Human Supervisory Control*, MIT Press, Cambridge, Mass, USA, 1992.
- [5] J. Y. C. Chen, E. C. Haas, and M. J. Barnes, "Human performance issues and user interface design for teleoperated robots," *IEEE Transactions on Systems, Man and Cybernetics C*, vol. 37, no. 6, pp. 1231–1245, 2007.
- [6] T. Fong, N. Cabrol, C. Thorpe, and C. Baur, "A personal user interface for collaborative human-robot exploration," in *Proceedings of the International Symposium on Artificial Intelligence, Robotics, and Automation in Space*, Montréal, Canada, June 2001.
- [7] H. K. Keskinpala, J. A. Adams, and K. Kawamura, "PDA-based human-robotic interface," in *System Security and Assurance*, vol. 4, pp. 3931–3936, October 2003.
- [8] J. A. Adams and H. Kaymaz-Keskinpala, "Analysis of perceived workload when using a PDA for mobile robot teleoperation," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '04)*, vol. 4, pp. 4128–4133, May 2004.
- [9] R. Gutierrez and J. Craighead, "A native iphone packbot OCU," in *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI '09)*, pp. 193–194, ACM, La Jolla, Calif, USA, March 2009.
- [10] B. J. O'Brien, C. Karan, and S. H. Young, "FOCUS—future operator control unit: soldier," in *Unmanned Systems Technology XI*, G. R. Gerhart, D. W. Gage, and C. M. Shoemaker, Eds., vol. 7332 of *Proceedings of SPIE*, p. 733, 2009.
- [11] R. R. Murphy and J. L. Burke, "Up from the rubble: Lessons learned about HRI from search and rescue," in *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society (HFES '05)*, vol. 49, pp. 437–441, September 2005.
- [12] P. Durlach, J. Neumann, and D. Billings, "Training to operate a simulated micro-unmanned aerial vehicle with continuous or discrete manual control," Tech. Rep., U.S. Army Research Institute for the Behavioral and Social Sciences, January 2008.
- [13] T. B. Sheridan, "Space teleoperation through time delay: review and prognosis," *IEEE Transactions on Robotics and Automation*, vol. 9, no. 5, pp. 592–606, 1993.
- [14] D. Pitman, *Collaborative micro aerial vehicle exploration of outdoor environments [M.S. thesis]*, MIT, 2010.
- [15] R. J. Jagacinski and J. M. Flach, *Control Theory for Humans: Quantitative Approaches to Modeling Performance*, Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2003.
- [16] T. Sheridan and W. Verplank, "Human and computer control of undersea teleoperators," Tech. Rep., Office of Naval Research, January 1978.
- [17] S. Bouabdallah, A. Noth, and R. Siegwart, "PID vs LQ control techniques applied to an indoor micro Quadrotor," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '04)*, vol. 3, pp. 2451–2456, June 2004.
- [18] M. Efe, "Robust low altitude behavior control of a quadrotor rotorcraft through sliding modes," in *Proceedings of the 15th Mediterranean Conference on Control and Automation (MED '07)*, pp. 1–6, July 2007.
- [19] S. I. Erwin, "UAV Programs Illustrate DoD's Broken Procurement System," *National Defense*, vol. 2009, 2009.
- [20] H. A. Yanco and J. L. Drury, "Rescuing interfaces: a multi-year study of human-robot interaction at the AAAI Robot Rescue Competition," *Autonomous Robots*, vol. 22, no. 4, pp. 333–352, 2007.
- [21] I. Jang and W. Park, "Gesture-based user interfaces for handheld devices using accelerometer," in *Advances in Multimedia Information Processing*, vol. 3331, pp. 359–368, PCM, 2004.
- [22] M. Rahman, S. Gustafson, P. Irani, and S. Subramanian, "Tilt Techniques: investigating the dexterity of wrist-based input," in *Proceedings of the 27th International Conference on Human Factors in Computing Systems (CHI '09)*, pp. 1943–1952, ACM, Boston, Mass, USA, April 2009.
- [23] S. G. Vandenberg and A. R. Kuse, "Mental rotations, a group test of three-dimensional spatial visualization," *Perceptual and Motor Skills*, vol. 47, no. 2, pp. 599–604, 1978.
- [24] 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, no. 3, pp. 260–264, 2008.
- [25] M. Peters, B. Laeng, K. Latham, and M. Jackson, "A redrawn vandenbergh and kuse mental rotations test—different versions and factors that affect performance," *Brain and Cognition*, vol. 28, no. 1, pp. 39–58, 1995.
- [26] M. Peters, J. T. 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, no. 2, pp. 251–260, 2007.

- [27] M. Kozhevnikov and M. Hegarty, “A dissociation between object manipulation spatial ability and spatial orientation ability,” *Memory and Cognition*, vol. 29, no. 5, pp. 745–756, 2001.
- [28] M. Hegarty and D. Waller, “A dissociation between mental rotation and perspective-taking spatial abilities,” *Intelligence*, vol. 32, no. 2, pp. 175–191, 2004.
- [29] A. N. Service, *Smartphones for All “Makes Sense in Long Run”*, U.S. Army News, 2011.
- [30] C. Heininger, *Army Develops Smartphone Framework, Applications for the Front Lines*, U.S. Army News, 2011.