

Development and Testing of a Quad Rotor Smartphone Control System for Novice Users

M.L. Cummings, K. Jackson, P. Quimby and D. Pitman

Reprinted from

International Journal of Micro Air Vehicles

Volume 4 · Number 3 · September 2012



Multi-Science Publishing
ISSN 1756-8293

Development and Testing of a Quad Rotor Smartphone Control System for Novice Users

M.L. Cummings, K. Jackson, P. Quimby and D. Pitman
77 Massachusetts Ave., 33-311
Cambridge, MA 02139

Received 10 June 2012, Accepted 18 August 2012

ABSTRACT

With recent regulatory efforts to reduce restrictions placed on the operation of Micro Air Vehicles (MAVs) in the United States, it is likely that in the next few years, these vehicles will become commonplace in the commercial marketplace as they are in military environments. In order to reduce the barrier to entry for operations of MAVs, customers of these systems will require ease of operation as well as minimal training time in order to reduce costs. To this end, a smartphone application was developed to control a quadrotor remotely in the exploration of an unknown environment, and tested for users with only three minutes of training. Initial motion capture room tests produced encouraging results for localization and target identification tasks, however, such environments are inherently artificial and the extensibility of such results is limited. A follow-on outdoor field study was conducted in order to compare the indoor and outdoor results and to assess operator performance in a realistic environment. Performance on the outdoor localization tasks was comparable to the indoor study, however, participants generally performed slightly worse on the target identification task in the outdoor experiment, attributed to camera image quality and GPS localization issues. Other issues such as wind and flight safety considerations are discussed.

1. INTRODUCTION

Soldiers, Special Weapons And Tactics (SWAT) teams, and natural disaster first responders are examples of teams of people operating in dangerous and potentially hostile environments who quickly need information about their local environment. These personnel often quickly need aerial imagery of their environments to answer relatively simple questions such as “Is my path blocked?” or “Is there a threat on top of that building?”

The use of Unmanned Aerial Vehicles (UAVs) to support such personnel has become commonplace in the military and with recent regulatory rulings [1], will likely be increasing in the next few years in civilian sectors. Micro Aerial Vehicles (MAVs) that weigh only a few pounds have also been introduced in these settings. However, both traditional small UAV and MAV systems require a dedicated operator with extensive training and control stations that require either a bulky laptop or a two handed controller that resembles popular gaming controllers.

Personnel in these situations have a primary mission other than operating the vehicle, i.e., soldiers are looking for insurgents, first responders are searching for victims, etc. Requiring such operators to have a specialized skill set for operating a MAV or UAV, or requiring additional dedicated personnel to just operate the vehicles places additional demands on resources and could unnecessarily put more people at risk. Moreover, in order for such systems to become viable in the commercial marketplace, the control devices will need to be easy to use with minimal training and low operational costs.

One solution to these problems is to make the control of such systems sufficiently intuitive so that almost any computer-literate person can operate them with little to no training. Such a system could reduce training time, manpower requirements, and costs, as well as reduce hardware requirements if such a control system leveraged off-the-shelf components. For example, the Army has recently announced that it will be equipping all soldiers with smartphones [2]. Given that adding any additional

weight to soldiers' packs is of critical concern [3], leveraging an existing lightweight platform to control a UAV or MAV has obvious advantages.

To this end, a smartphone application was developed to allow users with minimal training to use a quadrotor to remotely explore an unknown environment. This system was initially tested in a motion capture room with encouraging results [4], however, such environments are inherently artificial and the extensibility of such results are limited. To address this gap, we conducted an additional follow-on field study that tested novice users in an actual outdoor setting with an iPhone® application guiding an Ascending Technologies (AscTec)® quadrotor (Figure 1). We demonstrate that with just 3 minutes of training, novice users can effectively use this system to remotely explore an unknown environment, although there were some degradations in performance in the outdoor setting as compared to testing in a motion capture setting.



Figure 1. AscTec quadrotor equipped with onboard camera and GPS receiver

2. BACKGROUND

Unfortunately, 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 helicopter and DraganFly's DraganFlyer are publicly available, these use proprietary interfaces and no information is available about their development. Thus, in order to determine what design constraints and variables should be considered when developing a MAV interface for novice users, research from the field of teleoperation was evaluated.

The term teleoperation was first introduced by Sheridan in his work on levels of automation and human supervisory control [5]. 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 cognitive problem, especially when compounded with the practical constraints encountered in practice (i.e. time delays in communications, low bandwidth for information).

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) [6]. Many of these constraints are still relevant to the case of an autonomous MAV delivering live imagery to an operator. Fong et al. proposed teleoperation with a semi-autonomous robot, which may reinterpret or ignore teleoperation commands from the operator [7]. 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.

Several researchers [8-11] have investigated controlling a ground-based robot from a hand-held device. These interfaces used classical What-You-See-Is-What-You-Get (WYSIWYG) controls and widgets (i.e., sliders, buttons, scroll bars). For those efforts including user studies (which are not common due to the difficulty of conducting human studies), results show that teleoperation is difficult in two dimensions and that cognitive workload is high, primarily because of the need for almost continuous attention for the control device.

Although there has been little published research on teleoperating MAVs, past teleoperation research indicates that controlling a MAV in three dimensions will be a difficult problem. Operators experience high cognitive workload when engaged in teleoperation, exacerbated by the constraints such as time delay, frame rate, sensor FOV, and orientation of the MAV [6-7]. In order for a user, particularly one that is not highly trained, to effectively control a MAV, this previous research suggests that teleoperation, in the classical sense, will not be an effective control choice.

In the only other known published study on MAV interface design and human testing [12], operators were directed perform a search mission in a simulated environment with a simulated Honeywell RQ-16 MAV, with either a mouse or game controller. The study specifically looked at whether discrete or continuous input teleoperation controls yielded better performance. Seventy-two participants completed 7 practice missions, which took two hours to complete before testing could begin. The results showed that operators in the continuous input teleoperation mode control performed statistically significantly better overall, however, over half of all participants in both modes crashed, again raising the issue of whether teleoperation is a viable for remote control of MAVs. Perhaps most importantly for the purposes of our design, Durlach et al. [12] identified a common participant strategy of using gross control movements to approach a target, then hovering and switching to fine-grained teleoperation controls to obtain the necessary imagery.

This previous literature suggests that teleoperation, at least in the classic sense, is not the optimal strategy for MAV control, particularly if the goal is to develop a system that requires minimal training for a large cross-section of users (which is expected to be the market in the near term). We propose that a better solution is to develop a time-invariant Human Supervisory Control (HSC) framework such that users can give high level spatial commands (both horizontally and vertically), akin to the discrete commands explored by Durlach et al [12]. In a HSC setting, an operator intermittently commands an intermediate agent to respond to new goals or changing conditions [5], and the system does not require continuous, or even periodic attention. Such a HSC system reduces cognitive workload by offloading the actual control of the system to automation, while preserving the user's ability to command higher-level goals. The design of such a system is detailed in the next section.

3. MICRO AIR VEHICLE VISUALIZATION OF UNEXPLORED ENVIRONMENTS (MAV-VUE)

In order to provide an operator with no special training the ability to control a MAV, the Micro Aerial Vehicle Visualization of Unexplored Environments (MAV-VUE) application was developed. While the application was designed for a second-generation iOS® environment, it is platform agnostic and can be adapted for any mobile computing system such as a smartphone. MAV-VUE has two modes of operation, Waypoint Control and Nudge Control, discussed in detail next. These two modes were motivated by the user strategies exhibited in previous research [12], where operators tended to use gross localization movements, before attempting to fine tune their positions.

3.1. Waypoint Control

In Waypoint Control, a map of the environment occupies the entire display, (320x480 pixels (px)). The map displays relevant features of the environment, as well as the location of the MAV and waypoints entered by the user (Figure 2). Given the small display size, the user may zoom in and out of the map

by using standard 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® and other touch interfaces.

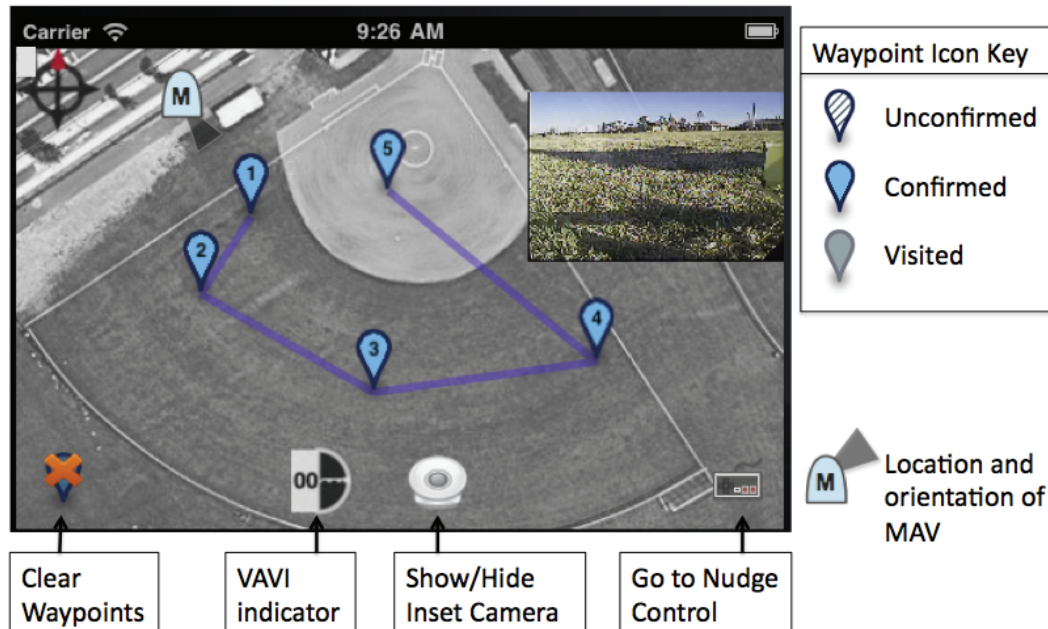


Figure 2. MAV-VUE Waypoint Control Mode

A grey arc shows the current orientation of the MAV's camera (Figure 2). The spread of this arc is an accurate representation of the field of view of the on-board camera. The map is intended primarily for gross location movements, i.e., letting the user command a general area of interest, while the Nudge Control mode (described in more detail in the next section) is intended for more precise movements using the camera view. As such, the map allows the user to construct a high-level flight plan using waypoints.

In keeping with our HSC design approach, the MAV autonomously flies between each waypoint, avoiding obstacles if present, without requiring additional action by the user. A teardrop icon represents each waypoint. Waypoint icons change color depending on the MAV's current state as shown by the key in Figure 2. The flight plan is displayed as a solid line connecting waypoints in the order they will be visited. Users double-tap on the map display to create a waypoint. This waypoint is then added to the queue of waypoints and transmitted to the MAV.

Embedded within the Waypoint Control display is a Vertical Altitude and Velocity Indicator (VAVI) (Figure 2). 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 [13]. 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 a larger version of the VAVI by tapping on the VAVI indicator.

3.2. Nudge Control

Nudge Control provides an operator the ability to more precisely position the camera (and thus the MAV) in all three dimensions, in order to better see some object or person of interest. Nudge Control is accessed by tapping on an icon at the bottom of the Waypoint Control display (Figure 2). Within the Nudge Control display (Figure 3), users are shown feedback from the MAV's camera.

MAV-VUE Nudge Control 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, and an Inertial Measurement Unit (IMU), or equivalent technology to provide information on its orientation in three dimensions. If not equipped with such technologies, the CT mode only requires conventional multitouch technology. The remainder of this paper will assume NG capabilities. For more details on the CT mode, see [4].

To activate Nudge Control, the user presses and holds down the Fly button. As long as the button is pressed, MAV movement commands are issued once per second, which was an empirically determined value from user testing and prototyping. The red dot in the center is a visual joystick, providing directional control feedback as the device is tilted, in effect translating lateral position and rate of change information to the user. In Nudge Control, a pinch or stretch gesture decreases or increases altitude. 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 control circle increases to provide visual feedback so that the user can direct the MAV. The opacity of the directional controls was purposely chosen to partially obfuscate the camera 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 camera or commanding vehicle movement, this trade-off prevents the user from experiencing mode confusion, or becoming cognitively over-tasked and trying to examine a small, imperfect imagery feed while moving the vehicle.

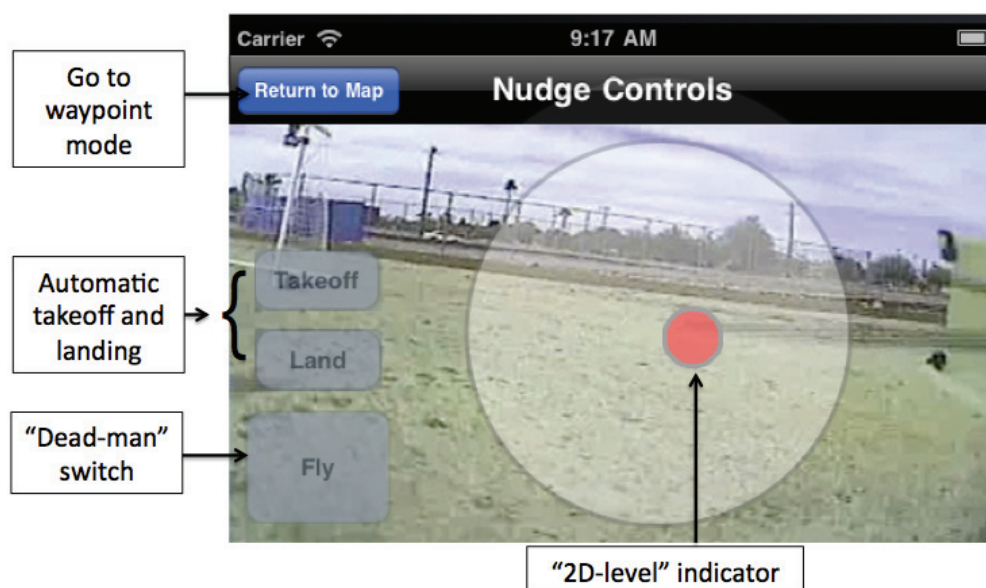


Figure 3. MAV-VUE Nudge Control Mode

The control paradigm in MAV-VUE is termed Perceived First Order (PFO) control, which enables MAV-VUE to be effectively used by operators with no piloting skill sets. While the operator sees commanded rates of change in the movement of the red dot (otherwise known as first order control), a software constraint layer translates the operator’s rate commands into position commands (otherwise known as 0th order control.) These 0th order commands are time-invariant and thus far safer in a remote-controlled robotic system, which may suffer loss of communication or lapses in operator attention which makes rate command much more susceptible to human and system failures.

The 0th order system provides stable and safe control while allowing operators to perceive first order control, which allows them to more accurately and easily predict the movement of the MAV and quickly formulate plans. Pilots typically perform better using first order control [14] so this design ensures safe operations, particularly for non-pilots, but allows them to have the perceptual feedback needed for precise control.

3.3. MAV-VUE Architecture

The MAV-VUE application is implemented using the iOS[®] SDK and open-source frameworks in Objective C. This application relies on a server application, MAVServer, which acts as a middleware layer interfacing between the iOS[®] application and the MAV’s software (Figure 4). The MAVServer exists as a means to offload computationally intensive tasks from the handheld device, log experimental

data, and for ease of implementation. However, in the future as the computational power and developer environment mature on hand-held devices, this server could be eliminated entirely.

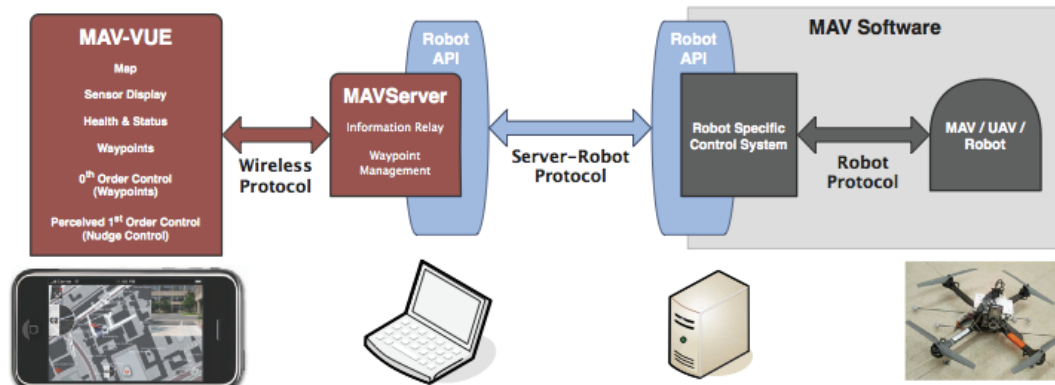


Figure 4. MAV-VUE Architecture

The MAV-VUE architecture is vehicle agnostic and treats the MAV as a subclass of a generic robot. Communication between the iPhone and the MAVServer (Figure 4) 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 [15]. Camera images are transmitted in JPEG format, while map images are transmitted in PNG-24 format.

4. INDOOR AND OUTDOOR EXPERIMENTS

4.1. Indoor Experiment

In order to test our hypothesis that MAV-VUE and the embedded Perceived First Order design could be effectively operated by novices with no flight experience and minimal training, we first conducted an experiment in a motion capture environment (4.6m x 7.8m) that simulates a GPS environment [16]. An AscTec Hummingbird AutoPilot (v2) quad rotor was customized with foam bumpers and reflective dots to function in the motion capture room and the GPS module was removed. 3-Cell Thunderpower™ lithium polymer batteries (1,350 milliAmperes (mA) and 2,100 mA capacity) were used to power the MAV.

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 by the combination of motion capture software and controls algorithms [7], which autonomously flew the MAV between waypoints. An RC transmitter was also used as a backup safety link, as required by the Hummingbird system.

The MAVServer was run on a MacBookPro, 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 access point/routers. The MacBookPro communicated with the motion-capture network over a 100Mb ethernet connection.

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 Hummingbird in a custom-built enclosure. Mounted on top 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 interface.

The webcam had a 60 degree field of view images were captured and transmitted in JPEG format, quality 90, via wireless using User Datagram Protocol (UDP) packets and a custom script based on the uvcapture software from Logitech™. The capture was limited to a maximum rate of 15 frames per

second (fps), although the frame rate experienced by the user was typically between 6 and 10 fps 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 motion capture system and the quality of the XBee radio link.

The actual experiment consisted of providing three minutes of training to 14 participants with MAV-VUE installed on a second generation iPod Touch® and the vehicle apparatus as previously described. The Touch was used instead of the iPhone in testing in order to prevent any confounding effects from cellular service. Participants were offered \$15 compensation and the additional incentive of a \$100 gift card for the top performer. Participants, located in a room separate from the motion capture room, were given a map that identified the general locations of a traditional Snellen eye chart and an 8"x11" picture of a Person of Interest (POI), recessed in a shadow box. The heights of these two targets were not specified and they were not the same. Participants never saw the motion capture room or the vehicle.

Each test session was limited to six minutes of flight time. Participants were expected to use the "TakeOff" button to get the MAV airborne, and then use Waypoint control to get to the general area of the eye chart from the takeoff spot. Using the paper map provided, participants would double tap on the spot where they wished to insert the waypoint. During the MAV's autonomous flight between waypoints, participants simply monitored the progress of the vehicle and could change the altitude if so desired.

Once they felt they were in the right area, participants were then expected to use Nudge Control mode to read the lowest, i.e., smallest, line of the eye chart. In Nudge Control, participants would tilt the device to position the MAV laterally, and use pinch or stretch gestures to obtain the desired height above the ground. Participants were then expected to go back to Waypoint control to get to the general area of the POI picture, switch to Nudge Control to see the picture, and then finally land in place using the "Land" button. Participants were to identify the POI from a contact sheet after the experiment was over. The POI and the others on the contact sheet were randomly picked graduate students unknown to the participants.

4.1.1. Indoor Experiment Results

All participants successfully found and identified the eye chart, and were able to read between lines 2 and 6 of the eye chart, with a median of line 4. 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 yaw. This motion caused the webcam image to blur at times, which often prevented participants from obtaining clear imagery.

Once participants finished reading the eye chart, their next objective was to fly to the POI area using the Waypoint Control mode, then switch to Nudge Control to find the 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, 12 correctly identified the POI from the photo contact sheet shown to them after the experiment. None of the participants had any piloting or remote control flight experience, and none crashed the MAV.

These results, coupled with post-experiment feedback that 13 out of 14 participants felt confident or very confident about their performance with MAV-VUE, demonstrate that with just three minutes of training, persons with no special skill set can effectively control a MAV and perform a search and identification task. However, the generalizability of these results is unclear. The motion capture room allowed for precision control to within 2-3 cm of the expected position, which is not at all realistic when dealing with a GPS-guided system, even with a Wide Area Augmentation System. Moreover, other environmental disturbances such as wind and network instabilities will assuredly cause additional problems.

To this end, we adapted the indoor experiment to an outdoor setting to determine how these indoor results, obtained in a precisely controlled environment, would generalize to more realistic settings. This experiment and the results are described in the next section.

4.2. Outdoor Experiment

Several system modifications were necessary to allow MAV-VUE to operate outdoors. The AscTec Hummingbird has an onboard Global Positioning System (GPS) unit and an Inertial Measurement unit (IMU), so the system was altered to obtain position information from a fused GPS and IMU estimate instead of from the indoor motion capture system.

In order to reduce the payload on the vehicle and increase the flight time for the outdoor distances travelled, the camera system used in the outdoor environment was switched from a traditional webcam to one that transmitted a standard NTSC signal, which was converted to a digital format on the ground with an external video capture card (Hauppauge WinTV-PVR-USB2). This reduced the mass of the onboard camera hardware to 100g, from 215g and also increased the typical frame rate experienced by the user to between 12 and 16 fps.

The Ground Control Station consisted of a laptop running the MAVServer application (as in the indoor experiment), a video receiver, an analog to digital video capture device, and a wireless access point router. All wireless communication occurred over an independent local area network. The interface on the iPod® Touch remained the same as in the indoor experiment.

The field test layout is shown in Figure 5. The vehicle was restrained to the flight region (an outdoor track area located in Arizona) via both a physical tether and a software boundary, centered at the Takeoff position as shown in Figure 5. The tether consisted of a lightweight, high visibility fishing line tied to the underside of the quadrotor and attached to an anchor in the middle of the flight region. The tether was approximately 35 meters in length, allowing the vehicle to fly freely within the area bounded by software constraints. Thus for the second experiment, the total flight area was 3,846m² while the area for the motion capture room in the indoor experiment was substantially smaller at 36m². Because of this, the experiment time was increased from 6 to 9 minutes.

For consistency, the tasks for the outdoor experiment mirrored the tasks performed in the earlier indoor experiment, but with a few minor alterations. First, the POI image was changed and the contact sheet of faces was modified so that the faces were more homogeneous, making it harder to pick out the POI. Additionally, the eye chart was augmented with an extra, larger letter (for viewing from a further distance) and used a non-reflective surface to allow for easier viewing outdoors. Figure 6 contains representative images of both visual targets, as seen from the MAV's webcam.

The POI and eye chart were located at points (1) and (2) in Figure 5 respectively. The order of identification was changed from the indoor experiment since people typically took longer to read the eye chart and flight time was a significant experimental concern. Participants were located inside a sheltered and visually obscured area ~45m to the west of Point 1 in Figure 5. As in the indoor experiment, participants did not see either the course or the vehicle in advance of the experiment.

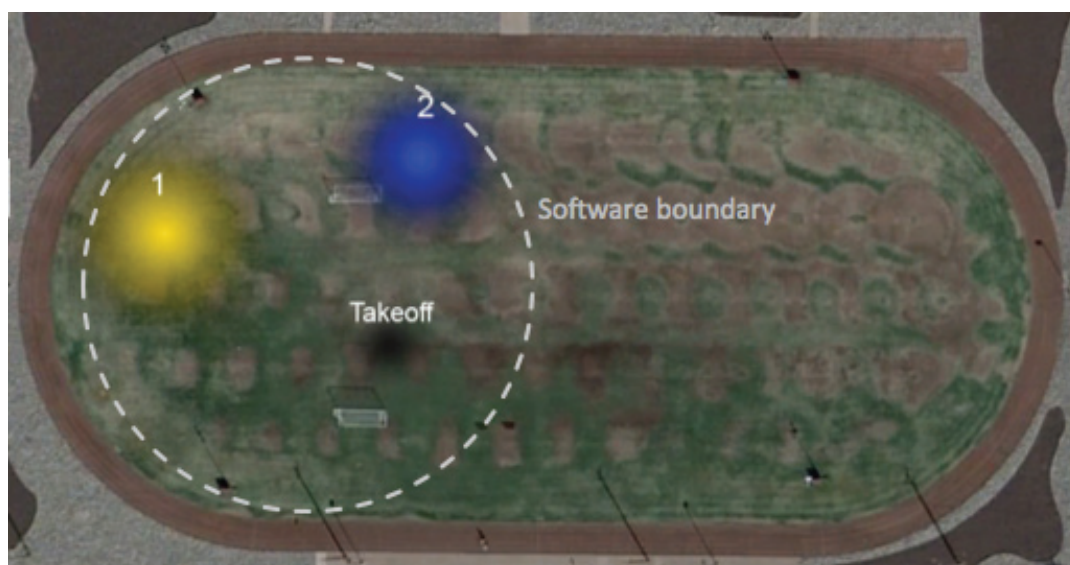


Figure 5. Outdoor Field Testing Layout (Diameter of the circle, the flight area, is 70m)



Figure 6. Person of Interest (a) and the eye chart (b), as see from the camera on the quadrotor.

Just as in the indoor experiment, each participant completed a three-minute practice flight where they were given a chance to test out each part of the interface and encouraged to ask questions. Each participant was provided with a paper map of the course as shown in Figure 5 but without the boundary indicated, which they could reference during the flight. The Waypoint Control interface on the handheld device displayed a similar map, but without the markings for each target. Participants were told that both maps were approximate and could contain errors. The compensation and incentives for the top performer were the same as for the indoor study.

Each participant was given 9 minutes for the test flight, beginning with the takeoff and then the use of Waypoint Control to approach the POI. After inserting the first waypoint and monitoring the progress of the MAV to the POI region, participants were to use Nudge Control to maneuver both laterally and vertically to view the POI. Once they felt confident in their ability to identify the face after the experiment, participants were to use Waypoint Control to navigate to the general area of the eye chart and again use Nudge Control to maneuver to read as many letters as possible from the lowest line legible on the eye chart. When 9 minutes elapsed, participants were instructed to land in place regardless of task completion.

4.2.1. Outdoor Experiment Results

Twelve participants were recruited to participate, often as they walked by the experiment location. The results of their outdoor flight tests are summarized in Table 1. Overall, participants performed similarly on the localization tasks in both the indoor and outdoor tests, with all participants finding the eye chart and all but one finding the POI. As in the indoor flight test, there were no operator-induced crashes, which is a notable result given the often-gusty wind conditions and typical GPS inaccuracies. Moreover, participants in the outdoor condition were under more time pressure since they had to cover greater distances with less available time because of battery limitations. The typical distance covered between the two targets in the indoor study was ~3m as compared to ~35m in the outdoor study. However, even with reduced time to search for targets, participants in the outdoor study were still able to localize the targets as well as those in the indoor study.

Table 1: Comparison of Indoor and Outdoor Results

	Indoor	Outdoor
Located the POI	13 of 14	11 of 12
Identified the POI	12 of 13	3 of 11
Felt confident enough to guess the POI	All	6 of 11
Located the eye chart	14 of 14	12 of 12
Eye chart lines read	Between 2 & 6 (median 4)	Between 2 & 4 (median 3.5)*

*Note that an extra line was added to the top of the eye chart in the outdoor experiment.

These results demonstrate that in a realistic operational environment, persons with no specialized skill set can perform exploration and identification tasks with MAV-VUE with only three minutes of training. However, these outdoor flight tests also highlight that such realistic environments introduce new layers of complexity, particularly those stemming from GPS localization, lighting, and wind, which are discussed below.

4.2.2. GPS Localization

GPS localization issues caused the most significant control problems in the outdoor experiment, as the position estimates in this study had much larger uncertainties than in the indoor experiment. In the motion capture room, position accuracy was estimated at 2-3 cm. In the outdoor environment, localization was achieved through the combined use of the onboard GPS and IMU. GPS can have an error of several meters (a much larger error than in the indoor motion capture facility). Additionally, the noise of the GPS drift combined with environmental factors like wind made the vehicle drift in unpredictable ways.

To counter this problem, virtual safety boundaries were programmed, which were needed to ensure the vehicle would not inadvertently crash into the targets. During the flight tests, 40% of the Nudge Control inputs were constrained by the invisible barrier on average, indicating that operators tended to inadvertently drive the vehicle towards the target. Such problems could be practically overcome by the use of some sort of obstacle detection sensor, but was not an option for these experiments given the need to maximize flight time (and thus minimize weight.)

The primary problematic manifestation of the GPS localization issue was several inadvertent landings. These incidents occurred when the vehicle was flown low, generally while traveling between waypoints. Even though operators would be shown a vehicle height well above the ground, because of the GPS inaccuracies discussed previously, the MAV would fly low enough to the ground that it would simply land (this occurred on average once per trial). None of the incidents caused any damage to the vehicle, and the participants were directed to take off again and continue the task with a thirty second extension to their allowed flight time. Again, such problems could be prevented by more accurate height control.

A second limitation of the system is the degree of accuracy of the vehicle's position in relation to the human operator's reference frame. Outdoors, accuracy was measured in meters, which meant that two identically logged positions for two different participants could be offset by as much as a meter, resulting in very different vantage points. The MAV position was logged at a frequency of approximately 2 Hz. Therefore there was an element of randomization in the surveillance perspective obtained by each operator. This did not impact real-time operations, but made post-experiment analysis more difficult since two identical logged trajectories of the vehicle may not have been identical on the field, resulting in difference experiences for the operators.

As an illustration, comparing the length of the actual logged path with the commanded path based on the waypoints sent to the vehicle generates one measure of the uncertainty and noise. The average commanded path was 416 m with a standard deviation of 200 m. However, the average logged path was 80 m longer than the commanded path, but with a standard deviation of 139 m, meaning that sometimes the actual path was shorter. The discrepancy between these distances can potentially be explained by several factors such as wind and the GPS drift mentioned previously.

4.2.3. Lighting

By far the environmental factor that had the largest impact on operator performance was lighting. Ideally we would have tested at the same time every day, but this was simply not feasible due to delays from hardware troubleshooting, and participant and field scheduling. In the outdoor experiment, lighting conditions were not constant over all tests, as the sun angle shifted throughout the day. All participants completed the experiment between 12pm and 5pm, and sunset occurred at approximately 6pm.

The shifts in the direction and angle of sunlight had several effects in the experimental setup. First, late in the day the sun was setting and as participants rotated to view the POI, the sun could oversaturate the image and cause a loss of contrast. Additionally, although the camera would automatically correct for changing light conditions, the contrast of the camera dropped as the sun set, so flights later in the day did not benefit from the same video quality.

This contrast problem was likely the primary reason for degraded Person of Interest identification

performance between the indoor and outdoor studies. As seen in Table 1, only 3 of 11 participants correctly identified the POI, as compared to 12 of 13 in the indoor setting under tightly controlled lighting conditions. Also as indicated in Table 1, in the outdoor study, half of participants did not feel confident enough to even wager a guess, which was not a problem in the indoor study. This was likely also caused by the more difficult contact sheet.

These results should be contrasted with the eye chart results, which did not appear to be impacted to the same degree by the changing light conditions. Participants in the outdoor setting performed almost as well as those in the indoor, overall reading just one line above that in the indoor study. The similarity in performance was likely due partially to the high contrast (black on white) of the eye chart, so the changing light conditions were not as critical. Additionally, the sun moved through the sky in approximately the same plane as the eye chart, whereas the POI was perpendicular to this plane, causing greater contrast changes.

These results demonstrate that humans can adapt to some types of poor image quality, but that the image identification results of controlled indoor studies can change significantly when placed in a field environment. Moreover, changing weather patterns and changing lighting conditions are realities in real world operations so this is a major concern that should be addressed in future work, likely through improved camera design and post-processing.

4.2.4. Wind

The wind speed, measured hourly from the nearby airport, varied both over the course of each test as well between each of the 4 testing days. Wind speeds varied between 5 and 11 knots. Previous research has demonstrated that the Hummingbird can hold its position fairly well in a stable wind, but any gusts can cause deviations in position that take time to correct [17].

Participants reported frustration when they felt that the vehicle was unresponsive, which was typically due to wind gusts that the operators could neither see nor feel. The latency in vehicle corrections only added to this perceived lack of responsiveness. However, there was no statistically significant effect between wind speed and the time spent observing either the POI or the eye chart. There was also no statistical correlation between wind speed and number of waypoint commands or number of nudge control commands, which means we cannot conclusively show how the wind affected the user's actions.

It should be noted that the wind speeds were averaged over an hour and did not capture local gusts, which were the primary source of user frustrations. This subjective feedback serves to emphasize that users need some feedback as to when wind could be affecting the vehicle's performance. However, the fact that participants performed nearly identical between the indoor and outdoor experiments in locating the POI and the eye chart even in the presence of wind gusts provides evidence as to the robustness of both MAV-VUE and the Hummingbird in realistic settings.

4.2.5. Flight Safety

A physical tether was attached to the MAV in order to maintain positive control and a safe flight test environment. However, its presence had two notable impacts on the system. The first was the tether mobility. The tether allowed the vehicle to fly freely within the software constraint zone so it never directly restrained the flight of the vehicle. However, sudden movements or changes in commanded direction occasionally caused the tether to drag across the ground and become caught on the grass. Wind gusts could also exacerbate this problem. As a result, a field supervisor was on the field at all times while the MAV was flying to free the tether when necessary.

The second problem occurred during one of the practice flights when the vehicle changed direction suddenly, causing the tether to become wrapped in the propeller and the vehicle to crash, breaking a carbon fiber arm. In this instance, the participant was required to return the following day to complete his test flight.

Overall the problems caused by the tether were surmountable, and caused minimal flight test delays. However, such a system has practical limitations in terms of field test size and possible mission types. The tether worked well for our simple target finding and recognition mission, but for more complex missions that require more maneuvering, achieving positive control through a tether is impractical, particularly in obstacle fields.

4.2.6. Subjective Feedback

Participants were directed to use Waypoint Control to navigate to the general area of the intended targets. In general participants found Waypoint Control very intuitive, with 75% reporting that they had “Good” or “Excellent” control of the MAV in this mode. Participants also tended to use Waypoint Control in order to reestablish their spatial orientation when they got disoriented in the environment. This was a behavior not seen in the indoor experiment.

The same result was seen in Nudge Control in that 75% of participants felt the system was responsive to their inputs, with the sources of frustration stemming from winds and the software barrier as previously discussed. One other likely cause of operator frustration with the perceived lack of responsiveness were the problems with the GPS localization, both laterally and vertically.

The strength of the GPS signal varied over time due to the number of available satellites, and during periods of weaker signal, the vehicle sometimes drifted in a circular motion. This resulting MAV motion, seen through the small handheld display, made the vehicle seem less predictable and less stable from the operator’s perspective. Moreover, as discussed previously, the inherent GPS vertical errors also led to operator frustration, as it appeared the MAV was not responding to altitude changes.

The most common complaint of participants concerned the quality of the onboard camera. The image quality was particularly bad when the vehicle was close to the ground and would typically degrade slightly over the course of a flight due to battery voltage drain. There was also a significant amount of random interference noise in the signal.

5. CONCLUSION

Using a quadrotor equipped with a camera, we have shown in both an indoor laboratory experiment and a more realistic outdoor field experiment that novice operators with basic smartphone experience and three minutes of training can effectively search for and identify targets using a handheld device. More specifically, the performance on the localization tasks during the outdoor study was comparable to the indoor study, while participants generally performed slightly worse on the identification portion in the outdoor experiment, which we attribute to camera image quality and GPS localization issues.

One critical performance metric that should be noted is that no participant caused the MAV to crash in either experiment. This is noteworthy because in the only other known published research report examining the impact of handheld display design on an operator’s ability to effectively control a MAV [9], more than half of 72 participants crashed. This previous testing was conducted just in simulation with more than two hours of training for each operator and in all cases, operators were given the ability to rate command the simulated vehicles. This provides further evidence for our hypothesis that to effectively allow even trained operators the ability to remotely operate a small UAV, particularly using a handheld device, the control paradigm should be one of position, not rate, control.

While not explicitly tested in either of our two studies, it should also be noted that this application was used to control an AscTec Hummingbird, but it could be used to control any UAV, as long as the platform leveraged GPS navigation. Such scalability issues are important since operators in the military will likely have access to many different vehicle types, and having a single application that can be used across platforms reduces training demands as well as provides substantial cost savings.

These results have obvious significance for military applications, since reduction in weight of current portable UAV systems is of great importance, as is the need for the reduction of training demands and cost. Moreover, as UAVs begin to become more commonplace for civil applications, which are expected to grow quickly with recent FAA regulation reforms for small UAVs [1], it will be critical that control devices be designed for a wide-ranging population who cannot be expected to have the same level and length of training as military personnel.

While there are many possible extensions to the work reported here, which span both hardware and software possibilities, we are currently focused on two areas in response to our user testing results: 1) Wind alerting, particularly in the likelihood of gusts that could cause significant control problems [8], and 2) Development of a LIDAR-based haptic and visual alerting display to provide operators with additional feedback that they are potentially too close to an object.

The most important improvement that could be made to this system stems from the major flight limitation for all small unmanned systems, which is battery life. The 1-2 lb weight of these systems is very attractive for dismounted field soldiers and other first response personnel who are weight constrained, but the obvious trade is the limitation of flight time. Without substantial improvements in flight time, applications for quadrotors will be necessarily limited in operational settings.

ACKNOWLEDGEMENTS

The Office of Naval Research and the Boeing Company sponsored this research. We would also like to thank Prof. Jon How, Buddy Michini, and other members of the MIT Aerospace Controls Laboratory who were instrumental in helping with the indoor experiment. In addition, we would like to thank Prof. Nancy Cooke of Arizona State University and the Cognitive Engineering Research Institute, especially Mistey Taggart for their assistance in the outdoor experiment.

REFERENCES

- [1] S. Magnuson. (2011, 12 January). *FAA Sets Date for Small Unmanned Planes in U.S. Airspace*. Available: <http://www.nationaldefensemagazine.org/archive/2011/October/Pages/FAASetsDateforSmallUnmannedPlanesinUSAirspace.aspx>
- [2] S. Ackerman. (2011, Army Taps Android Phones for ‘Wearable Computers. *Danger Room*.
- [3] National Research Council, “Tactical Display for Soldiers Human Factors Considerations,” National Academy of Science, Washington, D.C.1997.
- [4] D. Pitman, “Collaborative Micro Aerial Vehicle Exploration of Outdoor Environments,” Electrical Engineering and Computer Science, MIT, Cambridge, MA, 2010.
- [5] T. B. Sheridan, *Telerobotics, Automation and Human Supervisory Control*. Cambridge, MA: The MIT Press, 1992.
- [6] 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, 2007.
- [7] T. Fong, C. Thorpe, and C. Baur, “Collaborative Control: A Robot-Centric Model for Vehicle Teleoperation,” AAI, Technical Report1999.
- [8] T. Fong, N. Cabrol, C. Thorpe, and C. Baur, “A Personal User Interface for Collaborative Human-Robot Exploration,” presented at the International Symposium on Artificial Intelligence, Robotics, and Automation in Space, Montreal, Canada, 2001.
- [9] H. Keskinpala, J. Adams, and K. Kawamura, “PDA-based human-robotic interface,” presented at the IEEE International Conference on Systems, Man and Cybernetics, Washington DC, 2003.
- [10] R. Gutierrez and J. Craighead, “A native iPhone packbot OCU,” in *ACM/IEEE International Conference on Human Robot Interaction*, 2009, pp. 193-194.
- [11] B. J. O’Brien, C. Karan, and S. H. Young, “FOCU:S - Future Operator Control Unit: Soldier,” in *SPIE: Unmanned Systems Technology XI* 2009.
- [12] 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 January 2008.
- [13] M. L. Cummings and C. A. Smith, “Improving Precision Vertical Flight Operations through a Direct Perception-Action Display,” *AIAA Journal of Aircraft*, vol. 46, pp. 1270-1279, 2009.
- [14] R. J. Jagacinski and J. M. Flach, *Control Theory for Humans: Quantitative Approaches to Modeling Performance*. New Jersey: Lawrence Erlbaum Associates, Publishers, 2003.
- [15] L. P. Deutsch, “RFC 1952: GZIP file format specification version 4.3,” Internet Engineering Task Force1996.
- [16] J. How, B. Bethke, A. Frank, D. Dale, and J. Vian. (2008, April) Real-Time Indoor Autonomous Vehicle Test Environment. *IEEE Control Systems Magazine*.
- [17] M. Habib, P. Quimby, K. Jackson, S. Chang, and M. L. Cummings, “Wind Gust Alerting for Supervisory Control of a Micro Aerial Vehicle,” presented at the IEEE Aerospace Conference, Big Sky, Montana, 2011.

