



# CODA: Mobile interface for enabling safer navigation of unmanned aerial vehicles in real-world settings

Erin T. Solovey<sup>a,\*</sup>, Kimberly J. Ryan<sup>b</sup>, M.L. Cummings<sup>c</sup>

<sup>a</sup> Computer Science Department, Worcester Polytechnic Institute, Worcester, MA, USA

<sup>b</sup> Draper Laboratory, Cambridge, MA, USA

<sup>c</sup> Mechanical Engineering & Materials Science, Duke University, Durham, NC, USA

## ARTICLE INFO

### Keywords:

Unmanned aerial vehicle  
Micro aerial vehicle  
Mobile  
Collision avoidance  
Drone

## ABSTRACT

Recent legislation opening commercial operations to small drones, also known as Micro Aerial Vehicles (MAVs) under 55 lbs, is leading to increased use across many facets of society including first responders, news agencies, commercial entities, and hobbyists. However, the widespread use of such systems will be limited unless a few key hurdles are overcome, namely improved usability and support for safe operation in dynamic environments. To this end, we present the Collision and Obstacle Detection and Alerting (CODA) display, a novel interface that enables safe piloting of MAVs with a mobile device in obstacle-dense real-world settings. We describe the system design, architecture, and development. In addition, we present observations from a proof-of-concept technical demonstration and an empirical study with minimally-trained users. Results showed that CODA reduced collisions and improved operator awareness of obstacle proximity, confidence going around corners and perceived performance in avoiding walls and obstacles.

## 1. Introduction

Recent advances in drone technology have resulted in widespread field use for a variety of military and civilian applications. For example, smaller vehicles, called Micro Aerial Vehicles (MAVs), also known as drones, have been utilized for visual search tasks in rugged, potentially dangerous environments to find missing persons as well as to monitor wildfires (Adams et al., 2009; Goodrich et al., 2009). Although the term MAV originally referred to a vehicle less than six inches in diameter, it can now refer to a broader range of small drone systems.

For these smaller, portable unmanned vehicle systems (UAVs), the paradigm has shifted from one where a pilot (or team of operators) controls a vehicle remotely from thousands of miles away to one where the person on the ground can directly use the system's capabilities in a local area. The user interface technology has also shifted toward small, portable handheld devices (Funk, 2018; Merrell, 2018; Williams, 2007). By allowing the user to obtain immediate and current information about his or her surroundings, such systems are ideal for local observation and surveillance tasks, whether on the battlefield, in a disaster area, or for scientific observation.

Recent legislation has opened commercial operations to small drones under 55 lbs (14 CFR 107, 2016) so there is increasing use of these systems in the civilian sector. Organizations such as first

responders, news agencies, commercial entities, and hobbyists have expressed interest in taking advantage of the capabilities drones offer. However, the widespread use of MAV systems will be limited unless a few key hurdles are overcome, namely improved usability and support for safe operation in dynamic environments. This paper addresses these challenges.

In this paper, we extend previous work on MAV user interfaces by presenting techniques for supporting obstacle awareness and collision avoidance using visual and haptic mechanisms on a small mobile device. We introduce the Collision and Obstacle Detection and Avoidance (CODA) system which incorporates these techniques in a working MAV controller. We demonstrate the feasibility of our approach through the design and development of an actual system that assists MAV users in preventing collisions in real-world environments with only three minutes of training. A technical demonstration is discussed that shows how CODA can be integrated into an existing user interface on a mobile device, and support navigation and visual search tasks in a real-world outdoor environment.

Lastly, empirical results from a user evaluation study with 35 participants provide evidence that CODA can make users more aware of their proximity to obstacles and result in fewer crashes in people who had mastered the task. In addition, this study found that scores on a mental rotation test were predictive of the number of collisions

\* Corresponding author.

E-mail address: [esolovey@wpi.edu](mailto:esolovey@wpi.edu) (E.T. Solovey).

<https://doi.org/10.1016/j.ijhcs.2020.102508>

Received 3 October 2019; Received in revised form 10 July 2020; Accepted 11 July 2020

Available online 15 July 2020

1071-5819/© 2020 Elsevier Ltd. All rights reserved.

occurring during the mission. With this work, we advance understanding of how to make MAV systems usable enough to operate safely, effectively, and consistently, in order to facilitate wider adoption of MAV technology.

## 2. Background

As drone usage moves from highly trained individuals in the military to everyday users in the real world, the user interface needs to have minimal training requirements, especially since the FAA does not require any special training for commercial drone operation (14 CFR 107, 2016). A recent study demonstrated that by leveraging automation in the system, operators with only three minutes of training and practice could effectively use a MAV to accomplish visual surveillance tasks in both a controlled lab environment (Pitman and Cummings, 2012) and a more realistic outdoor field environment (et al., 2012). This demonstrates the potential of a well-designed user interface in enabling wider use of MAVs.

However, these earlier tests occurred in a structured environment with no obstacles, and software boundaries were set up to prevent crashes and constrain the vehicle to the experiment area. In a real-world scenario, such constraints are not feasible, as the environment may not be known in advance or may contain dynamic obstacles that the user would need to avoid. Most MAV systems lack the capability to detect objects in the environment and rely solely on operator skill to avoid collisions. A few models of small commercial MAVs have some on-board obstacle avoidance technology, including stereoscopic camera vision, LIDAR, and infrared sensors. However, such technologies are notoriously unreliable and lacking in explicit feedback to the operator to aid them in adjusting their flight path accordingly (Braasch, 2016; Snow, 2016).

MAVs are used in many operations requiring flight in close proximity to existing structures, such as during the assessment of structural damage of buildings and bridges after a disaster. In these cases, typical autonomous collision avoidance, which limits the MAVs movement, may not be feasible or desirable. Per the FAA regulations, the vast majority of flights are limited to the operator's visual line-of-sight. The skill and attention required to maintain a safe standoff distance while correcting for wind deviations and avoiding obstacles can cause increased stress and fatigue in the user, which is only exacerbated by unreliable obstacle detection systems (Murphy et al., 2008). For MAV systems to successfully operate in obstacle-dense areas, unknown battlefield environments, or constrained indoor spaces, they must have the ability to cope with uncertainty and unexpected obstacles. These requirements will become even more important as companies desire to move to beyond-line-of-sight (BLOS) operations. Thus, user interfaces that actively support collision avoidance are essential for effective operation, and the key to future operations.

The challenge is integrating additional information about potential obstacles into the users' display without affecting the usability of the interface. For most MAVs, the small form factor and portability of a mobile display provides a key advantage but limits the display size and screen real estate. Prior work has explored approaches for displaying off-screen information. For example, *Halo* overlays arcs on a small screen to indicate off-screen objects on a street map, inspired by a streetlamp metaphor (Baudisch and Rosenholtz, 2003). The arc size, shape, and translucency are mapped to the distance of the off-screen object so that closer objects have smaller, more curved arcs that are almost opaque. By using a simple arc, the main task screen is not obscured significantly by the off-screen indicators. However, for obstacle avoidance, it is desired that imminent collisions have the most salient features. This might not be the case with *Halo*, since further objects have larger representations on the screen. *Wedge* is a similar visualization technique that was proposed to overcome challenges with overlapping arcs that clutter the interface (Gustafson et al., 2008). Instead of arcs, an acute isosceles triangle is used, where the tip maps to

the off-screen location. The triangles can be rotated until the overlap is avoided, and this was shown to increase accuracy over *Halo* in a user study. *Canyon* was proposed as another approach for showing off-screen information by using a folded paper metaphor, but was designed for large displays without screen size limitations (Ion et al., 2013).

For augmented reality displays, designs have been proposed for off-screen objects including small mini-maps and arrows directly on the main display (Schinke et al., 2010). However, these take up valuable screen real estate and have mainly been studied for informing users of points of interest and not for obstacle avoidance. *Aroundplot* takes 3D spherical coordinates and maps them to a 2D orthogonal fisheye where they are shown as simple dots in the periphery. These augmented reality designs show benefits of having display components directly on the real-world view, but none were designed for a moving object with collision avoidance as the goal.

To support collision and obstacle awareness and avoidance, it is necessary to represent both direction and distance of off-screen obstacles during real-time navigation of a MAV, without increasing an operator's mental workload. Such an interface should also not dramatically increase required training. By presenting this information to operators in an intuitive, embedded way that does not increase mental workload, the system could improve the effectiveness of operators and lead to further adoption of UAVs and MAVs in a wider range of applications.

## 3. CODA user interface

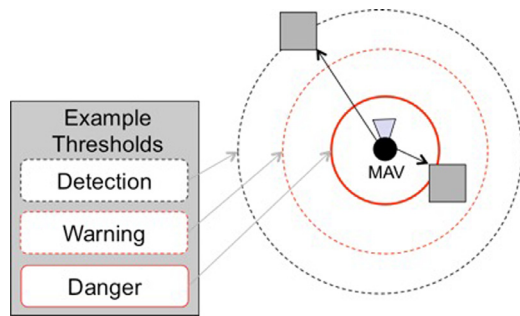
Results from extensive testing of MAV control in both indoor and outdoor (Cummings et al., 2012; Pitman and Cummings, 2012) testing environments were analyzed, including objective performance metrics and subjective responses from MAV operators. While these preliminary studies showed that participants with no experience and about three minutes of training could navigate to locate targets, several observations were made about areas needing further exploration.

During outdoor studies, the flight area was an open field with an invisible software barrier to constrain the vehicle to the test area, as well as a physical tether. This was done for safety purposes, and is common in systems that have the ability to autonomously avoid obstacles and limit the motion of the vehicle. However, the operators became frustrated when they did not have the information conveyed in the controller as to why the system either was not able to respond or was intentionally altering the desired inputs for safety.

In addition, observations indicated that operators need to be aware of objects outside of the field of view. Because a quadrotor MAV is capable of motion in any direction (not just in the forward direction), it is possible to collide with an obstacle that is not in the field of view. While completing a visual task of reading a sign, participants would often move the vehicle side-to-side to align to the proper viewing angle. In a more constrained environment, this could be disastrous to the vehicle if obstructions are present outside the field of view presented to the users.

These observations motivated the need for obstacle awareness, leading to the development of a set of requirements for an obstacle avoidance display and guided the design of a display. The requirements were distilled into the following: (1) The display must warn the user of potential collisions in the vicinity, both within and outside the field of view, (2) The display must show information about location and distance of potential obstacles, (3) The display must integrate effectively into an existing display on a mobile handheld device, to enable the operator to stay focused on the primary navigation task while maintaining obstacle awareness.

To meet this last requirement, the resulting CODA interface was designed to operate on a mobile device, which could be a handheld tablet or a smartphone, due to its portability, functionality, and commercial availability. This device would serve as the hardware platform for controlling the MAV, and many commercial drones have just such

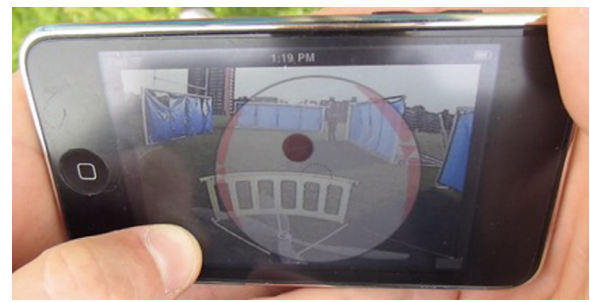


**Fig. 1.** Alerting structure with three thresholds (3, 1.7 and 1 m), corresponding to distances from the vehicle. Distances were chosen that gave the highest signal to noise detection threshold, but could be adjusted per the capabilities of the onboard sensors.

control devices.

In order to meet the first CODA requirement of warning the operator of a potential collision, we assumed that the system would be equipped with one or more sensors that provided information about objects in a two-dimensional plane. Some distance sensors, like a laser rangefinder, can return many distance measurements every second, which may cause information overload if displayed to the user directly. To simplify the information presented to the operator and in keeping with well-established human factors alerting research principles that recommend a staged caution-warning-danger alerting approach for dynamic alerts (Wickens et al., 2004; Wogalter, 2006), the alert system had three stages (Fig. 1) which transition the user from a lower-priority caution stage at 3 m, to a medium-priority warning stage at 1.7 m, to the highest-priority danger phase at 1 m. At 3 m, operators need to know they are near an obstacle, but not in immediate danger. At 1 m, operators are at high risk for a collision, so they need to act immediately to prevent a collision. These distances were chosen because they gave the highest signal-to-noise detection threshold, but they could be adjusted per the capabilities of the onboard sensors. The visual indicators dual-code each threat level using size and opacity, in that the indicators become larger and more opaque as the threat level increases (Fig. 2). This staged approach allows operators to immediately assess the threat environment and adjust their responses accordingly.

The theme of redundant coding was used through CODA, and understanding that the system should not rely solely on visual alerts, especially if the lighting conditions changed, other modalities of alerting were explored. While auditory alerts may be useful in certain settings, they were not explored here because they would not be effective in noisy environments, which are common in MAV operations, and may even be harmful to the goals of the mission (e.g. if the system is used for covert surveillance). Thus, we explored haptic feedback to



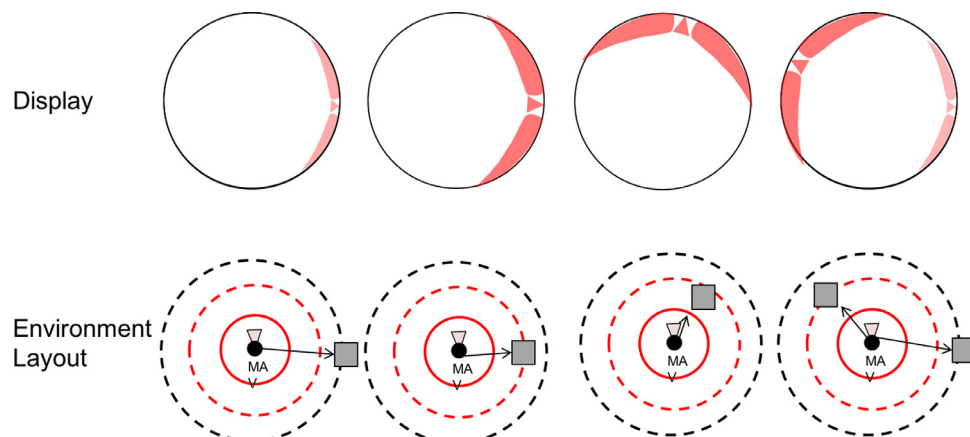
**Fig. 3.** The MAV interface supports tilt commands where tilting the interface forward and to the left would command the MAV to move in that direction. The dot indicates the direction of the tilt or “nudge” command. Here, the CODA display is integrated into this interface to indicate obstacles detected on both the left and right.

complement the visual display. The haptic alert is triggered when the user enters the highest alert threshold (*Danger*), and occurs simultaneously with the appearance of the largest visual display indicator. The alert used the standard vibration supported by iOS4, which consists of a single vibration of fixed intensity, which occurs for a duration of approximately 1.5 s. While different repetitions could have been employed for the different alert levels, users during pilot testing described the alert as “startling” and disruptive to operation. As a result, it was most effective to incorporate a single vibration at the onset of the highest alert level, where disrupting the current course of action is necessary to avoid a collision.

Along with the initial alert that appears on the screen, a circle also appears with the assumption the MAV is at the center. In this circle are a triangle and two arcs, as shown in Figs. 2 and 3. The triangle precisely indicates the obstacle location, and the arcs increase in size and salience as the distance to the obstacle decreases, without obscuring the main control interface. Due to the limited screen size, the indicators are overlaid on the camera image instead of positioned elsewhere. The visual alert indicators are placed on the top of the navigation circle, rather than at the edge of the screen, which allows them to be more central to the field of view.

#### 4. Study 1: proof-of-concept technical demonstration

For the purpose of demonstration and experimentation, the CODA interface was integrated into the Micro Aerial Vehicle Visualization of Unexplored Environments (MAV-VUE) interface, an iPhone-based application that could be used by a minimally-trained operator to accomplish local surveillance tasks. For a full description, the reader should refer to (Pitman and Cummings, 2012). The screen is 2-inches by



**Fig. 2.** Examples of the CODA display for various environmental conditions with the obstacle location (below) and resulting indicator (above).

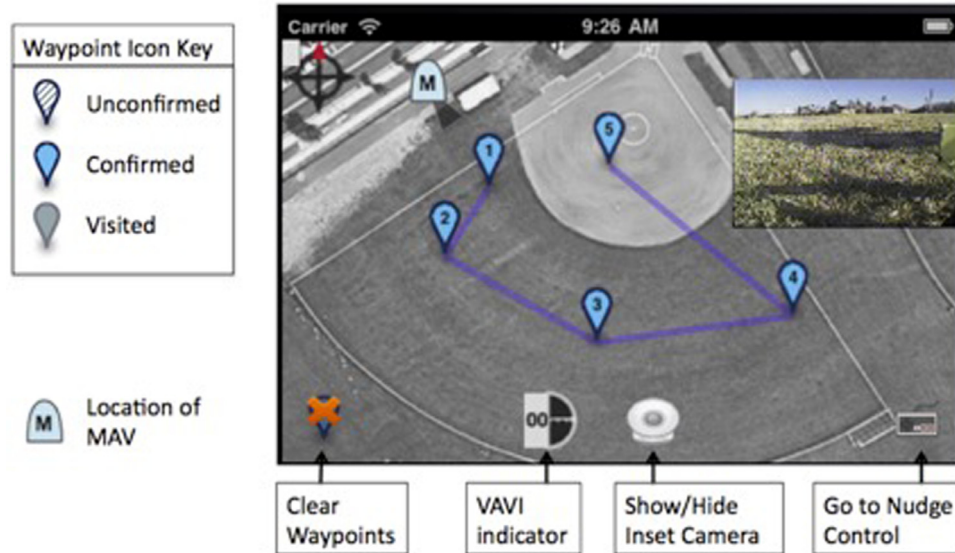


Fig. 4. MAVVUE waypoint control interface.

3-inches, so screen real estate is limited. The system has audio capabilities, and can play a number of built-in alert tones along with an unlimited number of sound files, but is volume-limited.

MAV-VUE has two modes of control. The *waypoint control* interface supports high-level control of the vehicle (Fig. 4). Users place waypoints at desired locations by tapping on the screen, and the MAV autonomously traverses to the locations in the order of creation. This high level of automation results in low pilot workload while the vehicle travels to the target, and the operator may attend to other tasks. The inset camera view (top right corner) provides the view from the MAV camera, and is meant to provide high level situation awareness while the vehicle is in transit.

The *nudge control* mode allows for fine-tuned position inputs once the vehicle reaches an area of interest (Fig. 5). This mode, which is where CODA would be overlaid, allows an operator to explore an unknown area out of his or her line-of-sight, relying on visual feedback from the device. The user gives flight controls through natural gesture inputs. To enable the controls, the user presses the ‘dead-man switch’, which is a type of fail-safe that requires constant active input to remain active and prevents unintentional control commands from affecting the system. While holding this switch, translational movement commands

are given by tilting the device in the desired direction of motion, with the degree of tilt corresponding to the magnitude of the input. Rotational commands require a one-fingered swiping motion around the circle in the center of the display. Altitude commands involve a pinching motion, where the magnitude of the resulting command is proportional to the size of the pinching input. In all three cases, the interface provides visual feedback that the desired inputs have been received.

## 5. Method

As an initial proof-of-concept step, we conducted a sample task scenario in an outdoor environment to ensure the platform functioned as intended. The MAV-VUE interface with CODA embedded was used to control a quadrotor helicopter, the Ascending Technologies (AscTec) Pelican, which can carry up to 500 g of payload in addition to its built-in autopilot, inertial measurement unit sensors, and global positioning system (GPS) receiver. The Pelican was equipped with a Hokuyo UTM-30X laser scanner, which uses a rotating single point laser to sweep out an arc in the horizontal plane. The sensor has a 270-degree field of view and a maximum range of 30 m.

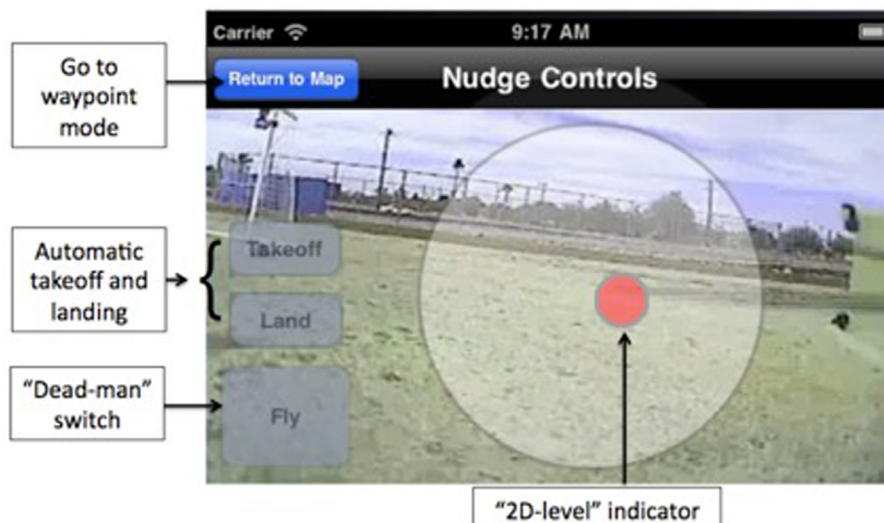


Fig. 5. MAVVUE nudge control interface.



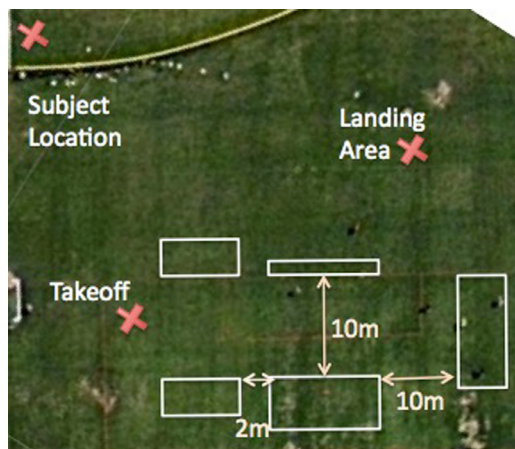


Fig. 6. Map of field layout for outdoor proof-of-concept demonstration.

The quadrotor MAV used a 2.4 GHz XBee radio to communicate with a server program running on a MacBook Pro. The quadrotor's onboard computer had an Atom board processor. Onboard computation, communication, and processing used the Robot Operating System (ROS) framework (Quigley et al., 2009). ROS collected data from the Hokuyo laser scanner and the AscTec autopilot, sent control commands to the autopilot, and transmitted and received data through the XBee serial interface. A 2.4 GHz analog video transmitter was used to send the video feed from the forward-facing onboard camera to the ground based receiver, where the analog video feed is converted to discrete JPEG frames by an external video capture card attached to the server computer. The frames were then sent to the iPhone via UDP.

## 6. Task description

To demonstrate the feasibility of the system, a simple course with obstacles was set up outdoors on an athletic field, and the layout is shown in Fig. 6. Obstacles were constructed using soccer goals with plastic tarps to create a solid reflecting surface for the LIDAR sensors (Fig. 7). Although the system had obstacle sensing capabilities, “no-fly zones” were implemented in the software (in the white boxes in Fig. 6) to prevent damage to the vehicle. These zones were calibrated at the beginning of each test session using the GPS locations of the obstacles since these locations can drift over time. The two demonstration participants were instructed to take off, navigate down the corridor, turn to the left, and land. Participants were located on an adjacent field to the flight area and could not see the vehicle during the task. The user relied solely on the iPhone interface for feedback about the environment during the flight.



Fig. 7. View of outdoor environment with MAV centered at takeoff location.

## 7. Results

We established that obstacle data was successfully transmitted from the onboard LIDAR sensors to the CODA interface in real time, via the ground station, with all components running in full operational mode. The lag experienced ranged from 0.25 to 0.75 s for the control inputs and from 0.5 to 1.5 s for the video and CODA display. The CODA interface successfully represented the obstacles in the mobile display so that the user could take advantage of this information during operation. The two users successfully navigated the vehicle through the outdoor environment without collisions.

As shown by the flight paths in Fig. 8, both participants were able to navigate through the corridor and turn the corner, although both did drift into the ‘no-fly zones’. The zones included a buffer around the actual obstacle, and no actual collisions occurred in either case. This is due not to the interface, but due to limitations in the hardware platform and its sensing capability. The Pelican relied on GPS for position control, and GPS accuracy alone is not sufficient for maneuvering around buildings and structures. With the MAV, we saw position drift of two to four meters, which increased if the windspeed was greater than 8 mph. Given the scale of the course, this could easily cause the system to drift into an obstacle. Additionally, the “no-fly zones” were calibrated using GPS, and the locations would drift over the course of the test flight. This had two effects: (1) The obstacles would no longer be within the zones, causing potential collisions, and (2) The drift would cause ‘obstacles’ where a clear path existed. Further development is needed on the hardware and sensors to increase robustness and improve repeatability of the setup in order to isolate usability problems of the interface from technology and system limitations.

Despite these challenges, this proof-of-concept real-world demonstration confirmed that the CODA design was a viable option for future MAV operations in flight environments with potentially threatening obstacles. We showed that the interface can be successfully integrated in the real world operational environment and receive data from on-board sensors to adapt its display accordingly. Since this paper focuses not on development and improvement of sensor hardware, which will continue to improve, but on the CODA collision avoidance user interface, we developed a simulation environment to be used for repeatable studies to explore whether the CODA display has an effect on MAV operation. The user evaluation of the interface using this simulation is described below.

## 8. Study 2: user evaluation

In order to test the effectiveness of the CODA display in a more rigorous and controlled setting, a user evaluation was conducted. The experiment involved navigating a simulated MAV through a simulated indoor course (Fig. 10). The objectives of this experiment were to assess whether the addition of the CODA display would objectively improve performance and examine how the display impacted user experience.

### 8.1. Experimental design

The experimental factor was the presence of the CODA interface. In the control condition, participants interacted with the vehicle via the original MAV-VUE interface (Pitman and Cummings, 2012) without the CODA display. In the experimental condition, participants used the CODA display integrated with the MAV-VUE system. The experiment was within subjects, with each participant completing the visual search tasks for both conditions. The setup was also counterbalanced, with half of the participants completing the control condition first and half completing the experimental condition first.

### 8.2. Experimental task

To assess the effectiveness of the system for a simulated indoor

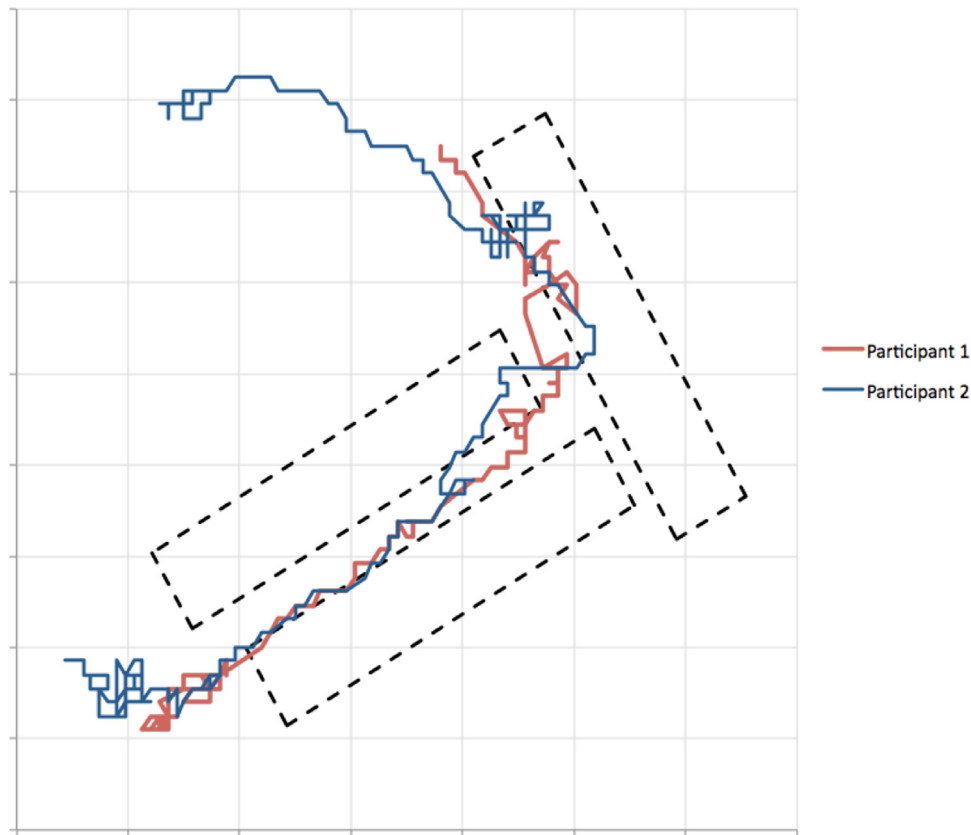


Fig. 8. Flight paths from outdoor pilot testing with two participants in the outdoor environment. The dotted lines represent the no-fly zones.

search mission, experimental tasks were constructed which consisted of locating and observing visual targets in the simulation environment.

The simulation was configured to mimic the capabilities of the hardware platform as closely as possible. Unified System for Automation and Robot Simulation (USARSim) provided a suitable simulation environment with built-in vehicle and sensor configurations (Lewis et al., 2007). The platform is built on the Unreal Tournament Engine and has been previously used in the RoboCup Urban Search and Rescue Challenge. Fig. 9 shows an example screenshot of the simulated indoor environment with the CODA interface. The vehicle used in the simulation was the AirRobot, an existing robot in the USARSim program that is modeled after a real system developed by AirRobot GmbH & Co. The AirRobot is a quadrotor vehicle with a diameter of 1 m. The

predefined settings file includes the robot structure, a forward-facing camera, and a ground-truth sensor. In order to mimic the capabilities of the Pelican-based hardware platform, the existing settings file was augmented with a range scanner with range and resolution properties identical to the Hokuyo UTM-30X. The simulation engine ran on a Dell desktop computer running Windows XP. The USARSim version used was compatible with Unreal Tournament 2004. Screenshots from the simulation were sent over the network as low-quality JPEG images. Communication between USARSim and MAVServer occurred via local wired ethernet with network communications routed through a wired/wireless router.

For each condition, the participant had to complete two visual search tasks, which involved locating a sign on the wall of the environment and reading the word on the sign aloud to the experimenter, which was “FINISH,” “DONE,” or “END.”

Fig. 10 displays the layout for the practice and test courses with dimensions. Participants were instructed to visit the tasks in the specified order. Although the target indicators were in the same general area on the map, participants were told that the targets might not be in exactly the same place in both trials. Additionally, the words printed on the visual targets were different in each condition so that the participants could not rely on memory when reading the targets.

If a participant crashed into a wall, the vehicle would be reset at the ‘save-point’ location corresponding to the most recently crossed threshold. The thresholds and ‘save-points’ were manually placed in roughly even spacing, just before areas that were likely to cause crashes (i.e. turning a corner, entering or exiting the room). Due to the time necessary to reset the simulator, the restart process took approximately three seconds. Once restarted, the participant needed to take off, re-orient themselves, and continue on the mission.

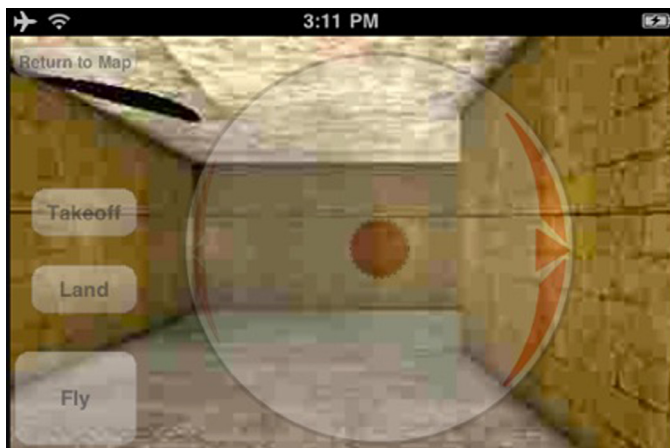


Fig. 9. Example of CODA indicator display in a simulated indoor hallway, as was used in the user evaluation.

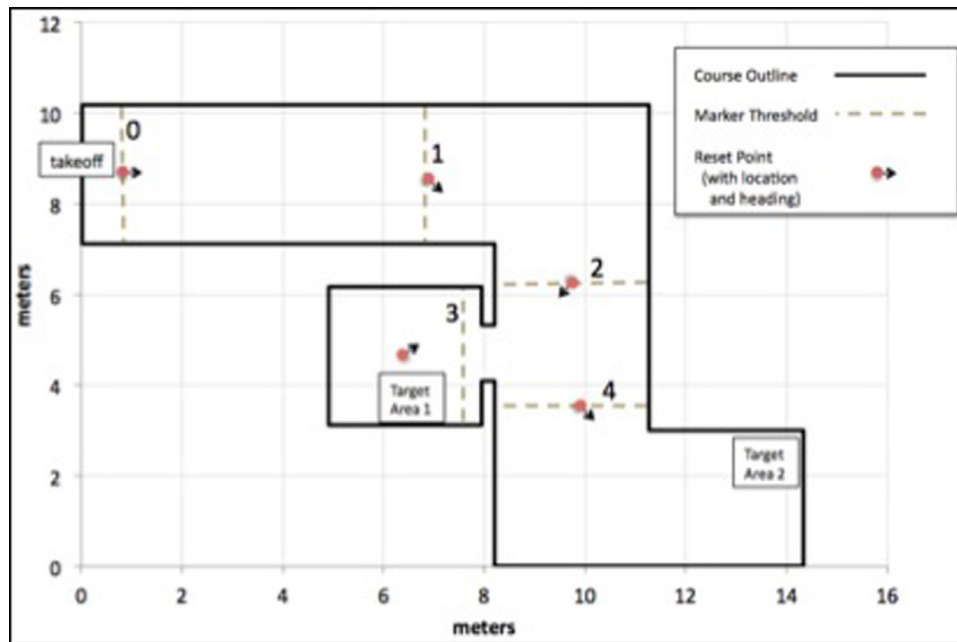


Fig. 10. Indoor course map.

### 8.3. Data collection

The telemetry data from each flight as well as each participant's command inputs were logged as text files on the MacBook computer. All of the simulated onboard video frames were recorded to the MacBook and saved as time-stamped JPEG image files to allow post-flight reconstruction from the participant's perspective. An external video camera recorded the participant's interactions with the device. The experimenter took notes during testing and during the post-flight interview to record observations and additional comments from each participant.

### 8.4. Dependent variables

The dependent variables analyzed can be separated into the following categories: performance metrics, control strategy metrics, and qualitative metrics. These are detailed below.

### 8.5. Performance metrics

- **Task completion:** A primary metric of interest was task completion, or how many participants could complete the mission in the allotted time in each experimental condition.
- **Number of collisions:** The ability to navigate a course without hitting obstacles is key to completing a mission successfully. Collisions occurred when a participant crashed into a wall. In the simulation, participants were alerted to an impending collision but were not prevented from actually hitting the wall. The hypothesis is that fewer crashes would occur with the CODA display, since it provides extra warning.
- **Task completion time:** It was hypothesized that the CODA display would help participants complete the course in a shorter amount of time. Overall task completion time was measured as the time from initial takeoff to final landing after viewing both targets. For participants who did not complete both tasks, completion time was capped at seven minutes (the maximum allotted time), and these participants were not included in the final analysis. Task completion time did not include the server reset time after each crash.

### 8.6. Control strategy metrics

The nudge control inputs for each participant were recorded, both with and without the addition of CODA. This data reveals information about how hard the participants had to work to control the system as well as any underlying control strategies that emerged.

- **Number of nudge controls:** The number of nudge controls required to complete the mission were recorded for the two experiment conditions. This provides a proxy measure for workload, measuring how many commands were required to complete the specified task. The hypothesis was that the CODA display would decrease the number of control inputs required to complete the tasks by allowing operators to navigate more efficiently.
- **Magnitude of nudge controls:** In previous work (Pitman and Cummings, 2012), correlations existed between nudge control commands and performance metrics, which led to the conclusion that smaller, more consistent inputs correlated with higher task performance. In this study, the hypothesis was that this correlation would still exist, and that the CODA display might cause a notable difference in control strategy. Although the participant perceives nudge control commands as velocity inputs, each command actually sends a waypoint to the vehicle, so the magnitude of the command is measured as the distance between the current location and the commanded waypoint.
- **Total path length:** The total path length travelled to complete the course was analyzed. The hypothesis was that the presence of the CODA display would affect the path length, because participants would not cut corners and would take a path that stayed further away from obstacles since they would be more aware of them. Path length included the cumulative path traveled from initial takeoff to final landing, including segments generated by multiple attempts after crashing.

### 8.7. Qualitative measures

Subjective feedback was collected using a survey administered at the end of each trial, consisting of Likert-scale questions for multiple categories relating to usability and preferences. Categories consisted of user confidence, perceived difficulty of the task, awareness of obstacle

proximity, perceived performance, frustration, time pressure, and understanding of the interface. In addition, the experimenter conducted an interview at the conclusion of both flight tasks. Field notes were also taken throughout the experiment.

### 8.8. Procedure

The experiment ran between 50 and 75 min for each participant, depending on how long the participant took to complete the flight tasks. One of two researchers conducted the experiment following the same protocol. Participants completed a consent form and a demographic survey, and then were briefed on the interface and the functions of each of the controls. The experimenter then demonstrated how to use each of the controls by interacting with the simulation in the practice course. This demonstration phase took approximately three minutes.

Participants were allotted three minutes to fly through the practice course and test out the controls. Participants were also instructed to crash into a wall in order to see the system reset behavior. A time of three minutes was selected to mirror the practice time given with earlier versions of the MAV-VUE interface (Pitman and Cummings, 2012). Participants could ask questions during this stage.

For both test flights, the participant had seven minutes to find two visual targets in the real course. During this test portion, the experimenter did not answer any questions or give any advice. Following the completion of each experiment condition, the participant was asked to fill out an evaluation questionnaire. Once the participant completed both conditions, the experimenter conducted a brief verbal interview to get general subjective feedback from the participant.

Finally, the participants completed the two pencil-and-paper spatial tests, the Mental Rotation Test (MRT) (Vandenberg and Kuse, 1978) and the Perceptive-Taking Spatial Orientation Test (PTSOT) (Hegarty and Waller, 2004; Kozhevnikov and Hegarty, 2001). The MRT measures spatial visualization capabilities by asking participants to compare three-dimensional rotations of an object. The version used in this research is a reconstructed version, since the original version has been lost due to deterioration of the existing copies (Peters et al., 1995). A higher score on the MRT represents higher performance.

The PTSOT measures perspective-taking abilities by asking participants to visualize themselves in a given reference frame. The test is scored by adding together the error in each answer, so a lower score on the PTSOT represents higher performance.

These tests were completed at the end of the session to reduce the risk of task performance impacts based on perceived performance on the spatial tests.

## 9. Participants

For this experiment, 47 participants were recruited from a north-eastern U.S. university undergraduate population. Of these, twelve had to be excluded from the quantitative data analysis for logging issues or bugs in the simulation. The 35 participants used in the analysis were between the ages of 18–47 ( $M=23$  yrs,  $SD=6.4$  yrs). Self-reported video game usage varied with 5 people never playing, 20 people playing monthly, 7 weekly, 3 frequently and 0 extreme players. Self-reported iPhone comfort varied from 1 to 5 (on a 5-point scale, where 1 is 'little comfort' and 5 as 'most comfort'), with two people rating themselves as 1 or 2, 5 people rating themselves as a 3, and 28 rating themselves as 4 or 5.

## 10. User evaluation results

### 10.1. Analysis of primary performance metrics

The following section describes the analysis of the performance metrics, including task completion, collision avoidance, mission completion time and total path length. Unless otherwise stated, an  $\alpha$  value

of 0.05 is used for determining significance.

### 10.2. Task completion

Of the 35 participants, twenty-one performed the same in both experiment conditions. Of these, twelve were successful in both trials, and nine participants could not complete either trial. Overall, there were no significant differences between the experiment conditions in terms of task completion. Six participants performed better with the CODA display, finding both targets compared to none in the control condition. Eight participants performed better in the control condition. A significant difference in mental rotation test scores was found between the participants who were able to complete both tasks ( $M = 12.8$ ,  $SE = 1.1$ ) and those who were not ( $M = 9.00$ ,  $SE = 0.80$ ),  $t(22) = 2.82$ ,  $p = 0.01$ ,  $r = 0.512$ .

### 10.3. Collision avoidance, nudge controls, completion time and path length

For participants who completed the task, a two-way mixed ANOVA with the CODA display as a factor, blocked by experimenter, showed a main effect of the CODA display on the number of crashes  $F(1,34) = 8.16$ ,  $p = 0.007$ ,  $f = 0.79$ . We blocked on experimenter since training was potentially affected by the person demonstrating the system. There was no significant effect of the experimenter on any measure and no interaction effect. In addition, the display did not significantly affect any of the other performance measures (nudge control count, completion time or path length). However, with the CODA display, 9 of the 35 participants did not crash at all, compared to only four in the control condition. Only two participants managed to complete the mission in both experiment conditions without crashing.

Given that additional practice could improve flying skills and lower the probability of crashing, it is important to examine whether a learning effect was present between the first and second trial. However, based on a paired  $t$ -test, there was no significant learning effect on the number of crashes between the experiment trials ( $M_{T1} = 2.63$ ,  $SE_{T1} = 0.38$ ;  $M_{T2} = 2.89$ ,  $SE_{T2} = 0.46$ ,  $p > 0.05$ ). In support of the previous results that showed the potential predictive ability of the mental rotation test, there was a marginal moderate negative correlation between participants' mental rotation scores and number of crashes for the experimental condition with the CODA display (Pearson Correlation =  $-0.329$ ,  $p = 0.053$ ). Thus, those participants with higher MRT scores had fewer crashes.

### 10.4. Subjective measures

A Wilcoxon signed-rank test ( $Z$ ) was used to assess whether any of the subjective Likert-scale responses were significantly affected by the introduction of the CODA display. This test was used because the Likert data is not from a normal distribution and so non-parametric methods are required. There were significant differences in several of the measures (Table 1).

The Wilcoxon Signed-rank test showed a significant increase in confidence in navigating around corners when the CODA display was present ( $Z = 2.13$ ,  $p = 0.040$ ,  $r = 0.25$ ). In addition, there was a significant increase in obstacle proximity awareness when the CODA display was present, in general ( $Z = 3.42$ ,  $p < 0.00$ ,  $r = 0.41$ ) as well as when turning corners ( $Z = 2.30$ ,  $p = 0.023$ ,  $r = 0.27$ ), approaching the doorway ( $Z = 3.27$ ,  $p < 0.001$ ,  $r = 0.39$ ), and entering the room ( $Z = 3.48$ ,  $p < 0.001$ ,  $r = 0.42$ ). There was a marginally significant increase when in the straight hallway ( $Z = 1.91$ ,  $p = 0.074$ ,  $r = 0.23$ ) and when locating the target ( $Z = 1.86$ ,  $p = 0.071$ ,  $r = 0.22$ ). There was also a significant increase in self-report assessment of performance when avoiding obstacles and walls ( $Z = 2.49$ ,  $p = 0.013$ ,  $r = 0.30$ ).

Three participants commented on the shape or design of the alert, remarking that the triangle portion of the indicator made sense, but that the "arms" got in the way as they took too much space and made the



**Table 1**

Differences in subjective responses between the CODA condition and the control condition. Significant results are indicated by bold.

Question	Without CODA mean (SD)	With CODA mean (SD)	Z	p	r
1. How confident were you while navigating the MAV? (No Confidence, Minimal Confidence, Somewhat Confident, Mostly Confident, Absolutely Confident)					
• Overall	3.00 (0.87)	3.06 (1.06)	0	1	
• In a straight line	3.80 (0.78)	3.74 (0.86)	-0.5	.804	
• <b>Around corners</b>	<b>2.69 (0.99)</b>	<b>2.97 (1.0)</b>	<b>2.13</b>	<b>.040</b>	<b>0.25</b>
• Through a narrow doorway	2.14 (0.97)	2.24 (0.96)	.721	.512	
2. How difficult was it to navigate the MAV? (Not Hard At All, Slightly Hard, Somewhat Hard, Fairly Hard, Very Hard)					
• Overall	3.06 (1.0)	2.83 (1.1)	-1.84	.096	
• In a straight line	1.6 (0.88)	1.6 (0.82)	0	1	
• Around corners	3.03 (1.1)	2.89 (1.3)	-.639	.579	
• Through a narrow doorway	3.91 (1.1)	3.88 (0.98)	-.293	.776	
3) Do you feel like you were aware of your proximity to obstacles in the environment? (Not Aware, Slightly Aware, Somewhat Aware, Mostly Aware, Absolutely Aware)					
• <b>Overall</b>	<b>2.57 (1.1)</b>	<b>3.2 (1.0)</b>	<b>3.42</b>	<b>&lt;.001</b>	<b>0.41</b>
• In the straight hallway	3.46 (1.1)	3.66 (1.2)	1.91	.074	0.23
• <b>While turning corners</b>	<b>2.66 (1.1)</b>	<b>3.17 (1.2)</b>	<b>2.30</b>	<b>.023</b>	<b>0.27</b>
• <b>While approaching the doorway</b>	<b>2.54 (1.2)</b>	<b>3.24 (1.1)</b>	<b>3.27</b>	<b>&lt;.001</b>	<b>0.39</b>
• <b>While entering the room</b>	<b>2.40 (1.2)</b>	<b>3.03 (1.1)</b>	<b>3.48</b>	<b>&lt;.001</b>	<b>0.42</b>
• While locating the sign in the room	2.74 (1.2)	3.18 (1.3)	1.86	.071	0.22
4. How well did you feel you performed on the following aspects of the task? (Very Poor, Poor, Satisfactory, Good, Excellent)					
• Controlling the MAV using nudge controls	2.63 (1.2)	2.83 (0.98)	1.33	.227	
• <b>Avoiding walls and obstacles</b>	<b>2.09 (0.92)</b>	<b>2.46 (1.0)</b>	<b>2.49</b>	<b>.013</b>	<b>0.30</b>
• Locating and reading the sign	2.86 (1.2)	2.91 (1.1)	.088	.951	
5. Frustration: How frustrated were you during the task? (Not at all, Minimally Frustrated, Somewhat Frustrated, Moderately Frustrated, Very Frustrated)	2.83 (1.1)	2.63 (1.1)	-1.56	.139	
6. Time: During the task did you feel rushed or like you wouldn't be able to finish in time? (Not at all, Only a little, Neutral, Moderately rushed, Very rushed)	2.80 (1.3)	2.66 (1.4)	-.852	.418	
7. Please indicate how well you understood each of the following parts of the application. (Poorly Understood, Somewhat Understood, Well Understood, Did not use)					
• Moving the helicopter	2.83 (0.45)	2.83 (0.51)	0	1	
• Rotating the helicopter	2.66 (0.59)	2.69 (0.63)	.447	1	
• Changing Altitude	3.14 (0.91)	3.06 (0.94)	0.84	0.53	

indicator feel less precise. One participant suggested that the indicators be shaped such that there is an outline that “fills up”, to establish a baseline and make the relative levels clearer.

Eleven participants commented on the vibration feedback during the post-study interview. Eight of them appreciated having the extra vibration alert as it was “helpful” and “good for noticing things”. However, there were some criticisms, including that it was “too sensitive” and “confusing.” One participant stated that after the initial alert, there was not enough information to determine what to do as a result of the alarm. Two participants, who liked the vibration alert, suggested having the intensity or frequency increase (or decrease) as the alert level increased (or decreased), which had been considered in the initial alerting design, but was a limitation of the mobile device used in this experiment. One participant stated that the alert was surprising at first, but felt comfortable after the practice session. One participant noted that they reacted to the alert by stopping before considering the next action and then moving on.

During the debriefing interview, participants were asked which parts of the course they found the CODA display most and least helpful. Participants were able to choose multiple options and some also did not provide answers.

Their responses showed that participants primarily relied on the alerting interface in two areas of the course. First, 19 participants (54%) used it when entering the room to help them center on the opening before attempting to pass through the door. Second, 14 participants (40%) found the CODA display useful when turning a corner. The camera had only a 60° field-of-view, so participants cannot see directly to the side of the vehicle. Without the CODA display, participants would often turn too soon and have to maneuver sideways before going down the next hallway. Five participants (14%) found it useful when approaching a wall in a room.

However, some participants did not like the CODA display in cases where it provided too much information or when information was deemed unnecessary. Twelve participants (34%) said it was not useful on the straight hallway, claiming that information was not necessary. Nine participants (26%) did not like the CODA display when entering the doorway, compared to the nineteen who did find it helpful in the doorway. Because of the small opening, the alert remained at the highest level and these participants found that it did not provide additional useful information, as they just had to “ignore it and go for it”. Four participants (11%) did not find it useful when in a corner and two (7%) said that the CODA display was least helpful in the room with walls on all sides. Finally, one participant (3%) indicated that they did not find it useful during takeoff.

Twelve out of 35 participants remarked that lag was a problem. The simulator was designed to have lag that mimicked the lags seen in outdoor testing, although the lag in the simulation environment was typically more consistent. The framerate was set at 7 frames per second. This is lower than in previous studies (Pitman and Cummings, 2012), but the noise from signal interference that occurred in the outdoor studies was not present. Although participants found the lag to be an annoyance, the task performance indicates that most were able to overcome or adjust to the lag.

## 11. Conclusion and future work

We presented the CODA interface, which provides visual and haptic information about obstacles in the environment and warnings of potential collisions, without substantially increasing the required training or mental workload. A preliminary proof-of-concept study demonstrated feasibility of the CODA system in a real-world, outdoor scenario. A subsequent formal user evaluation in a simulated indoor environment

showed that the CODA interface reduced the number of collisions during a mission and improved an operator's awareness of obstacle proximity, confidence going around corners and perceived performance in avoiding walls and obstacles. Further, we show that scores on the mental rotation test were predictive of the ability of participants to complete a complex obstacle-rich navigation course when using the CODA user interface.

While we showed in this work that such representations through CODA can be assistive, it was clear that operators desire more explicit representations of recommended courses of action. Such recommendations would likely need to integrate local area maps with reliable real-time location information, which would provide complementary functionality to support MAV navigation. Additional highly salient cues could also be evaluated, such as the use of flashing indicators for imminent collisions and giving the operator estimates of time to collision, rather than distance. While the use of audio alerts could and should be explored, operators using drones for first person inspection tasks often work outdoors, often in loud environments so it would be critical to consider the operational context.

It was not clear from this effort to the extent to which the haptic alert was effective, and whether this alert could have been improved. Research in related settings such as driving has shown that haptic alerts can provide superior response times for collision warnings (Fitch et al., 2011; Meng et al., 2015). One study looking at the value of auditory versus haptic alerts in warning roadside police of a potential threat from an oncoming car determined that haptic alerts were superior to audio alerts (Solovey et al., 2017). Moreover, even though the tablets used as the ground control stations are small, it would also be useful to explore hemispherical haptic alerts which would help to localize an obstacle's potential location. Such use of spatial haptic alerts for localized warnings has been shown to improve performance in other settings (Prinet et al., 2016; Spence and Ho, 2008).

The results of this study have implications for the design of mobile user interfaces that facilitate safe and effective piloting of drone systems by minimally trained users with only three minutes of training. In 2018, the Occupational Safety and Health Administration (OSHA) approved the use of drones for workplace inspections, so the use of drones in applications such as pipelines, rooftops, solar farms, and bridges, is rapidly expanding. Thus, designers of these systems need to focus on giving operators feedback about the vehicle's perception and proximity of obstacles both for improved job efficiency and also public safety.

## CRedit authorship contribution statement

**Erin T. Solovey:** Methodology, Validation, Formal analysis, Data curation, Writing - review & editing, Supervision, Project administration. **Kimberly J. Ryan:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **M.L. Cummings:** Conceptualization, Methodology, Validation, Formal analysis, Resources, Writing - review & editing, Supervision, Funding acquisition.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

We thank Kyle Moss and Hank Huang for their contributions. This work was funded by Boeing Research and Technology.

## References

- CFR 107. (2016). Small Unmanned Aircraft Systems.
- Adams, J.A., Humphrey, C.M., Goodrich, M.A., Cooper, J.L., Morse, B.S., 2009. Cognitive task analysis for developing unmanned aerial vehicle wilderness search support. *J. Cognit. Eng. Decis. Mak.* 3 (1), 1–26. <https://doi.org/10.1518/155534309X431926>. <https://doi.org/>
- Baudisch, P., Rosenholtz, R., 2003. Halo: a technique for visualizing off-screen locations. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 481–488. Retrieved from. <http://dl.acm.org/citation.cfm?id=642695>.
- Braasch, M., 2016. Obstacle Avoidance: The Challenge for Drone Package Delivery. The Conversation Retrieved from. <https://theconversation.com/obstacle-avoidance-the-challenge-for-drone-package-delivery-70241>.
- Cummings, M.L., Jackson, K., Quimby, P., Pitman, D., 2012. Field testing of a quad rotor smartphone control system. *Int. J. Micro Air Veh.* 4 (3), 165–177.
- Fitch, G.M., Hankey, J.M., Kleiner, B.M., Dingus, T.A., 2011. Driver comprehension of multiple haptic seat alerts intended for use in an integrated collision avoidance system. *Transp. Res. Part F: Traffic Psychol. Behav.* <https://doi.org/10.1016/j.trf.2011.02.001>. <https://doi.org/>
- Funk, M., 2018. Human-drone interaction: Let's get ready for flying user interfaces. *Interactions*. <https://doi.org/10.1145/3194317>. <https://doi.org/>
- Goodrich, M.A., Morse, B.S., Engh, C., Cooper, J.L., Adams, J.A., 2009. Towards using unmanned aerial vehicles (UAVs) in wilderness search and rescue: lessons from field trials. *Interact. Stud.* 10 (3), 453–478. <https://doi.org/10.1075/is.10.3.08goo>. <https://doi.org/>
- Gustafson, S., Baudisch, P., Gutwin, C., Irani, P., 2008. Wedge: clutter-free visualization of off-screen locations. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA. ACM, pp. 787–796. <https://doi.org/10.1145/1357054.1357179>. <https://doi.org/>
- Hegarty, M., Waller, D., 2004. A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence* 32 (2), 175–191. <https://doi.org/10.1016/j.intell.2003.12.001>. <https://doi.org/>
- Ion, A., Chang, Y.-L.B., Haller, M., Hancock, M., Scott, S.D., 2013. Canyon: providing location awareness of multiple moving objects in a detail view on large displays. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA. ACM, pp. 3149–3158. <https://doi.org/10.1145/2470654.2466431>. <https://doi.org/>
- Kozhevnikov, M., Hegarty, M., 2001. A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory Cognit.* 29 (5), 745–756. <https://doi.org/10.3758/BF03200477>. <https://doi.org/>
- Lewis, M., Wang, J., Hughes, S., 2007. USARSim: simulation for the study of human-robot interaction. *J. Cognit. Eng. Decis. Mak.* 1 (1), 98–120. <https://doi.org/10.1177/155534340700100105>. <https://doi.org/>
- Meng, F., Gray, R., Ho, C., Ahtamad, M., Spence, C., 2015. Dynamic vibrotactile signals for forward collision avoidance warning systems. *Human Fact.* <https://doi.org/10.1177/0018720814542651>. <https://doi.org/>
- Merrell, T.W., 2018. Evaluation of Consumer Drone Control Interface. Wright State University Retrieved from. [https://etd.ohiolink.edu/etd.send\\_file?accession=wright1526381891672951&disposition=inline](https://etd.ohiolink.edu/etd.send_file?accession=wright1526381891672951&disposition=inline).
- Murphy, R.R., Pratt, K.S., Burke, J.L., 2008. Crew roles and operational protocols for rotary-wing micro-UAVs in close urban environments. In: *Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. New York, NY, USA. ACM, pp. 73–80. <https://doi.org/10.1145/1349822.1349833>. <https://doi.org/>
- Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., Richardson, C., 1995. A redrawn Vandenberg and Kuse mental rotations test: different versions and factors that affect performance. *Brain Cognit.* 28 (1), 39–58. <https://doi.org/10.1006/brcg.1995.1032>. <https://doi.org/>
- Pitman, D., Cummings, M.L., 2012. Collaborative exploration with a micro aerial vehicle: a novel interaction method for controlling a MAV with a hand-held device. *Adv. Human-Comput. Interact.* 2012, 15 pages.
- Prinet, J.C., Wan, Y., Sarter, N., 2016. Tactile spatial guidance for collision avoidance in NextGen flight operations. In: *Proceedings of the Human Factors and Ergonomics Society*. <https://doi.org/10.1177/1541931213601069>. <https://doi.org/>
- Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Ng, A.Y., 2009. ROS: an open-source Robot Operating System. In: *Proceedings of the 2009 ICRA Workshop on Open Source Software*. Kobe, Japan. 3. pp. 5. <https://doi.org/http://www.willowgarage.com/papers/ros-open-source-robot-operating-system>.
- Schinke, T., Henze, N., Boll, S., 2010. Visualization of off-screen objects in mobile augmented reality. In: *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*. New York, NY, USA. ACM, pp. 313–316. <https://doi.org/10.1145/1851600.1851655>. <https://doi.org/>
- Snow, C., 2016. Sense and Avoid for Drones is No Easy Feat. *Drone Analyst* Retrieved from. <https://droneanalyst.com/2016/09/22/sense-and-avoid-for-drones-is-no-easy-feat>.
- Solovey, E.T., Powale, P., Cummings, M.L., 2017. A field study of multimodal alerts for an autonomous threat detection system. In: *Proceedings of the 2017 Human-Computer Interaction International*. Vancouver, BC, Canada. 10276. pp. 393–412. [https://doi.org/10.1007/978-3-319-58475-1\\_29](https://doi.org/10.1007/978-3-319-58475-1_29). <https://doi.org/>
- Spence, C., Ho, C., 2008. Tactile and multisensory spatial warning signals for drivers. *IEEE Trans. Haptics*. <https://doi.org/10.1109/TOH.2008.14>. <https://doi.org/>
- Vandenberg, S.G., Kuse, A.R., 1978. Mental rotations, a group test of three-dimensional spatial visualization. *Percept. Motor Skills* 47 (2), 599–604. <https://doi.org/10.2466/pms.1978.47.2.599>. <https://doi.org/>
- Wickens, C.D., Gordon, S.E., Liu, Y., 2004. *An Introduction to Human Factors Engineering*. Pearson.
- Williams, K.W., 2007. An Assessment of Pilot Control Interfaces for Unmanned Aircraft. Office of Aerospace Medicine, Federal Aviation Administration, Washington, DC Retrieved from. [https://www.faa.gov/data\\_research/research/med\\_humanfacs/oamtechreports/2000s/media/200708.pdf](https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2000s/media/200708.pdf).
- Wogalter, M.S. (Ed.), 2006. *Handbook of Warnings*. Lawrence Erlbaum Associates, Philadelphia, PA.