Air-Ground Collaborative Surveillance with Human-Portable Hardware

Michael Dille*, Ben Grocholsky*, Stephen Thomas Nuske*, Mark Moseley†, and Sanjiv Singh*

Coordination of unmanned aerial and ground vehicles (UAVs and UGVs) is immensely useful in a variety of surveillance and rescue applications, as the vehicles' complementary strengths provide operating teams with enhanced mission capabilities. While many of today's systems require independent control stations, necessitating arduous manual coordination between multiple operators, this paper presents a multi-robot collaboration system, jointly developed by iRobot Corporation and Carnegie Mellon University, which features a unified interface for controlling multiple unmanned vehicles. Semi-autonomous subtasks can be directly executed through this interface, including: single-click automatic visual target tracking, waypoint sequences, area search, and geo-location of tracked points of interest. Demonstrations of these capabilities on widely-deployed commercial unmanned vehicles are presented, including the use of UAVs as a communication relay for multi-kilometer, non-line-of-sight operation of UGVs.

INTRODUCTION

Reconnaissance using small unmanned air and ground vehicles (UAVs and UGVs) is finding widespread use in a variety of surveillance and rescue missions. Due to their small size and weight, such robotic vehicles can be deployed quickly by small teams near where they are needed, without substantial infrastructure. Their relatively low cost also permits deployment in larger numbers and riskier missions in hostile zones or challenging terrain from which retrieval is impractical. Finding a single vehicle that can fulfill all aspects of a mission, such as both searching a large area efficiently and providing high-fidelity reconnaissance data such as imagery is difficult, however. UGVs can get up close to objects or targets of interest to provide high-resolution imagery, can carry large accurate sensors, and execute long run-time missions, but they provide a narrow field of view, are relatively slow-moving, and must avoid ground obstacles or danger zones. In contrast, small UAVs are fast-moving and can cover wide areas quickly, above most obstacles, but they can provide only distant, low-resolution images of targets. Therefore, it is advantageous to use both in a system to leverage these complementary properties.

Typically, however, operation of multiple vehicles quickly becomes unwieldy. Most operating procedures require at least one human operator per vehicle, and each type of vehicle usually uses a dedicated, proprietary communication link. Thus, multiple operators must manually coordinate all aspects of a mission, introducing the potential for errors resulting from operator overload or needlessly re-entered coordinates. Further, given the simplicity of their design, most such small vehicles require labor-intensive operator control to search an area, place robots for static observation, and identify or track targets.

*The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA (Email: mdille3@ri.cmu.edu)
†iRobot Corporation, Bedford, MA
‡Distribution statement A: Approved for public release; distribution is unlimited.

To address these issues, iRobot and Carnegie Mellon University, in collaboration with AeroVironment, have developed a unified system for mission planning and task allocation in surveillance-type applications that provides a single operator real-time control of multiple robots through a seamless interface. Target engagement with minimal operator workload is enabled through the use of semi-autonomous subtasks such as automatic visual target tracking, target geolocation estimation, and execution of area search and target pursuit. This is implemented as a single operator control unit (OCU) with subsystems for both UAVs and UGVs providing real-time visual target tracking, waypoint navigation, and pursuit behaviors for each robot. Respective subsystems also include features addressing the different properties of the two classes, such as a wide-area search algorithm for UAVs and obstacle detection and avoidance for UGVs.

This collaborative framework enables a single operator to perform complex surveillance missions in which UAVs and UGVs cue one another to the location of potential targets for further inspection by the other. Live runs of several such missions are demonstrated through tens of hours of field tests using presently-fielded vehicles with minimal special interfacing to demonstrate the practicality and immediate applicability of the system. Additional collaborative capabilities enabled by recent hardware advancements such as the use of a UAV for long-range, non-line-of-sight operation of a UGV are also presented.

This work presents contributions on many fronts. First, we present a seamless framework for collaborative control of both air and ground vehicles that accomplishes practical missions. We further apply recent work in computer vision to produce the best known automatic visual tracking results on limited hardware using lossy remote video streams. Substantial study of the effects of UAV sensor error produced several algorithms for high-accuracy target geolocation. Demonstrated and ongoing work applies information-theoretic approaches to the tasks of area search and target pursuit, considering terrain constraints on target location and possible escape routes. Finally, we demonstrate previously unattainable high-bandwidth transmission of video and control of a UGV at multi-kilometer range from the operator using a UAV communications relay. This effort complements and builds upon existing work in air-ground robot mission management[1] and cooperative self-localization,[2] navigation,[3] target tracking,[4] target localization,[5] and pursuit-evasion[6] by applying formal approaches and experimental lessons to demonstrate the completion of practical missions with readily deployable stock systems.

SYSTEM OVERVIEW

The collaborative system developed is demonstrated using minimally modified stock systems presently fielded by military and law enforcement agencies. This is comprised of a collection of man-portable equipment easily transported by a small team that includes a single UAV, a single UGV, and a single OCU laptop. Scenarios involving larger numbers of robots have also been tested in simulation to validate the algorithms developed.

UAV Hardware

The UAV platform selected for this system is the RQ-11B Raven produced by AeroVironment, Inc and pictured in Figure 1(b). With over 10,000 units delivered, it is possibly the most widely deployed UAV system in the world and is heavily fielded by many branches of the US military as well as several domestic agencies. Relatively inexpensive and easily transported by a single person, the Raven can be hand-launched from a small clearing, thus serving as an appropriate vehicle to emphasize the capabilities of the described system.

The man-portable stock Raven system includes the following components:[7]

- 2kg (4.5lb), 1.5m (4.5ft) wingspan UAV airframe
- Modular nosecone with fixed (gimbal-less) forward and side-look cameras
- Ground control station (GCS) for low-level communications and flight control
- Antenna to wirelessly link GCS with Raven
- Hand controller OCU for teleoperation during launch and landing, if desired

- FalconView PC software for live flight control or simulation

Both the stock RQ-11B using an analog wireless link as well as the recently released RQ-11B DDL using AeroVironment's Digital Data Link (DDL) have been evaluated, the latter permitting extended operating range and communications relay capabilities. Neither model of vehicle required hardware modification or special interfacing, permitting the use of any fielded Raven with the system.

UGV Hardware

The UGV platform utilized for this effort is a stock iRobot Explosive Ordinance Disposal (EOD) Packbot chassis with a development payload and several additional sensors to enhance its capabilities. The UGV is pictured in Figure 1(a).

This vehicle, easily transported in a Humvee and deployable by a single operator, includes the following components:

- 18kg (40lb) EOD Packbot chassis
- Computation payload with 1.4GHz dual core embedded computer
- MicroStrain 3DM-GX1 IMU
- Antaris Ublox GPS
- Hokuyo UTM-30LX scanning laser rangefinder for mapping and obstacle detection
- 3D volumetric sensor (iRobot SEER Payload with Tyzx stereo vision system)
- TracLabs Biclops pan/tilt unit, with the following sensors mounted:
  - Sony FCBEX 1010 block camera with 36x optical zoom and 12x digital zoom
  - Opti-Logic RS800 laser finder for target ranging

Two modes of wireless operation were demonstrated. The first is built upon standard IEEE 802.11b WiFi, providing line of sight operation up to 300m. For longer range operation up to several kilometers (and over 10km using a UAV relay), a DDL module was outfitted. Experiments testing these capabilities are described further in a later section.

OCU Hardware

The entire system is controlled from a single OCU computer (typically a laptop as pictured in Figure 1(c)) containing all the software required to complete the air-ground collaborative surveillance mission. Depending on logistical constraints and the desired level of capability (number of vehicles, visual tracking performance, planning horizon, etc.), this may vary from an easily portable laptop to a server installation in the rear of a Humvee. In the experiments conducted for this effort, both the commonly-fielded ruggedized Panasonic Toughbook CF-30 and a faster commercial Dell laptop were evaluated, the latter providing a smoothly scaling increase in algorithm performance.

(a) (b) (c)

Figure 1. The UGV, UAV and the computing hardware that comprises this heterogeneous collaborative surveillance system. All components are man-portable and battery operated.

OCU SOFTWARE

The OCU provides a central point of control of multiple UAVs and UGVs and presents feedback from all robots for a unified view of mission state. This consists of a control/video window for each vehicle and one high-level mission planning and tasking window formulated as an overhead map view. The OCU is built on the iRobot Aware 2.0 architecture using the Python scripting language, allowing additional functionality and customizations specific to new types of vehicles to be added.

Map View Window

The map view window, pictured in Figure 2, provides a top-down view of the mission area and integrates all available information about the world for maximum situational awareness. Overlaid on a satellite view of the area are the locations of all robots, their recent motion path histories, current waypoints for all vehicles, operator-designated areas of interest (AOIs), estimated locations of detected targets, and target track histories.

Basic controls for the map view include the ability to pan and zoom the map to study different parts of the environment. Additional buttons enable the operator to add and alter one or more desired waypoints for any vehicle as well as designate an AOI for it. A tasking interface, further described later in this section, permits the operator to select a vehicle and direct it to perform a behavior such as following a series of operator-provided waypoints, autonomously searching an AOI, or persistently pursuing a target. When doing so, the operator also selects the relevant AOI or target, as appropriate, since the system supports multiple of each.

Vehicle Control Windows

Each robot receives its own vehicle control window, which displays live video from cameras onboard that vehicle and specific status information such as signal strength, battery life, and numeric location coordinates. This includes a graphical depiction of the vehicle's orientation to rapidly perceive any unusual or dangerous conditions such as large roll or pitch. Low-level teleoperation control using an attached joystick is also available for emergency operation or fine maneuvers an operator may wish to perform manually. Access to additional functionality specific to a given vehicle type is also provided via this interface.

UAV Control Window Control windows specific to UAVs, such as that shown in Figure 3, display realtime position, altitude, heading, speed, and a graphical depiction of orientation that provides an immediate way to perceive whether the UAV is pitched to gain or lose altitude or banking for a turn. Low-level controls to change the throttle level and model-specific functions such as the Raven's autoland behavior are presented. The primary operator interaction with this interface otherwise is to initiate and monitor automatic visual tracking and geolocation of targets described in a later section.

UGV Control Window Control windows specific to UGVs, such as that shown in Figure 4, present live position and heading status, along with a graphical depiction of orientation that efficiently conveys whether the UGV is moving up or down an incline. Packbot-specific information includes battery status for each installed battery pack and distance to an object at the center of the camera's view as reported by the onboard target rangefinder. Also provided is a standard teleoperation interface validated in extensive operator trials and additional low-level control such as whether the vehicle's brake is engaged, whether obstacle detection is enabled, the position of the Packbot's flippers, and a simple interface to the pan-tilt-zoom camera that enables surveillance from a safe standoff distance.

Joint Tasking

Using the map view window, an operator may direct any agent to perform one of several tasks. Rather than requiring tedious teleoperation for each action, these tasks form a set of semi-

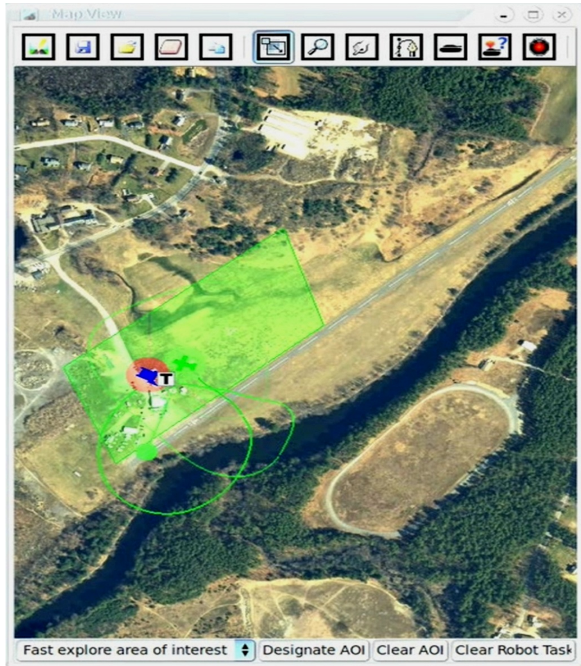


Figure 2. A screen capture of the map view window presenting consolidated mission state, here a partially completed search of the green-highlighted area. The operator can see current locations of all robots (UAV in green at the tail of the green trace denoting its recent path and the UGV in blue within the red circle) and targets overlaid on a satellite image and can initiate a variety of autonomous control tasks.

Figure 3. A screen capture of the UAV control window, containing its live wide-area video stream beneath realtime UAV status. An operator may monitor the video stream for targets and designate them for visual tracking using a single mouse click. Tracking and geolocation options are accessible via menus.

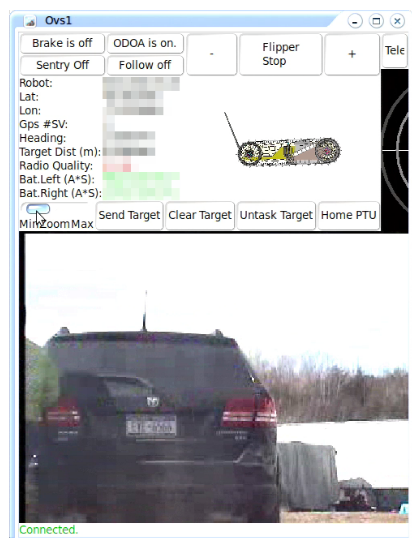


Figure 4. A screen capture of the UGV control window. The UGV's close-up video stream is displayed live beneath realtime status information. An operator may manipulate the pan, tilt, and zoom of the video, designate targets for automatic tracking, and perform standard teleoperation.

autonomous mission primitives that may be used to build up a complex practical mission. The underlying details of how each task is executed may depend on whether the selected robot is an air or ground vehicle, including which model thereof, but the collaboration system handles these without further effort by the operator.

Idle - The default task that is always active when no other task is in progress. When selected, the robot attempts to maintain its current view, corresponding to a stopped state for UGVs and a loiter pattern for UAVs.

Follow waypoints - The vehicle attempts to move through the indicated series of waypoints and idle at the last one.

Perimeter follow - Generates a loop of waypoints from the perimeter of a selected AOI and continuously traces it using the waypoint follow behavior.

Search - Executes a coverage pattern over an AOI to search for a potential target of interest using one of the algorithms specific to the vehicle described in a later robot capabilities section. At any time during the search, the operator may designate a target visible in the vehicle's video feed and, if desired, initiate a pursuit behavior.

Pursue target - Attempts to keep the specified target in view, moving the camera or vehicle as necessary to accomplish this. Several methods for estimating a target's location and moving to follow it are available and further described in later respective UAV and UGV capability sections.

Follow other vehicle - Causes the selected vehicle to maintain view of a secondary selected vehicle using the same algorithms as for pursing a target, should the operator wish to further study the area surrounding the second vehicle or if the second vehicle has entered a dangerous area in which it is vulnerable.

The remainder of this paper presents more detailed descriptions of the specific capabilities upon which these tasks are built, as well as broader missions these enable.

Firestorm Interface

The described collaboration system provides an interface to the Firestorm battlefield network developed by ARDEC using the Cursor on Target (CoT) protocol. This allows it to link with vehicles operated by other teams or a command center to broadcast vehicle position updates or target spot reports and accept remotely-provided AOIs or waypoint sequences on which vehicle tasks may be initiated. Other capabilities include responding to remote image requests from any vehicle's live video feed and providing battle damage assessment.

UAV CAPABILITIES

Implementation of the aforementioned subtasks required differing development attuned to the complementary nature of UAVs and UGVs, described in detail in this and the following section. The algorithms developed were optimized for and validated on the Raven and Packbot respectively, however the underlying designs may be applied to broad classes of each type of vehicle with minimal modification, primarily to interface with differing forms of telemetry streams.

UAVs of the size and cost appropriate for deployment by small forward teams are inherently not designed for the levels of precision targeting commonly provided by much larger drones with extensive sensor suites. In particular, such UAVs typically provide low-resolution video downlinks, use much simpler and less accurate GPS and inertial navigation sensors, broadcast telemetry at lower rates, are easily buffeted by wind, and provide limited and low-frequency control commands from the ground station. Many of these issues can be addressed by the addition of sensors to the vehicle or by performing targeting and trajectory planning onboard, however a major goal of this effort is to leverage the robustness of already-fielded hardware and demonstrate the ease with which capability can be added to existing systems without a costly hardware retrofit that would require the development of a different payload for each type of UAV. Further, many of the same limitations apply with onboard processing, merely to a lesser extent.

The algorithms developed fulfill several key aspects of a mission. These include providing an operator a low-effort way to designate and visually track targets from a UAV's video stream, geolocate these targets in the environment, pursue targets if they move, and search an area for unknown or previously-lost targets. As implemented, the software operates on video and telemetry adhering to the Motion Imagery Standards Profile (MISP)[8] produced by the Raven ground station and will operate with at most minimal modification on any MISP-compliant feed.

Visual Target Tracking

The primary external sensor assumed to be present on small UAVs is one or more cameras, as others such as radars and laser rangefinders are generally of prohibitive size and cost. Therefore, the targeting procedure envisioned is that as a UAV covers an area of interest, targets visible in the video feed are designated by an operator click or an automatic detection algorithm, these designated targets are then tracked through succeeding video frames without further operator intervention by one of several algorithms, and a tracking recovery procedure is executed to attempt to re-acquire targets that have become occluded by terrain or that have left the view of the camera. The system developed handles high levels of camera motion caused by wind, performs out-of-view tracking recovery that is frequently required in practice, and handles tracking and recovery through video corruption or blackouts caused by communications glitches.

Image Stabilization Possibly the greatest challenge in visual tracking from a lightweight UAV is rapid unwanted camera motion caused by wind-induced attitude disturbances. This causes an objects location within the image to change rapidly, creating failures in many traditional tracking schemes and also making it difficult for a human user to initially designate a target in the video stream. Therefore, the first stage of our tracking system performs image stabilization of the live video stream. This scheme operates with a planar ground assumption, which is not strictly true, but given both the altitude of the UAV is much higher relative to the undulations on the ground and also the rotational velocity of the UAV dominates translation, this assumption is approximately correct. Two neighboring image frames are analyzed to find the planar transform that minimizes some objective function. For us, the objective function is a global image difference. Global image alignment is more robust in such featureless scenes, where traditional approaches that track point features are more brittle. The Lucas-Kanade alignment[9] is one of the formative image alignment approaches. However, this conventional approach works via gradient descent and fails with large amounts of image motion, often the initial estimate is too far away from a global optimum, and the gradient directs the algorithm into a local optima away from the correct answer. We avoid local optima all together by sampling a discrete set of estimates from the objective function and choosing the optima from a global set.

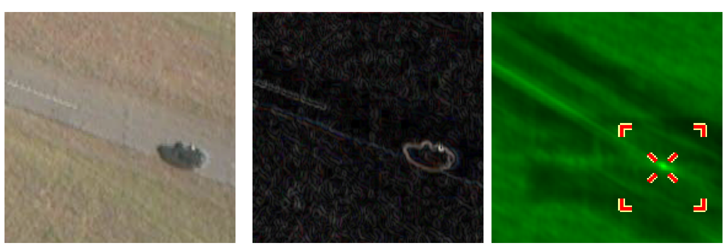
Target Designation In contrast to typical existing systems that provide for target designation in the form of highlighting a target in previously captured imagery or by using a joystick to continuously aim a gimbal or cursor, the system developed here provides a simple means for operator-assisted target designation. First, the operator enters target designation mode, during which the live video stream is stabilized as described to produce a synthetic view as though the UAV were not moving. Next the operator may either draw a box around the target that is then automatically sized to best fit the identified target or click on the center of the target, causing a fixed-size box to be placed around the target and then likewise resized. Resizing is performed by initiating one of the following tracking methods and sampling bounding box sizes to find the one providing the best separation of visual appearance between area within it and the surrounding area for the tracker. Once a target is designated, an automatic visual tracking method is applied to identify the target's location in succeeding video frames without the need for the operator to continuously highlight it. We have evaluated two such methods, both relying on the use of a template of the target generated from the initial video frame, and we describe and compare these methods in the following sections.

Mean-shift Color Tracking The first of these builds a color model of the target in the form of a histogram of color values appearing in the region initially designated as the target and finds the nearby region in succeeding video frames whose color distribution most closely matches the template histogram. We use a direct implementation of this concept[10] as well as a variation[11] that continuously reweights a set of features derived from functions of color values (encoding features such as intensity and chrominance) to explicitly maximize the template's discrimination between the target and the surrounding background. We can configure between both the direct and modified implementations and use the well-known mean-shift algorithm to onverge on the location of the target, starting at an initial guess location lying where the target lay in the previous image.

Spatial Patch Tracking An alternative approach to visual tracking is to construct a spatial appearance template of the target, treating it as a patch in the image that may move between video frames. We develop a few modifications to standard patch tracking methods, probably the most well-known of these is the KLT algorithm.[12]

The first, albeit slight, modification of our approach is to use a Sobel transform version of the template image as input to our objective function, depicted in Figure 5(b). The Sobel transform makes the approach more robust to fast illumination and fast exposure changes. The second modification provides greater robustness to large camera motions, because the KLT framework typically assumes that the deformation of the object's appearance will be minimal. However, in our case the camera motion is unusually large, and often the image motion can be on the order of the size the template itself, causing the KLT linearization for gradient descent to be meaningless and the translation solver to fall into an incorrect local minimum. Our alternative approach densely samples possible inter-frame motion parameters and chooses the best fit to the current image, avoiding local minima at the cost of some additional computation that in practice is insignificant given that one target is tracked per video stream. Figure 5(c) shows an example of the template comparison function output over the entire parameter space, with the best location designated by a crosshair. There is a notable peak at the optimum, but the peak is very narrow with other minima in the search window that gradient descent type algorithms are likely to get trapped inside of when the camera motion is high.

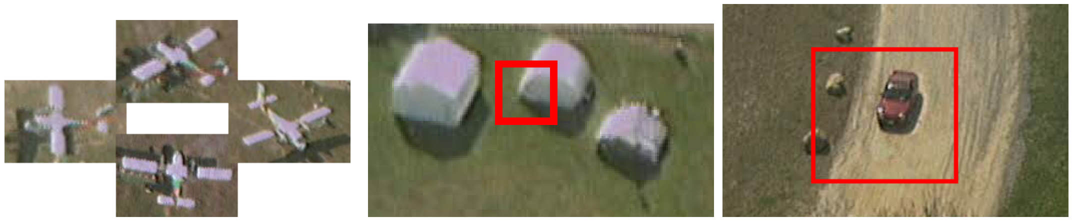
As the UAV flies around a target, the perspective change causes a change in the targets appearance, therefore we update the template of the target over time. The template update problem is a well documented[13] problem, requiring a compromise between adapting to the current appearance of the object and unwanted model drift. We update the template while avoiding the drifting problem by adjusting the template by mixing it with the current patch.



(a) Raw image (b) Sobel image (c) Tracking likelihood

Figure 5. Examples showing the spatial patch tracking algorithm, search for the patch within a region.

Tracker Comparison Clearly, there exist scenarios in which each type of tracker may perform better than the other. Bright, distinctly-colored targets are best tracked using a mean-shift color tracker, while the patch tracker performs well with indistinctly colored but consistently-shaped targets. More subtle strengths and weaknesses are present, however. For instance, as shown in Figure 6, the patch tracker requires continuous tuning of the template to handle changes in perspective that non-spatial mean-shift algorithms are largely unaffected by and performs poorly if too much background that will not move with the target lies within the template, while environments with similar objects confound non-spatial mean-shift. On average, the patch tracker performs best and is selected by default. At any time during a mission, the operator may choose an alternate tracker if appropriate.



(a) Changing perspective (b) Similar objects (c) Background in template

Figure 6. Examples showing cases of challinging tracking situations. (a) and (c), for which non-spatial tracking performs better, and another, (b), for which a spatial template is more appropriate.

Tracking Recovery Inevitably, a tracked target will often leave the view of a UAV's camera due to wind buffetting. We recover from out-of-view tracking losses by estimating the camera motion and predicting the new location of the target, and when the predicted location re-enters the view, we attempt to search for the target inside a sizeable search window. In addition to out-of-view events, other factors may induce tracker failure. For instance, analog and digital radio communications loss causes corruption in the video stream. We detect failures in the tracker when the template no longer matches well with the search window and then attempt to recover the target once the period of corruption is over. An example of this procedure applied to both video corruption and an out-of-frame event is shown in Figure 7.

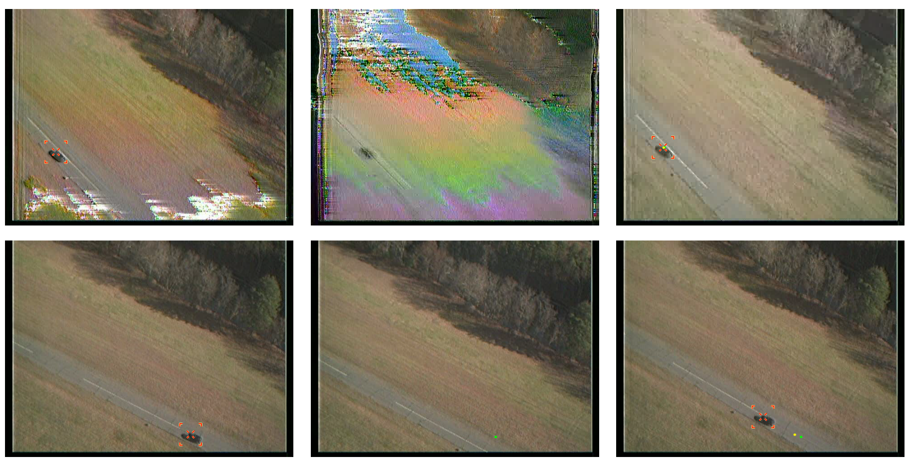


Figure 7. Different types of tracking recovery. The top row shows recovery after a period of communications corruption. The bottom row shows recovery after the target leaves the field of view.

High-accuracy Target Geolocation

Once a target is being observed within a UAV's video stream, the next task in the mission is to geolocate the target in a global coordinate frame. This task is greatly complicated given the limited-fidelity sensing aboard the class of vehicles considered here. Yet, for the collaborative tasking proposed in this paper, it is essential for estimate accuracy to be sufficient for another vehicle to reach the true location of a target if directed there. We refine existing and develop new strategies for target geolocation in the presence of large UAV state uncertainty.

Two broad approaches exist for mapping a pixel coordinate to a ground location: either registering live UAV video frames to previously geo-referenced aerial imagery [14] or using an estimate of the UAV's pose to generate a ray in world coordinates that is intersected with previous observations [15] or a prior terrain model. [16] Given the fragility and heavy computational requirements of the prior, as well as the typical availability of a compact Digital Terrain Elevation Database (DTED) for a mission area, the latter approach of ray intersection with a terrain model is adopted.

Given a series of highly uncertain observations of a target in world coordinates, a filtered estimate is desired. If the target is known to be stationary, these observations (e.g., from an orbit) may simply be batch-combined to remove symmetric biases in observations,[17] however in general, various forms of the Kalman filter framework have been proposed[18,19] with varying levels of success. A major shortcoming of existing geolocation literature is the lack of discussion of how observation uncertainty is modeled given uncertainty in UAV pose, for instance in some cases a ground measurement is modeled as a Gaussian distribution with fixed and arbitrary covariance. Yet, the observation function mapping image to world coordinates is highly nonlinear and greatly sensitive to UAV orientation, and state errors may be non-Gaussian.

Stationary Geolocation Filter Taking inspiration from both particle filter techniques, which can handle arbitrary observation functions and large uncertainty, and batch methods, which can solve for systematic biases in observations, the authors previously proposed a strategy[20] that estimates a stationary target's location using an accumulated evidence grid populated by sampling along each axis of state uncertainty (vehicle pose, camera pose, camera calibration, etc.) for each observation. An example of this filter in action is shown in Figure 8. This filter has proven to work extremely well in practice, handling non-symmetric biases in state caused by, for instance, time-varying compass heading error up to 30° that is nearly uniformly distributed.

Additionally, another filter was developed[21] by drawing from ideas previously applied to range-only localization problems. Rather than representing target location as a sampling of points, it is able to parametrically represent the crescent-like shapes of uncertainty distributions that typically occur from a series of highly bearing inaccurate observations. This allows efficient storage of distributions and fast evaluation of target probability at a given point, easing the system's expansion to larger numbers of targets and robots by reducing computational load and minimizing the amount of data that must be transmitted over the network.

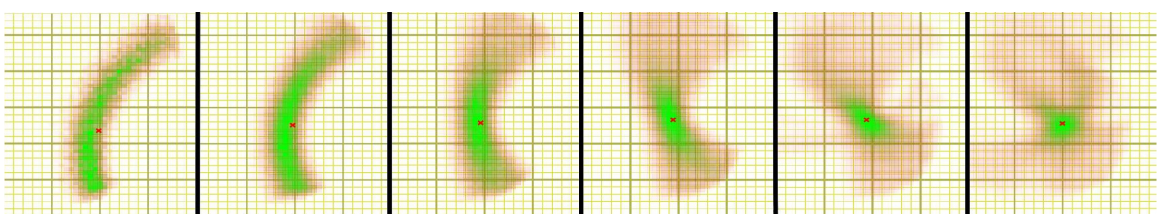


Figure 8. Visualization of the stationary target geolocation filter over a UAV's orbit of a target, in chronological order, left to right and top to bottom. The filter is a probabilistic 2D grid of cells (the two axes are the North and East location of the object). Cells vary from bright green (high target probability) to red (low probability) to blank (none). Scale: large grid cells are 50×50 meters.

Moving Target Filters For moving target geolocation, various forms of Kalman filters were initially evaluated. In general, these proved inadequate primarily due to their limited representation of target uncertainty as a mean location and Gaussian covariance, both for individual observations and the best estimate. Studying this, we observed that the primary source of geolocation error is inaccurate orientation, particularly heading, confirming the intuition behind the developed stationary target filter. Therefore, a filter was developed that represents observations as an uncertain range and bearing relative to the reported UAV position. Because of its ability to parametrically describe observations of the crescent shape merged in Figure 8, this representation has shown itself[22] to be superior to observation uncertainties produced by Jacobian linearization of the pixel to ground observation function or the use of the Unscented Transform,[23] as would be implemented for a conventional Extended or Unscented Kalman Filter (EKF or UKF), respectively. The filter itself is a conventional particle filter, estimating position and velocity of the target, with each particle representing a sampling of this space under the uncertainty in the current target state estimate. This is able to accurately incorporate range-bearing observations by reweighting particles to well approximate the oddly-shaped uncertainty distributions that may arise from a series of such observations. In practice, even without careful tuning, this filter provided considerably greater geolocation accuracy than a comparable EKF or UKF.

Because moving targets increase the challenge of producing accurate location estimates by providing fewer observations at a given location, taking advantage of natural terrain constraints to reduce the space in which a target might lie and avoid considering areas in which it cannot was applied as a simple yet highly effective means of reducing the impact of target geolocation uncertainty. The two primary such constraints are impassable environmental regions such as thick swamps or brush, and the case of a target vehicle that is non-evasive (oblivious to surveillance) or assumed to remain on a road because it is ill-equipped for off-road travel. Such information about the environment may often be known a-priori from topographical or infrastructure maps and changes slowly over time, unlike the mere appearance of a satellite map. Several experiments were run[22] applying such constraints, for instance by geometrically projecting observations of a target on a road to the nearest road or to the nearest cell of traversable terrain, while constraining location estimates or samples to lie within these as well.

Target Pursuit

Once an estimate of target location has been formed, it is typically desirable to maintain surveillance to observe its actions and follow it to provide continuously updated location estimates to other vehicles. Directing a UAV to pursue a selected target via the map window initiates this task.

Orbiting Predicted Best Estimate Intuitively, a UAV need merely stay on top of a target to pursue it. This is relatively straightforward for UAVs with a gimbaled camera, which need only stay in the vicinity and train their camera on the target, or those such as the Raven with a side camera, which can be directed to orbit the target, thereby keeping it in view.

Thus, the first pursuit strategy is to continuously transmit the best estimate of target's location, updated with each observation of the target, as the desired orbit location. For moving targets, this strategy has the drawback that due to control latency and half of any orbit's motion necessarily being in the opposite direction of the target's, it will tend to lag behind the target, increasing the risk of loss. To compensate, a configurable look-ahead time offset (typically 10 seconds) is added, so that the requested waypoint is the target's location predicted by the geolocation filter at this time in the future, ideally lying slightly ahead of the target to increase its proportion of time in view.

Uncertainty-minimizing Trajectory Planning Simply orbiting the predicted location of a target has two serious limitations. The first is that areas other than the most likely point of target location are observed only incidentally as the UAV moves to center that location in its view, ignoring other regions with nontrivial likelihood it may be aware of due to geolocation uncertainty or multiple hypotheses that may result from a target nearing an intersection. The other is that moving so as to keep the target in view instantaneously does not necessarily place the UAV where it can move to continue see the target in the future, given the typical motion constraints on UAVs such as a minimum turning radius and a minimum airspeed. Addressing either of these requires moving in a deliberative fashion, predicting both target motion and considering possible UAV actions that are likely to provide good observations in the future.

Choosing UAV actions that are likely to provide future benefit is accomplished using a trajectory planning framework that applies principles of information-theoretic pursuit.[24] This generates a sequence of turn or bank angles that optimizes some objective such as maximizing the expected proportion of time the target is likely to be in view, minimizing the expected geolocation estimator's uncertainty at the end of the trajectory resulting from predicted observations, or maximizing the probability of seeing the target at least once during the trajectory. While developing this system, a study was performed[22] evaluating several such metrics while comparing various geolocation filter representations. The primary conclusions from this effort are that applying terrain constraints greatly improves performance by channeling predicted target motion and reducing the area the UAV must plan to cover, that maneuvering to acquire observations that minimize geolocation filter uncertainty (particularly in target velocity) while sacrificing short-term view of the target outperforms trying to keep the target in view because it provides better predictiveness during inevitable observation outages, and that anything beyond short-term trajectory look-ahead provides rapidly diminishing returns given a tendency for the predicted target uncertainty distribution to flatten into a wide uninformative area. While promising, field experiments with the Raven exposed the challenges of planning in a less-agile action space such as waypoints and orbit points as provided by its autopilot and needing an accurate model of UAV response to these commands. Ongoing work in this area seeks to develop compact representations of multiple target location hypotheses and perform higher-level reasoning about target intent and capabilities to better infer likely target routes and reduce the dependence on a consistent stream of observations.

Perimeter Protection and Area Search

The final desirable capability of UAVs is to minimize the operator effort required to locate targets (search and coverage) or verify that none are present (area denial or perimeter security). Towards this end, the developed system provides UAV-attuned implementations of the general semi-autonomous perimeter following and area search tasks available to an operator.

Waypoint-based Perimeter Following Given an operator-designated polygonal AOI, this continuously directs the UAV in a loop consisting of the vertices of the polygon. If the UAV is equipped with a gimbaled or forward-pointing camera, the direction around the loop (clockwise or counter-clockwise) is typically irrelevant as the area surrounding the perimeter may be observed continuously. Otherwise, when for instance using a side-pointing camera, the direction may be chosen depending on whether detection of targets entering or leaving the area is desired.

Geometric Coverage Area Search For area search, a method was developed to rapidly and thoroughly cover an area while making minimal assumptions about the UAV's agility and available control actions. This takes the rectangular convex hull of an operator-designated AOI to produce the smallest bounding rectangle, then directs the UAV along smoothly precessing orbit points that themselves follow a lawnmower-like pattern that winds through the area. In addition to reliable behavior across various autopilots, this has the advantage over simpler waypoint-based methods that as the UAV covers the area, any given point on the ground is viewed from multiple angles and is in view roughly continuously until overflown, easing the task of operator or automatic target designation. An example of this algorithm in action is shown in Figure 9.

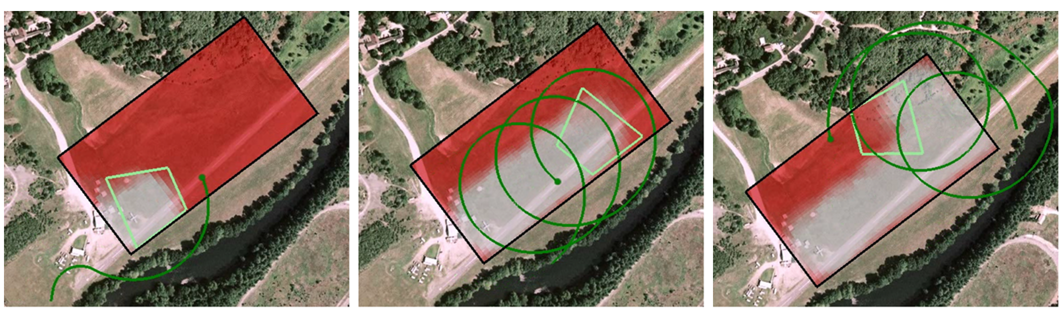


Figure 9. Snapshots during an execution of a geometric area of interest (AOI) search. The UAV is directed to a series of precessing orbits weaving through the bounding rectangle of the AOI. Here, red areas indicates unobserved terrain, progressing to white for adequately observed (cumulatively in view for approximately 10 seconds).

Information-theoretic Area Search While highly appropriate for uniform open area, geometric coverage search performs unnecessary work for irregularly-shaped AOIs or AOIs containing holes (impassable or previously-searched regions). Further, it makes no use of prior knowledge to prioritize searching areas more likely to contain a target. This was improved upon by applying information-theoretic search techniques similar to those used for pursuit, in which the probability of target presence is stored by a cellular map, particles, or continuous parametric distributions and updated with each observation (or negative observation).[25] Appropriate motions for the UAV can then be formally reasoned about, for instance by either following the gradient of target probability or planning a trajectory that maximizes the probability of detecting a target (intuitively, scooping up the most probability). Ongoing work applies this specifically to road networks, whose sparseness and topology lend themselves to irregular search patterns and UAV trajectories that stem the likelihood diffusion of a previously-observed moving target. Figure 10 provides an example of this search-to-capture approach applied to a simulated live pursuit to resolve intermittent target losses due to uncertainty and sharp turns the UAV cannot match.

Figure 10. An example of information-theoretic pursuit of a fleeing target in a road network. The UAV anticipates loss of view (area within the green trapezoid) and begins to turn sharply. While out of view, it predicts diffusion in possible target location (red probability mass) and chooses a search trajectory that maximizes the probability of re-detection in the shortest time.

UGV CAPABILITIES

The Packbot UGV is equipped with a similar suite of capabilities as those developed for the UAV. Unlike the UAV, the UGV has much more relaxed power, size and weight limitations but must be able to navigate cluttered environments with physical barriers to navigation in potentially GPS-denied areas. To provide the same capabilities present on the UAV, a computational payload and sensor payloads have been added to the Packbot. Additional capabilities on the UGV include obstacle detection and avoidance, indoor map building, localization, and path planning.

Navigation and Indoor/Outdoor Localization

The UGV is equipped with an IMU, GPS, 2D LIDAR, and 3D stereo camera for navigation purposes. Rotation estimates through encoders on skid-steer platforms is very inaccurate. Typical errors seen on standard road surfaces were approximately ± 15 degrees after executing a 90 degree rotation. To overcome the inaccuracies present in the encoded odometry, an improved odometry is calculated by integrating the track odometry distance with the gryro estimate of rotation around the gravity vector. Outdoors, the odometry estimate is combined with GPS to establish a global position. Indoors, if a globally referenced map had been established, the UGV also uses the 2D range scans to localize itself within the map.

Waypoint navigation is also implemented on the robot, with map-based path planning when indoors. At all times, obstacle detection and avoidance is present through a local 3D model made by combining the 2D LIDAR and 3D stereo camera.

Persistent Target Inspection

A 30x zoom camera and single-point laser ranger are mounted on a pan-tilt unit on the UGV. These sensors enabled long distance persistent stare capability on the platform, as well as input for target geolocation and tracking. Under the described configuration, in a persistent stare scenario the UGV was capable of operating for up to six hours.

Onboard Visual Tracking

Tracking from a ground vehicle is significantly different than tracking from a UAV. The UGV has higher quality video and typically tracks targets that take up a larger area of the video feed. However, tracking must now deal with a much more dynamic background, motion blur due to high frequency motion of both the platform and target, and the potential of significant change in object appearance due to changes in scale and physical rotation. In the tracking system developed targets within the video feed are designated by an operator and automatically tracked through succeeding video frames. Recovery is attempted when targets are lost due to occlusion, leaving the field of view, or due to dramatic changes in their appearance. A brief overview of the algorithm developed for this system is presented here. The algorithm combines two fast, non-parametric algorithms, a modified color histogram tracker and a Speeded Up Robust Features (SURF)[26] tracker to maximize their strengths and mitigate their drawbacks.

Collins Adaptive Mean Shift Tracker The first algorithm used as a foundation of the tracker is based on the mean-shift algorithm proposed by Collins.[11] A mean-shift tracker offers fast and robust tracking well suited for embedded systems. However, typical mean shift trackers are susceptible to occlusion and are often distracted by background areas similar in color to the foreground object. The extension proposed by Collins significantly improves performance in such situations by performing mean-shift on synthesized images derived from each video frame.

The tracker attempts to achieve the highest discriminative power between the object and the background by selecting the N best channel sums of every frame using the two-class variance ratio of the log likelihoods of each histogram bin. Mean-shift is performed on these channel combinations in the next frame, and the median of the mean shift results are taken to get an initial estimate of the kernel location.

Robustness Enhancements The approach suggested by Collins is to individually use each likelihood image as the source image to run mean shift. However, resizing the kernel on each likelihood image will naturally result in a focus on the peaks within the likelihood images. Therefore, a naive Bayes classifier was constructed using the likelihood images to create a mapping between each pixel coordinate and its likelihood of being part of the tracked object. The Bayes classifier allows for a more accurate prediction of kernel size by taking into account the inputs of all likelihood images.

Bias Model Further improvements are made on the tracker through the addition of a bias model when calculating the centroid of the object and an adaptive kernel resizing method to improve on scale and aspect-ratio invariance. Mean-shift struggles to track multimodal distributions as shown in Figure 11. The introduction of a bias model helps to counteract the tracker's desire to focus on only a part of a tracked object.

Figure 11. Unmodified Collins tracker with a multimodal distribution. In 30-60 frames the tracker has abandoned the top color.

Rotational and Scale Invariance SURF is the second algorithm used in the combined tracker. Unlike the adapted Collins tracker previously described, SURF provides a true scale and rotation invariant detector and descriptor,[26] but only provides rotation invariance in the image plane. Physical space rotation causes new features to be introduced with no provision for determining which features belong to the object versus the background. Figure 11 shows a typical SURF result when an object undergoes physical rotation.

Combined Tracking Algorithm The combined tracking algorithm employs the adaptive mean shift tracker on each incoming frame. A set of checks for the mean shift tracker are used to assess the likelihood that the object has been lost. Detection of track loss is based on a loss of discernibility in the log-likelihood images between foreground and background or sudden changes in the estimated position, scale, or aspect ratio of the target. SURF is used to buttresss the color tracker when background-foreground discernibility is particularly low. When potential tracker loss is detected SURF is used to find it again.

The two trackers combined in this manner are very good complements in performance and tracking quality. Color histogram tracking takes significantly less computation and is rotationally invariant, but prone to drift. SURF is scale and aspect ratio invariant but computationally expensive and not rotationally invariant. The overall algorithm meets the goal of maximizing real-world robustness by combining strengths and mitigating complementary deficiencies.

Figure 12. Examples of the combined tracker tracking a moving vehicle through a continuous sequence. The tracker is adapting to numerous changes in size, shape, and appearance

Operator Interaction Operator interaction with the tracking system is kept simple and similar to the UAV. High quality video is captured locally on the UGV computational payload. The video stream is downsized, compressed, and then transmitted to the OCU. The operator in the lower quality video stream simply draws a box around the target that initiates the tracker. The box position and size then update over time based on successive results of the tracker as new images are captured. A button exists for the user to then turn on target pursuit. If the UGV's brake is engaged, then it will attempt to keep the camera centered on the target by controlling the pan tilt unit. If the brake is disengaged then a combination of the tracks and pan tilt unit will be used to stay centered on the target and the robot will engage in pursuit of the target attempting to maintain a user defined stand off distance with distance being acquired through the calibrated single point laser ranger present on the pan tilt unit.

JOINT CAPABILITIES

The system allows for a wide array of potential mission scenarios. One of the goals of the project was to provide a framework for UAV/UGV collaboration that can be adapted to current mission needs and goals. A set of joint capabilities have been established to serve as a foundation and example of what can be done with a unified UAV/UGV communications and control architecture. Within this context, there is no discerning between a UAV or UGV as there is a common set of state feedback and autonomy employed by all assets and this information is shared between all vehicles. The common set of capabilities this enables are:

- Waypoint tasking
- Area of interest search
- Pursuit of other UAV/UGV targets
- Cuing to inspect any tracked target
- Follow other UAV/UGVs
- Stand-off stare at other UAV/UGV targets

Persistent Tracking and Pursuit

Using this set of capabilities, a wide range of mission scenarios are enabled, with several tested in front of audiences presented in the following sections. In general, the mission flow this system excels at is one in which an operator observes or is notified of events or targets of interest visible to the sensors onboard one of the robots, the operator designates or acknowledges this point to begin geolocation, and other robots or teams are directed to this location as desired. As other nearby points of interest are detected, or if tracking is lost due to rapid motion or environmental occlusion, any robot to which it is visible can be used to re-designate it and re-cue other vehicles.

Long-range Communications Relay Using DDL

The Small Unmanned Airborne System Digital Data Link (SUAS DDL), developed by AeroVironment and the SOCOM SUAS ACTD, allows the Raven to relay control and payload data to and from the PackBot over a low-latency bi-directional 3 Mbps digital link. In addition to video, command and telemetry channels for the Raven, the DDL offers an Ethernet bridge capability that allows it to be used as a control link for the PackBot or other UGVs, either for direct control between the UGV and its OCU or via relay through an overhead UAV and the OCU. By acting as an ethernet bridge, the DDL can also be easily integrated into other UAVs and UGVs that have a digital communications architecture, enabling the use of large heterogeneous teams over wide areas.

Use of the DDL to relay a UGV command link through a UAV allows the UGV to operate beyond the direct Line of Sight (LOS). This is applicable to both long range mission scenarios, far exceeding current capabilities of approximately 800m, and to urban operations where buildings sever operator line of sight. Currently, UGVs can't even be used on the other side of most small buildings.

Bandwidth allocation between members of the network can also be dynamically adjusted depending on mission needs. Both the UAV and UGV have the ability to throttle their video quality based on the bandwidth allocated in the DDL network. This allows the UAV to have high-resolution video during an area search or tracking, while switching to low-resolution when acting as a communications relay. Similarly, the UGV can provide high-resolution video when doing close-up inspection of a target, and low-resolution when navigating to waypoints or performing a persistent stare.

UGV Integration The AeroVironment DDL module is integrated into a single wide PackBot enclosure. The DDL module's Ethernet bridge has been interfaced to the Ethernet supplied by the PackBot payload port. This effectively establishes the DDL module as a network bridge between the PackBot's wired Ethernet and the OCU's wired Ethernet. Similarly, the OCU attaches to an Ethernet switch that is attached to the AeroVironment DDL antenna. The DDL modules are configured through an XML-based API designed by AeroVironment. Once the DDL modules have been configured and joined a DDL session, the DDL is completely transparent to both the OCU and PackBot. It appears as if they have been directly connected together over wired Ethernet.

The DDL network defaults to being configured for a relay scenario. However, the PackBot on joining the DDL network inspects the current members. If no Raven is found, it will assume ground-ground communications and configure the network appropriately. The interface also supports the ability to dynamically control the allocated bandwidth for the different nodes in the DDL network. Similarly, the UGV has the ability to adjust its video quality based on the available bandwidth to ensure that maximum video quality is received.

FIELD-DEMONSTRATED SCENARIOS AND RESULTS

The described system was evaluated in tens of hours of field trials, at several test sites. Varying wind, lighting, and terrain types including open fields, thick foliage, and desert exercised the robustness of the proposed algorithms. For rapid offline algorithm refinement and study, video and telemetry streams were captured that could then be played back in simulation, allowing complete validation of any changes prior to execution on hardware. Descriptions, screenshots, and experimental observations for several field-tested scenarios are provided here.

IED Inspection Scenario

In the first scenario considered, the mission is a response to reports of suspicious activity in an area that may include improvised explosive device (IED) emplacement. Due to the danger of traversing the uncleared area, the reconnaissance team takes up a position at a safe location just beyond it, and launches a UAV that is directed to perform a search of the area. During the search, the operator notices a suspicious package at the side of a roadway and initiates automatic visual tracking of it to produce an accurate geolocation estimate. Based on the cleared region the UAV provided up to this point, a safe path for a Humvee is taken near to this point, and a UGV is dropped off in the vicinity. The operator, now with the UGV under his command as well, instructs it to approach the location reported by the UAV, provide close-up imagery from a safe stand-off distance, and move around it as desired. Based on the results of this inspection, an appropriate response to the threat and/or a resumption of the area search may be taken.

Examples screenshots from field trials of this scenario are shown in Figure 13. Using the Raven, approximately one orbit of the target is needed to provide an accurate location estimate for the UGV, coincidentally providing the operator a view from all sides to better study it. Using the zoom camera onboard the UGV, a standoff distance of up to 20m allowed it to assess detailed features of the simulated package.

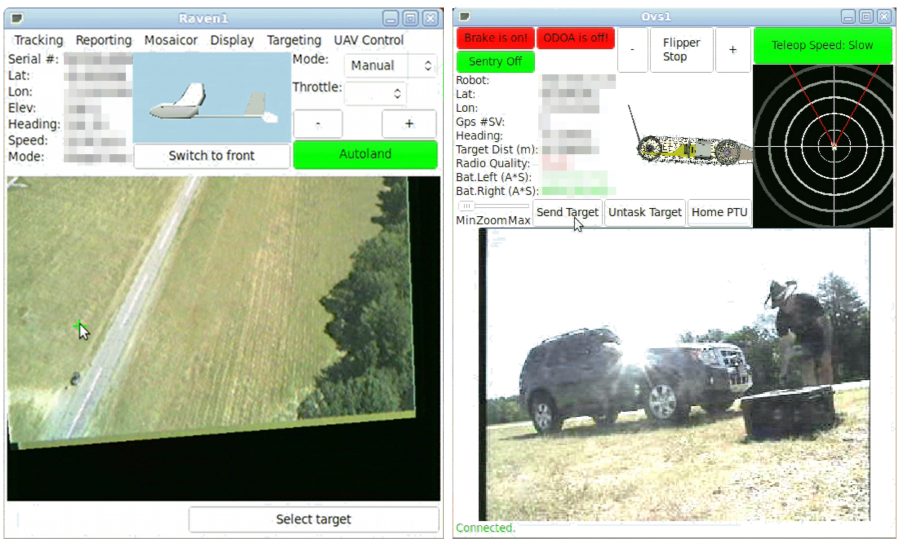


Figure 13. Screenshots of UAV and UGV control windows during an example IED inspection scenario. The operator designates what appears to be a suspicious package near a vehicle in stabilized incoming UAV video *(left)*, and its coordinates are transmitted to a UGV which moves to a nearby standoff position to provide more detailed imagery *(right)*.

Stake-out Scenario

As a substantial extension to suspicious package inspection, an alternative field-tested scenario is a more elaborate UAV/UGV urban stake-out mission that highlights many of the capabilities developed under the project. In this mission, the reconnaissance team seeks not only to assess a potential static threat but also to trace its origin and identify additional suspects that may be involved in planning other attacks. To accomplish this, the common operator interface, DDL communications relay, visual moving target tracking, target geolocation, and UAV/UGV tasking components were combined to collaboratively search for, geolocate, pursue, and inspect multiple ground targets. Screenshots from such a scenario are provided in Figure 14.

This scenario involves the following stages:

- A UAV passively acts as a communications relay to a UGV forward-deployed near a road
- The operator designates a search area along the road and initiates UAV coverage search

- The UGV is tasked to maintain persistent stare on any UAV targets to be discovered
- The operator designates a moving target along the road when the UAV passes over it
- Initial target location and velocity estimates trigger UAV moving target pursuit
- Continuous visual tracking during pursuit refines the target estimate
- The UGV watches the target vehicle pass and obtains a visual on the vehicle's occupants
- Once the vehicle stops, the UGV is tasked to its location for close-in visual inspection
- The UAV is tasked to loiter to maintain communications relay to the UGV
- The operator designates a dismount target seen leaving the vehicle and entering another
- The UGV pursues the moving target to the secondary vehicle and reports its location
- The UAV is cued to return for pursuit of secondary vehicle

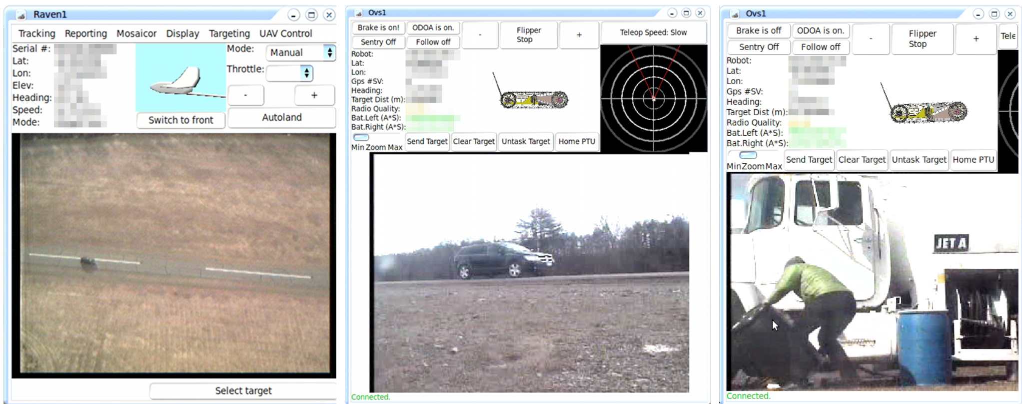


Figure 14. Screenshots of UAV and UGV control windows during an example urban stake-out scenario. The UAV pursues a moving target along a road *(left)*, during which a nearby UGV is directed to the road to provide a closer view, automatically panning its camera as the target passes by *(center)*. Once the target stops, a UGV is tasked to observe its actions *(right)*, in this case a rendezvous with a second vehicle and a hazardous materials transfer. UAV relay capability allowed the operator and UGV to be separated by a number of buildings.

Long-range UGV Remote Control

Long range UGV remote control was demonstrated using the DDL in a desert setting under two scenarios. The first was ground to ground with the OCU talking directly to the PackBot over DDL (no Raven relay). In this scenario, the UGV performed a set of simple tasks every 800m along a road that involved teleoperated commands requiring low latency and good visual feedback quality. The test involved precision turning, pan tilt unit control, and operator identification of a person. The test pattern was continued until the PackBot reached 3.2km (2mi), where communications were lost due to a dip in the road breaking line of sight.

The second scenario used the Raven UAV as a communications relay to the PackBot. In this configuration, the system was comprised of the Raven GCS (patch antenna) to a DDL Raven (Omni-directional antenna) to PackBot (Omni-directional antenna). The same tasks were executed as in the first scenario. For the first 4 km, the Raven was flown at orbiting waypoints lagging the PackBot's path down range. The Raven has weaker communications directly beneath it, so this was done to maximize signal strength. As the PackBot continued down-range from 4-8 km, the Raven maintained an orbit at 4km. Between 6km and 7km, the PackBot was driven around several one-story structures at the farthest point from line of sight to the Raven's orbit to demonstrate brief loss of communications and recovery that could occur. The PackBot then continued down range until reaching the end of the test area, at 8km. During the course of this demo, no increase in latency was perceived, and communications loss only occurred temporarily when loss of line of sight occurred around the one-story structures. Further testing at the sight demonstrated that the maximum distance achievable with the equipment is 14km.

Situational Awareness and BDA for Target Engagement

A team of vehicles running this system was also present at a simulated target engagement demonstration at the US Naval Air Weapons Station, China Lake. In this scenario, a group of targets among several buildings is engaged by a separate team, with this system providing stand-off reconnaissance and hypothetical battle-damage assessment (BDA). Just prior to the test, the primary larger UAV intended to provide overhead situational awareness suffered a crash landing, leaving this system as the sole source of situational awareness to the secondary team. It did so admirably, with the UGV providing close-up views complementing the wide-angle footage from the UAV and filling in for UAV observation outages during occlusions due to the tall buildings. Screenshots from this test are provided in Figure 15.

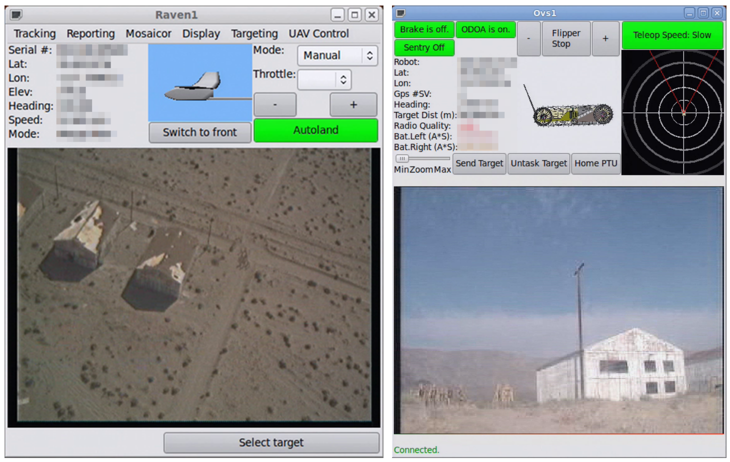


Figure 15. Screenshots of UAV and UGV control windows during a simulated target engagement scenario. The UAV offers wide-area situational awareness *(left)*, while the UGV provides detailed imagery from a safe stand-off distance *(right)*.

CONCLUSIONS AND FUTURE WORK

The collaborative framework presented provides unified control of multiple unmanned air and ground vehicles requiring minimal operator effort and maximizing situational awareness through high-level presentation of simultaneous vehicle and mission state. This is enabled by underlying algorithms implementing high-level capabilities permitting a variety of practical mission scenarios, a number of which were demonstrated in extensive field testing under varying conditions.

Despite the substantial capabilities of the system, this work has barely scratched the surface of the powerful potential of air-ground collaboration and centralized multi-vehicle control. Several future directions we anticipate pursuing include cooperative localization and navigation in which vehicles observe one another to provide improved state estimates and avoid obstacles visible only from other perspectives, communication-aware vehicle motion planning to keep vehicles positioned for minimal transmission losses, more complex reasoning for pursuit and capture to derive possible escape routes to block, extending view-predictive positioning to take into account known terrain occlusions, adding additional UGV capabilities such as more general outdoor autonomous navigation, and demonstrating the system on other types of air and ground vehicles.

ACKNOWLEDGEMENTS

This work is funded in part by contracts W15QKN-08-C-0029 and W15QKN-08-C-0497 coordinated by Jamal R. Musbeh, US Army RDECOM-ARDEC Network Lethality Division. The first author gratefully acknowledges support under a DoD National Defense Science and Engineering Graduate (NDSEG) Fellowship. The authors wish to extend their deepest gratitude to Earl Cox of AeroVironment, Inc. for his encouraging demeanor and constant assistance.

REFERENCES

[1]Hsieh, M. A., Chaimowicz, L., Cowley, A., Grocholsky, B., Keller, J., Kumar, V., Taylor, C. J., Endo, Y., Arkin, R., Jung, B., Wolf, D. F., Sukhatme, G. S., and MacKenzie, D., “Adaptive Teams of Autonomous Aerial and Ground Robots for Situational Awareness,” *Journal of Field Robotics*, Vol. 24, No. 11, 2007, pp. 991–1014.

[2]Vaughan, R. T., Sukhatme, G. S., Mesa-Martinez, F. J., and Montgomery, J. F., “Fly spy: lightweight localization and target tracking for cooperating air and ground robots,” *International Symposium on Distributed Autonomous Robot Systems*, 2000.

[3]Stentz, A., Kelly, A., Herman, H., Rander, P., Amidi, O., and Mandelbaum, R., “Integrated Air/Ground Vehicle System for Semi-Autonomous Off-Road Navigation,” *AUVSI Unmanned Systems Symposium*, July 2002.

[4]Jung, B. and Sukhatme, G. S., “A Generalized Region-based Approach for Multi-target Tracking in Outdoor Environments,” *IEEE International Conference on Robotics and Automation (ICRA)*, 2004.

[5]Grocholsky, B., Swaminathan, R., Keller, J., Kumar, V., and Pappas, G., “Information Driven Coordinated Air-Ground Proactive Sensing,” *IEEE International Conference on Robotics and Automation (ICRA)*, April 2005.

[6]Kim, H. J., Vidal, R., Rene, K., David, V., Shim, D. H., Shakernia, O., and Sastry, S., “A Hierarchical Approach to Probabilistic Pursuit-Evasion Games with Unmanned Ground and Aerial Vehicles,” *IEEE Conference on Decision and Control*, 2001.

[7]AeroVironment Inc., “Raven Product Data Sheet,” `http://www.avinc.com/downloads/Raven_Domestic_1210.pdf`, December 2010.

[8]DOD/NGIA Motion Imagery Standards Board, *Motion Imagery Standards Profile 6.1*, January 2011.

[9]Baker, S. and Matthews, I., “Lucas-Kanade 20 Years On: A Unifying Framework,” *International Journal of Computer Vision*, Vol. 56, No. 3, 2004, pp. 221–255.

[10]Comaniciu, D., Ramesh, V., and Meer, P., “Real-time tracking of non-rigid objects using mean shift,” *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 2, 2000, pp. 142–149.

[11]Collins, R. T. and Liu, Y., “On-Line Selection of Discriminative Tracking Features,” *IEEE Conf. on Computer Vision (ICCV)*, 2003, pp. 346–352.

[12]Tomasi, C. and Kanade, T., “Detection and Tracking of Point Features,” Tech. rep., International Journal of Computer Vision, 1991.

[13]Matthews, I., Ishikawa, T., and Baker, S., “The Template Update Problem,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, 2004, pp. 810–815.

[14]Conte, G., Hempel, M., Rudol, P., Lundström, D., Duranti, S., Wzorek, M., and Doherty, P., “High Accuracy Ground Target Geo-location Using Autonomous Micro Aerial Vehicle Platforms,” *AIAA Conference on Guidance, Navigation, and Control*, August 2008.

[15]Madison, R., DeBitetto, P., Rocco Olean, A., and Peebles, M., “Target Geolocation from a Small Unmanned Aircraft System,” *IEEE Aerospace Conference*, March 2008, pp. 1–19.

[16]Gibbins, D., Roberts, P., and Swierkowski, L., “A video geo-location and image enhancement tool for small unmanned air vehicles (UAVs),” *Intelligent Sensors, Sensor Networks and Information Processing Conference*, Dec. 2004, pp. 469–473.

[17]Barber, D. B., Redding, J. D., Mclain, T. W., Beard, R. W., and Taylor, C. N., “Vision-based Target Geo-location using a Fixed-wing Miniature Air Vehicle,” *J. Intell. Robotics Syst.*, Vol. 47, No. 4, 2006, pp. 361–382.

[18]Dobrokhodov, V. N., Kaminer, I. I., Jones, K. D., and Ghabcheloo, R., “Vision-Based Tracking and Motion Estimation for Moving targets using Small UAVs,” *American Control Conference*, June 2006.

[19]Ross, J. A., Geiger, B. R., Sinsley, G. L., Horn, J. F., Long, L. N., and Niessner, A. F., “Vision-Based Target Geolocation and Optimal Surveillance on an Unmanned Aerial Vehicle,” *AIAA Conference on Guidance, Navigation, and Control*, August 2008.

[20]Nuske, S., Dille, M., Grocholsky, B., and Singh, S., “Representing Substantial Heading Uncertainty for Accurate Target Geolocation by Small UAVs,” *AIAA Conference on Guidance, Navigation, and Control*, August 2010.

[21]Grocholsky, B., Dille, M., and Nuske, S., “Efficient Target Geolocation by Highly Uncertain Small Air Vehicles,” *To appear in IEEE International Conference on Intelligent Robots and Systems (IROS)*, September 2011.

[22]Dille, M., Grocholsky, B., and Nuske, S., “Persistent Visual Tracking and Accurate Geo-Location of Moving Ground Targets by Small Air Vehicles,” *AIAA Infotech@Aerospace Conference*, March 2011.

[23]Julier, S. J. and Uhlmann, J. K., “A New Extension of the Kalman Filter to Nonlinear Systems,” *International Symposium on Aerospace/Defense Sensing, Simulation, and Controls*, 1997.

[24]Ponda, S. S., *Trajectory Optimization for Target Localization Using Small Unmanned Aerial Vehicles*, Master's thesis, Massachusetts Institute of Technology, 2008.

[25]Tisdale, J., Ryan, A., Kim, Z., Törnqvist, D., and Hedrick, J. K., “A multiple UAV system for vision-based search and localization,” *American Control Conference*, 2008.

[26]Bay, H., Tuytelaars, T., and Gool, L. J. V., “SURF: Speeded Up Robust Features.” *European Conference on Computer Vision (ECCV)*, 2006, pp. 404–417.