GeoTrack: An Autonomous Closed-loop Target Tracking System for Small UAV Networks

Gaemus Collins · Prabir Barooah · João Hespanha

Received: date / Accepted: date

Abstract This paper describes a closed loop tracking and control system for a collaborating team of Small Unmanned Air Vehicles (SUAVs). The GeoTrack system includes automatic target detection and tracking with video, real-time video geo-registration, distributed sensor fusion, and bio-inspired collaborative UAV control. The nominal sensors employed by the system will be EO/IR video cameras on the SUAVs. Video tracking algorithms perform fully automated multiple-target detection and feature-aided tracking, geo-registration, and multi-sensor data fusion. Control algorithms are derived from swarming algorithms used by flocks of birds and schools of fish. These algorithms will enable SUAVs to move in formation and cooperatively track targets with minimal estimation errors, while relying solely on local communication between the agents. Implementation of the video tracking, compression, and geo-registration will be in a low-SW AP FPGA device that can run on-board a SUAV. This device and system is expected to improve the ISR capabilities and autonomy of SUAVs over the current state-of-the-art. This paper presents first-year preliminary progress and results.

Keywords Small Unmanned Air Vehicles (SUAVs) · Swarm control · Video tracking · Multi-sensor fusion and tracking · Geo-location · Embedded processing

1 Introduction

This paper describes an autonomous closed loop tracking and control system for a collaborating team of Small Unmanned Air Vehicles (SUAVs). Each SUAV sensor feeds streaming video to Toyon’s Image Plane Video Tracker (IPVT). The IPVT can automatically detect and track ground targets in the video imagery. Detections or target tracks generated by the IPVT are fed to Toyon’s Global Fusion and Tracking Center (GFTC), which uses DTED terrain data and camera state telemetry to geo-register target positions in an Earth-based coordinate frame (lat/lon/alt). The GFTC also performs multi-sensor data fusion, combining detections from multiple SUAVs to reduce target tracking errors. Toyon’s Cooperative Decentralized Asset Manager (CDAM) employs multi-UAV coordinated routing and sensor tasking algorithms to automatically search for, intercept, and track ground targets, even through urban environments with partial obscuration. The IPVT, GFTC, and CDAM system components interact with Toyon’s SLAMEM® simulation environment.

Thanks to the Institute for Collaborative Biotechnologies (ICB) and AFOSR for funding this work. ICB contract number W911NF-09-D-0001, AFOSR grant numbers FA9550-05-C-0180, FA9550-06-C-0119.

Gaemus Collins
Toyon Research Corporation
6800 Cortona Dr.
Goleta, CA, 93117
805-968-6787 x113
E-mail: gcollins@toyon.com

João Hespanha
Electrical and Computer Engineering Dept.
University of California, Santa Barbara (UCSB)
Santa Barbara, CA 93106-5100
E-mail: hespanha@ece.ucsb.edu

Prabir Barooah
Mechanical and Aerospace Engineering
University of Florida (UF)
Gainesville, FL 32611
E-mail: pbarooah@ufl.edu
SLAMEM can display UAV and sensor locations, target locations, and terrain in real time, and can simulate additional UAVs, sensors, manned aircraft, buildings, foliage, etc. in the same environment, so that the real entities can seamlessly interact with the simulated entities.

This paper provides background to the GeoTrack program, and highlights specific progress during the first-year of a three-year program:

- We have modeled video measurements for the purpose of optimizing SUAV swarm formations relative to the target position. Given the image pixel coordinates of the target, the intrinsic and extrinsic camera parameters, and terrain data, one can estimate the 3-dimensional location of the target in Earth coordinates and compute the associated error covariance.
- We have improved the frame registration stage of our automatic video tracker to better handle fast motion and jitter in SUAV video.
- We have improved geo-registration of SUAV video by digitizing and GPS time-stamping the video on-board the SUAV. The time-stamped digital video frames can tightly synchronized with sensor state telemetry, requiring minimal inter-frame state interpolation.
- We have developed a prototype FPGA-based video tracker for eventual installation on-board the SUAV. This embedded tracker operates within low size, weight, and power consumption (SWAP) requirements. This technology will enable moving the GeoTrack signal processing on-board a SUAV, improving the tracking capabilities and autonomy level of these low-cost assets.
- We have developed a track fusion algorithm that produces smoothed target state estimates superior to those produced by the individual trackers, using information received from nearby SUAVs at rates much lower than the individual measurements.
- We have tested portions of this system in simulation and in hardware, and verified some theoretical results.

2 Unsupervised Real-time Image Plane Video Tracking

The GeoTrack system employs advanced algorithms for video-based target tracking. Toyon has developed real-time video tracking software that performs fully automated multiple-target detection and feature-aided tracking. Included is a novel motion-based target detection algorithm that has proven to be effective at detecting humans and vehicles in video containing significant parallax effects and motion clutter due to moving vegetation, water, etc. The algorithm operates by estimating camera motion compensation parameters which are used to register a statistical background model which is updated with each new video frame. Pixel-level change detection is performed by evaluating each frame in the background model, and a target detection filter operates on the spatial properties of the change pixels to segment the moving targets. From the segmented pixels, target appearance features are extracted, which can be used to help resolve measurement-to-track data association ambiguities, and for performing feature-based detection.

Fig. 2 Toyon’s IPVT unambiguously detects and tracks multiple ground targets in SUAV video

3 Global Tracking and Automatic Target Geo-location

The GeoTrack system performs unsupervised real-time geo-location of targets in the image plane. Figure 3 displays the image plane target geo-location problem. Toyon’s IPVT generates video image plane detections (as described in Section 2) and transmits the detection information to the GFTC. Simultaneously, SUAV state data is measured by the aircraft navigation sensors and sent to the GFTC from the SUAV autopilot. The our GFTC software fuses the video tracking data with SUAV state data to geo-locate image plane detections and track targets in a global coordinate system.

Fig. 3 Image plane target geo-location

Solving the video geo-location problem in real-time requires extremely tight synchronization of the SUAV video frames with the SUAV state data. We currently accomplish this synchronization by first digitizing the video on-board
Fig. 4 Toyon’s GFTC application geo-registering detections from SUA V video. Red dots represent target detections (and some false detections). The white dot is the sensor location, and the polyhedron shows the sensor field of view on the ground.

the SUAV. Analog video is converted to MJPEG, and each JPEG frame in the MJPEG stream is time-stamped with GPS time. Meanwhile, SUA V position and orientation of the aircraft, and orientation of the sensor, are measured by the GPS, IMU, and pressure sensors on the aircraft. This aircraft and camera state data is GPS time-stamped by the autopilot, and transmitted to the ground control station (GCS) on a separate data channel. The two data streams can then be synchronized by the GFTC, with only small (<1 second) inter-frame state interpolation required.

As mentioned, the result of processing video is one or more detections and an associated uncertainty (i.e., covariance ellipse) about that detected location. These detections are in the image plane itself and are two-dimensional measurements (in units of pixels) with an associated covariance. For state estimation of a ground vehicle’s state using image plane detections, we must define a measurement equation that has the general form:

\[
\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k, \beta_k) + \mathbf{n}_k
\]

in which \( \mathbf{x}_k \) represents the vehicle’s state (position and velocity) in the TCS frame, \( \mathbf{z}_k \) are the pixel coordinates of a detection in the image plane, \( \mathbf{n}_k \) is an additive noise value (in pixels) to account for the imperfect location of the detection with respect to the true location of the vehicle in the image plane, and \( \beta_k \) is a vector of parameters (i.e., rotation angles, focal length, translation parameters) that characterize the particular mapping of 3D points to image plane locations. Some of these parameters may be known with sufficient accuracy and treated as deterministic, while some are modeled as random with given estimates and uncertainties about those estimates (often referred to as nuisance parameters). We can use this measurement equation for the state estimation algorithm but would have to define the appropriate covariances (e.g., the innovation covariance), taking into account the errors on these nuisance parameters.

In order to use these image plane detections for tracking an object in a three-dimensional frame, we must define a model for how the three-dimensional state of the object is mapped to a two-dimensional measurement in the image plane. Such a model has multiple nonlinear transformations which are shown in Figure 5.

An alternative to the nonlinear measurement model is to convert the image plane measurements to a different measurement that has a different measurement equation. For example, suppose we take the image plane measurements and clamp them to the TCS ground plane to create a “measurement” of the target location in the tracking coordinate system. The measurement equation in this case is simply

\[
\mathbf{z}_k = \begin{bmatrix} 1 & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \end{bmatrix} \mathbf{x}_k + \mathbf{n}_k
\]
While the process of creating a converted measurement appears straightforward, we also must determine the uncertainty in the converted measurements. The original measurements that result from the sensor data processing have associated uncertainties (i.e., a measurement error covariance). Therefore, when we apply a transformation to these measurements to create a new (converted) measurement, we must also modify the error characteristics of the original measurements in order to derive a measurement error covariance for the converted measurements. Further complicating the problem are the presence of uncertainties in the meta-data (camera location and orientation).

There are a few approaches for propagating the uncertainties through transformations. In the case of a linear transformation, this process is straightforward and exact. For nonlinear transformations of a random vector (e.g., the process of clamping detections), we must approximate the statistics of the transformed vector. One approach relies on determining the Jacobian of the nonlinear transformation and evaluating the Jacobian at the mean of the random vector subjected to the transformation. An alternative is to use a numerical approach referred to as the unscented transform (UT) [Julier et al(2000)]Julier, Uhligmann, and Durrant-Whyte]. In the UT, we generate a set of deterministically chosen sample points from the probability density function (PDF) of the original random vector. These samples are then subjected to the nonlinear transformation to generate a new set of sample points from which a sample mean and covariance are generated. This sample mean and covariance then represents the mean and covariance of the transformed random vector.

In our problem, the random vector that we want to transform is the 2D vector representing the location of the target in the image plane. We have the mean and covariance of this vector from the image processing step. Ostensibly, then we would generate samples from this 2D PDF and subject them to the nonlinear transformation (i.e., the ground-clamping algorithm). However, this clamping process involves other random quantities such as the sensor position and orientation, which cannot be ignored if an accurate representation of the uncertainty in the ground-clamped estimate is desired. We have derived how we would use the UT to account for the uncertainties in these nuisance parameters and have used this approach to generate a mean and covariance for the ground-clamped location of the target. The ground-clamped estimate is \( z_k \) in (2), while the uncertainty in this estimate is the covariance matrix associated with \( n_k \) in (2).

A single GFTC can receive detections from multiple IPVTs, and fuse the different data streams to provide higher target estimation accuracy. These IPVTs can be connected to moving and/or stationary cameras, allowing for future expansion of GeoTrack to a mixed air- and ground-based surveillance system. The GFTC can also fuse data from other types of sensors, such as GMTI radar. We have successfully demonstrated a remote queuing surveillance scenario in which a simulated Global Hawk sends GMTI radar detections from SLAMEM to a GFTC. The GFTC instantiates track(s) for the detections sequence(s) and feeds them to our CDAM software. CDAM automatically plans routes for the SUAVs and sends the waypoints and sensor tasks to the SUAV autopilot for execution. The SUAVs then fly out to investigate and image the targets that were detected by the Global Hawk.

### 4 Distributed Sensor Fusion over Networks

In [Barooah and Hespanha(2007)], [Barooah et al(2006)Barooah, da Silva, and Hespanha] we developed distributed algorithms for data fusion in networks of relative measurements. These estimation algorithms were inspired by swarming behavior observed in flocks of birds and schools of fish. One of the most well-known biologically inspired models for swarming was introduced by Reynolds in 1987 [Reynolds(1987)] for the creation of computer animations of flocking behavior. Reynolds’ key observation was that cohesion motion for a swarm was possible based solely on local interactions between an agent and its neighbors. For birds flying in a V-formation, each animal senses its neighbors through vision and turbulence in the airflow. Reynolds’ heuristic model for swarming was based on three rules to govern flocking behavior: cohesion, which causes an individual to stay close to nearby flock-mates; separation, which prevents collision with nearby flock-mates; and alignment, which is responsible for matching velocity with nearby flock-mates. Although a complete analysis of these types of algorithms is still unavailable, some of its simplest forms are sufficiently well understood to be applied to distributed estimation and control problems [Olfati-Saber(2006)], [Barooah and Hespanha(2007)], [Barooah and Hespanha(2006a)].
A network of relative measurements, like those in the biological examples above, arises in the multi-sensor fusion and tracking problem of Figure 6(b). Although position estimates for each target can be obtained by a single SUAV (as described in Section 3), by fusing measurements across a SUAV network, a more accurate target position estimate can be computed. Track ambiguities, possibly resulting from vehicles that travel very close together, can often be resolved by using information from several SUAVs.

Our research has also shown that Reynolds’ rules can be implemented in an asynchronous manner, meaning that multi-sensor fusion can improve tracking performance even if the exchange of information between agents are not synchronized by a common clock. This implies that SUAVs can utilize these rules in GPS-denied regions or over unreliable and latent communication networks.

An advantage of using multiple SUAVs in target tracking with noisy information can be understood by Figure 6(b). To achieve accurate target geolocation, target positions must be estimated by fusing the positions of the SUAVs and relative position measurements between SUAVs and the targets. In our first-year prototype GeoTrack system, distributed sensor fusion resides at the output of the GFTC for each SUAV, as depicted in Figure 7. Ground track data is provided by the GFTC at approximately 10Hz. These tracks are preprocessed inside a Track Fusion Module (TFM), and the processed information will be exchanged between multiple TFMs at a rate of approximately 1Hz. The local and received data streams are then fused in the TFM, which produces smoothed target state estimates superior to those produced by the individual GFTCs. The output of the fusion block is a smoothed estimate, meaning that current measurements are used to improve past estimates.

The pre-processing block of the TFM collects measurements from the GFTC. Each measurement is time-stamped, and arranged into a “comb graph,” as in Figure 8. In this graph, each vertical edge represents a relative measurement between the UAS and the target, and each horizontal edge represents a dynamical model, in this case of the motion of the target. Moving horizontally to the right corresponds to the passage of time.

![Figure 6 Distributed Estimation in a biological network and an SUAV network](image)

![Figure 7 The communication topology used to fuse the information from two UAS simultaneously imaging a common target.](image)

![Figure 8 The comb-like graph of the relative measurements between one UAS (red triangles) and the target (black squares). The horizontal axis corresponds to the passage of time. In total, this image depicts relative measurements at 11 distinct time instants. Note that no communication has taken place yet.](image)
The utility of the comb-graph is that an optimization can be performed to determine optimal estimates of the target positions, factoring in the available relative measurements and the dynamical model. For the particular case of linear measurements and dynamics with Gaussian noise, as assumed here, the optimal estimates correspond to the output of a Kalman smoother. Kalman smoothing is the best one could hope to obtain from a single UAS.

The situation becomes more interesting when more than one SUAV is imaging the target. When measurements are made more quickly than they can be shared via communication, one must select which information to share. Here, we elect to share measurements directly, along with corresponding time stamps. The measurements are easily included in the comb graph, which now takes on a new structure we call an “extended comb graph structure,” as in Figure 9. With the extended comb graph, we can again perform optimal estimation to obtain the Kalman smoothing result. If more bandwidth is available between the UAS, additional measurements can be transmitted resulting in even better joint estimation.

![Extended comb graph](image)

**Fig. 9** The extended comb graph for multi-UAV fusion from two SUAVs is depicted. The top SUAV and target are as previously shown in Figure 8, and three received measurements from the adjacent SUAVs are shown on the bottom of the graph.

Two additional properties of the comb graphs are noteworthy. The first property is that the Kalman smoothing result can be obtained in a distributed (but iterative) manner using Jacobi iterations. While not immediately necessary for a few UAS tracking a single target, this iterative solution could prove useful in larger UAS formations. The second property relates to the structure of the comb graphs. Here, they appear as trees, which is a particularly simple case. In general, the graph of relative measurements can be more complicated, and may even contain cycles. In the cycle case, we have developed a distributed estimation algorithm that can reach the optimal solution much quicker than traditional methods [Russell et al(2010)]

5 Embedded Video Tracking and Geo-registration

There is a significant performance advantage associated with operating the entire IPVT, video-state synchronization, and geo-registration processing chain on-board the SUAV platform. While primarily useful for increasing the accuracy of tracking results, particularly in terms of tracks relative to platform position, there are other additional benefits. These include the ability to process raw, uncompressed video at full frame rate. In terms of data transmission, it is possible to reduce frame rate and video quality, thereby relaxing radio requirements resulting in either reduced power and/or increased range.

To this end, Toyon is developing a low-SWaP (size, weight, and power) embedded version of our IPVT and GFTC that will run on-board a SUAV. This embedded GeoTrack device will receive a direct camera feed and perform image plane tracking on the uncompressed video. It will also interface directly with the SUAV autopilot via RS-232 serial connection to obtain camera state data for target geo-registration. The video will then be compressed as H.264 to reduce bandwidth consumption. The data transmitted by the GeoTrack device to the SUAV operator will be comprised of compressed video and geo-registered target track information.

While traditional embedded video tracking solutions have focused on the use of digital signal processors, Toyon has pursued a heterogeneous approach. The motivation is to apply different processor technologies to different architectural elements of the video tracking solution. Generally, the video tracking problem is broken down into three parts: frame registration, detection and segmentation, as well as tracking. A field programmable gate array (FPGA) handles the bulk of front-end camera operations and parallel, integer-based mathematical operations. An FPGA-based 32-bit RISC soft-processor with floating-point hardware acceleration provides for computation of high dynamic range numbers. Finally, an ARM processor with DSP acceleration handles the bulk of non-synchronous operations associated with data association and tracking. This combination maximizes performance while minimizing power, a major consideration for small-platform UAS.
Our current GeoTrack prototype implements frame registration using the KLT algorithm, partitioned appropriately between available hardware logic and microprocessor operations. A probabilistic data association (PDA) technique is used to compute the likelihood that a detection belongs to a specific track. When multiple tracks cross or overlap, ambiguities can arise. To deal with this situation, each feasible target to track association is treated as a separate hypothesis. The probability of each hypothesis is computed and carried forward to the next video frame.

The target tracking software uses an approach derived from Kalman filtering, capable of tracking multiple targets, even in an environment with significant process noise and measurement noise. In each video frame, the position and velocity of the target is used to predict its position in a subsequent frame. Noise is minimized by statistically weighting the predicted vs. measured position.

6 Distributed Swarm Control and Multi-UAV Cooperation

The GeoTrack system employs algorithms for cooperative decision and control of SUAVs and their sensors. These algorithms will facilitate cooperative UAV missions supporting Intelligence, Surveillance and Reconnaissance (ISR). For small collections of agents, our CDAM architecture can retain cooperative control performance similar to that achieved by a centralized controller, using only intermittent or asynchronous communications.

It is believed that coordinated behavior can be observed in over 10,000 species, and may be the result of evolutionary survival mechanisms. Migratory birds provide some of the best examples of motion coordination. Especially in long-range migrations, these birds fly in a V-shape formation with a bird in the lead followed by the remaining birds trailing behind in two straight lines (see Figure 10). It has been theoretically shown that the increased aerodynamic efficiency of formation flight can allow 25 birds in a flock to increase their range by about 70% as compared to each bird flying alone.

This control problem of maintaining an airborne formation using relative position measurements is solved very successfully by birds in flocks. In [Barooah and Hespanha(2006b)], [Barooah and Hespanha(2006a)] we analyzed algorithms believed to be used by schools of fish and flocks of birds to maintain tight groups and specific formations, an instance of swarm control. In [Barooah and Hespanha(2006a)] it was shown that tight formations can be maintained using only communication between direct neighbors, even in the presence of measurement noise and communication losses.

Coordinated target tracking using multiple SUAV surveillance vehicles is a challenging problem that has received significant attention in the past decade. The accuracy of target geo-location achieved by fusing measurements from all the SUAVs will depend on the positions of the SUAVs relative to the targets and the measurement error models. Given the image pixel coordinates of the target, the intrinsic and extrinsic camera parameters, and terrain data, one can estimate the 3-dimensional location of the target in Earth coordinates and compute the associated error covariance. The extrinsic camera parameters of UAV position and sensor attitude are the primary sources of geolocation error, and are thus treated as the exclusive sources of noise corrupting the target geolocation estimate in our model.

The estimation errors that arise in vision-based tracking are generally characterized by error covariance matrices corresponding to ellipses, with the major axis aligned with the vector from the camera position to the target. These types of error footprints shown in Figure 11 are especially common when cameras view the ground from low grazing angles: the projection of the image frame onto the ground produces errors that are exaggerated in one direction. Consequently, a majority of coordinated target tracking algorithms are based on view angle coordination. For example, with two UAS tracking a target, the natural coordination is to steer the pursuit vehicles so that they observe the target from orthogonal directions at a fixed standoff distance, as shown in
6.1 Problem Formulation

The problem formulation begins with $N_u$ UAS that are tasked with tracking a single target. Each UAS has nonholonomic dynamics modeled as a constant-speed planar kinematic unicycle,

$$
\begin{align*}
\dot{x}_i(t) &= v_a \cos(\psi_i(t)) \\
\dot{y}_i(t) &= v_a \sin(\psi_i(t)) \\
\dot{\psi}_i(t) &= u_i(t).
\end{align*}
$$

(3)

Here $x_i, y_i$, and are the position and heading of the $i$th SUAV, and $v_a$ is the SUAV constant forward speed. The single control input, $u_i$, affects the curvature of the vehicle, and is constrained to live within some set, $u_i \in [U_{\text{min}}, U_{\text{max}}]$.

We assume each UAS has a camera-like sensor co-located with the vehicle. We show in [Quintero et al(2010)](Quintero, Papi, Klein, Chisci, and Hespanha) that the dominant source of error in the geolocation estimate for a single UAS arises from the uncertainty in the sensor attitude angles. These uncertainties are mapped to the ground plane, as described in Section 3, and used to formulate an optimization over control inputs $u_i$ that minimizes the total tracking error.

We have employed dynamic programming to optimize the control input sequence. We define a cost function $J$ for dynamic programming to be the mean trace of the fused error covariance over a horizon of $N_s$ stages. The optimal control input sequence $u^* = \{u(k)\}_{k=0}^{N_s-2}$ is found by solving

$$
u^* = \arg \min_{\nu^{(k)}} J,$$

(4)

from a specific initial condition.

6.2 Results

To demonstrate the effectiveness of the dynamic programming approach to UAS coordination, we have performed initial analysis on this method using an example with two UAS tracking a single target, assuming a fixed travel speed of $v_a = 15$ m/s and quantized in time steps of one second.

In the first scenario, the target travels at 5 m/s, and we optimize over a horizon of 90 seconds. Sample trajectories for the UAS tracking the target are shown in Figure 12, and corresponding data are given in Figure 13. This data includes the planar distances-to-target, denoted $d_1(t)$ and $d_2(t)$, the coordination angle $\phi(t)$ between the two UAS, and lastly, the instantaneous trace of the fused error covariance.
In Figure 12, we observe the UAS making alternating passes over the target. Also, from Figure 13, the issue of keeping a 90° angle between the UAS is clearly not important. In fact, the trace of the fused error covariance has its minimum values when at least one UAS is on top of the target, i.e., $d_1(t) = 0$ or $d_2(t) = 0$. As an illustration, by looking at the times of 20 and 40 seconds in both figures, one can see that the value of the trace is comparatively lower than other time instances, as UAS 2 is close to being over the target at these times.

Fig. 13 Distances, relative angle, and cost performance for two UAS tracking a single target with a speed of 5 m/s.

7 High-Fidelity Simulation and Mixed HWIL-with-Simulation

7.1 SLAMEM

The SLAMEM simulation is a powerful computational tool that facilitates the exploration of many issues regarding the architectures and concept of operations (ConOps) of surveillance and targeting systems, the utility and performance requirements of individual sensor types, the effects of terrain and foliage on surveillance effectiveness, and the role of threat tactics in degrading the quality of surveillance and attack operations.

The SLAMEM environment module includes models of terrain, road network, foliage, and the atmosphere. Standard NIMA products are used. SLAMEM features models of a wide variety of surveillance assets, including UGS, UGVs, and UAVs. UAVs use a 6-DOF flight model, have finite fuel resources, and their flight paths must obey maximum g-limits. Wind can affect the velocity of small, light UAVs. Sensors are characterized by parameters such as range and azimuth scan limits, scan rate, minimum and maximum detectable velocities, wavelength, azimuth beamwidth, boresight angle and foliage penetration depth, if any. Probabilities of detection can be functions of sensing geometry, mode of operation, resolution, and target type. Obscuration due to terrain, foliage, buildings and – in the case of EO/IR, ladar, and video sensors – atmospheric effects (e.g., smoke and clouds) can prevent detection of ground targets. Asset positions and sensor pointing directions may be known imperfectly, with random and bias errors.

The ground vehicle model in SLAMEM is the Ground Vehicle Simulator (GVS™). GVS uses acceleration states to model starting and stopping motions as well as normal driving speed fluctuations. Vehicles also reduce speed when navigating along road bends. Fast vehicles are not allowed to pass slower vehicles and must reduce their speed to avoid a collision.

SLAMEM has been used as the center for our system development on this project. Our control and data fusion algorithms work with SLAMEM sensor and communications models to emulate a complete UAV autonomous control
system. Video sensor exploitation models produce measurements to determine the presence of objects in a region, and communications models provide realistic data link constraints on the network to determine the feasibility of our decentralized processing and control architecture. SLAMEM asset models represent Toyon’s UAV platforms and other standard military platforms in the simulation and execute commands from our control algorithms. GVS supplies realistic targets of interest and confuser vehicles to the simulation. Figure 14 shows SLAMEM running a simple simulation test case.

Fig. 14 A SLAMEM screenshot showing 2 SUAVs and 2 targets. One SUAV is tracking one target, while the other SUAV is in a search mode. SUAV paths are automatically generated by the GeoTrack routing algorithms, and displayed in black. Terrain and a road network is also shown.

7.2 Toyon’s UAV Testbed

The UAV platform employed by our team is based on the Unicorn expanded polypropylene (EPP) foam wing [Unicorn Flying Wing(2009)], shown in Figure 15, outfitted with a gimbled video camera and digital video encoder. This durable platform is well-suited for testing prototype tracking and navigation control algorithms because of its ease of use. We have flown over 100 test flights with this platform, and successfully tested portions of the GeoTrack system including automatic routing, sensor tasking, and video tracking.

Fig. 15 Toyon’s Unicorn UAV test platforms

7.3 Hardware-in-the-Loop with SLAMEM

SLAMEM can act as a hardware-in-the-loop (HWIL) display and controller for a complete hardware/software system, as described in [Collins et al(2007)Collins, Vegdahl, and Riehl]. We can integrate real UAVs, sensors, targets, wind data,
etc. together with simulated aircraft, vehicles, sensors, etc. into a single demonstration. In particular, this capability allows our system to simulate buildings, weather, foliage, etc. in the same environment where real UAVs are flying. This functionality has enabled HWIL demonstrations of our algorithms for the Air Force and SOCOM (at TNT events [Tactical Network Topology(2009)]), and will provide a testbed for the development and demonstration of GeoTrack.

7.4 Results

Initial testing of the GeoTrack system and SLAMEM framework was performed, while SUAV position and orientation, the target position, and the sensor data were recorded and used to compute baseline parameters for the performance of the GeoTrack system. Observed parameters of interest include

- UAS-to-target distance,
- UAS-to-target orientation,
- distance from target position to camera boresight ray intersected with Earth,
- position of target in camera Field of View (FOV),
- % time sensor is viewing target,
- target track lifetime, and
- target track accuracy.

We set up SLAMEM simulation runs to test the new metrics. The first set of plots shown below were generated by running GeoTrack algorithms in a SLAMEM simulation scenario. 2 small UAS are being controlled with our SIF algorithm [Collins(2010)] to search an Area of Interest (AOI) and locate and follow targets. There are four targets moving within the AOI: two targets of interest (TOIs) and two confuser vehicles. The UAS are launched and fly for 1 hour. (The simulation entity numbering scheme is non-sequential, so the UAS are labeled “UAS 5” and “UAS 6”; the targets are labeled “Target 3,” “Target 4,” “Target 8,” and “Target 9.”)
Figure 16 Shows the 2D distance between UAS 6 and each target, over time. From this plot we see that UAS 6 located and followed target 3 from about $300 < t < 1200$, then switched to follow target 4 from $1200 < t < 3600$.

![Graph showing 2D distance between UAS 6 and each target over time. From the plot, it is seen that UAS 6 located and followed target 3 from about $300 < t < 1200$, then switched to follow target 4 from $1200 < t < 3600$.](image-url)
It is typically impossible for the UAS to maintain a constant distance from the target for several reasons:

- the exact target location and speed are unknown to the UAS routing algorithm – it only knows the estimated target position and speed (i.e. the track);
- the target is often moving slower than the minimum UAS speed, so the UAS must loop or “S-turn” to stay nearby the target.

This problem can be partially mitigated with a gimbaled UAS sensor, by constantly re-pointing the sensor directly at the target, independent of the UAS position. Figure 17 shows the distance between (a) the target, and (b) the sensor boresight ray intersected with the Earth. Ideally, this distance should be zero during UAS Follow mode. However, tracking errors combined with limited gimbal range of motion will generate nonzero errors in this statistic, as evidenced by Figure 17.

**Fig. 17** 2D distance between (a) the target, and (b) the sensor boresight ray intersected with the Earth, plotted over time
Although Figure 17 shows nonzero errors between the target position and sensor boresight position, Figure 18 shows that, most of the time, the target stayed within the sensor FOV. This is a critical metric for performance of the GeoTrack system.

**Fig. 18** Target(s) within the sensor FOV
The SLAMEM infrastructure and mixed HWIL capabilities (Section 7.3) enable performance metrics to be computed from simulated and real data. Preliminary flight testing of the GeoTrack system was performed in July and September 2009, with a single SUAV tracking a single target. Hardware results are shown in the following figures.

Similar artifacts can be observed in both the simulation and hardware results, which helps to verify the accuracy fidelity of the SLAMEM environment. Also, comparing Figures 19 and 22 indicates that the cyclical SUAV-to-target distance resulting from looping induces a corresponding cyclical track position error. This result stresses the importance of the distance coordination work in Section 6.

**Fig. 19** A time snapshot plotted from recorded flight and target data from Camp Roberts UAS testing, July 2009. This data shows a single pass down General’s Road in Camp Roberts with one UAS following one target. The UAS position is plotted in blue, the UAS sensor boresight is in red, and the true target trajectory is shown in green.

**Fig. 20** 2-D distance between the UAS and target, plotted over time
8 Conclusions and Future Work

We have presented the components of a closed loop tracking and control system for a collaborating team of SUAVs. The Geotrack system includes automatic target detection and tracking with video, real-time video geo-registration, distributed sensor fusion, and bio-inspired collaborative control. Most of the Geotrack components currently run on ground-based computers, but we are in the process of moving much of this computation into a low-SWAP device that can run on-board a SUAV. This device is expected to improve the ISR capabilities and autonomy of SUAVs over the current state-of-the-art.

Our future work is focused on improving, implementing and testing distributed formation control algorithms and distributed sensor fusion algorithms in CDAM and GFTC, respectively. We have used intrinsic and extrinsic camera

Fig. 21 2-D distance between the target and the sensor boresight ray (intersected with the Earth)

Fig. 22 Track 1 position error (relative to vehicle truth). Comparing to Figure 19 indicates that the cyclical SUAV-to-target distance resulting from looping induces a corresponding cyclical track position error.
parameters to build video detection error models, from which optimal formations for the SUAVs can be computed. These algorithms will be implemented in CDAM and flight tested with our SUAV testbed. Toyon’s GFTC software will be augmented with distributed estimation algorithms for data fusion in networks of relative measurements. Using intermittent local communication, we anticipate that these algorithms will significantly reduce the anisotropic estimation errors common in vision-based sensors.

Acknowledgements
The authors would like to thank the Army Institute for Collaborative Biotechnologies for their generous support of this research.

References


