AIAA Guidance, Navigation, and Control Conference
2 - 5 August 2010, Toronto, Ontario Canada

AIAA 2010-7886

Representing Substantial Heading Uncertainty for Accurate Geolocation by Small UAVs

Stephen Nuske, Michael Dille, Ben Grocholsky and Sanjiv Singh
Robotics Institute, Carnegie Mellon University *

Geolocation of a ground object of interest from live video is a common task required of small and micro unmanned aerial vehicles in surveillance and rescue applications. The small low-cost sensors these vehicles carry provide low accuracy when mapping an image coordinate to a world location. Frequently, a primary source of such inaccuracy is error in vehicle heading. Filtering methods that inadequately represent the resulting nonlinear uncertainty distributions of geolocation measurements will produce inconsistent and inaccurate estimates. This paper presents a geolocation filter with a discretized solution space that correctly handles sampled nonlinear observations. The filter achieves higher accuracy when compared to alternative linearized methods. Assessment of the improved solution accuracy for stationary objects is provided through flight experiments using a commercial human-portable fixed-wing UAV system.

I. Introduction

Autonomous small unmanned aerial vehicles and micro aerial vehicles (SUAVs and MAVs) are at the forefront of interest in many surveillance and rescue applications for their attractive combination of light weight, low cost, low operator burden, and up-close low altitude operation. However, lightweight and low-cost vehicles typically restrict the available sensing hardware to low-end, consumer-grade devices that offer limited precision and therefore higher uncertainty when performing tasks such as deriving estimates of the location of a ground-based object of interest in a global coordinate frame. Typical sensor suites for such UAVs will consist of a GPS unit for position; accelerometers, gyroscopes and magnetometers for attitude; and a barometer for altitude. Combining these vehicle state sensors with a camera and visual tracking algorithm allows for the estimation of the location of an object by projecting a ray from an image coordinate and intersecting it with a terrain map or rays from previous observations. However, each of the considerable inaccuracies introduced by each sensor in the suite must be taken into account in the final estimate, lest it be uselessly inaccurate.

This paper demonstrates improvements in geolocation accuracy by more correctly propagating uncertainty in the vehicle's state-space into an estimation filter in the solution-space (the object's world location). Typically the heading estimate of a low flying UAV is poor, and we develop a representation that explicitly models large heading uncertainty. We present a scheme that passes a discrete sampling of the non-linear state uncertainty to a grid-based filter which maintains a geolocation estimate. The accuracy improvements of our approach are proven over extensive trials of a small commercial UAV by comparison with implementations of conventional approaches. Field tests are conducted over several months and several flights in a range of different conditions. The results from these field tests illustrate heightened accuracy in the geolocation estimate over more common naïve filters, and in some cases the accuracy gains are dramatic. All improvements are achieved without any attempt to improve the state estimate of the vehicle provided by the UAV's sensors, which is often impractical.

II. Related Work

Two broad approaches may be found in the literature for UAV-based geolocation of a ground object whose image coordinates are already provided by a separate visual tracking algorithm or operator input.

*Pittsburgh, Pennsylvania, USA, 15213 (Email:nuske@cmu.edu)

One is to register previously geo-referenced satellite imagery with the live UAV video streams, demonstrated for instance by Conte et al.[1], which conveniently provides a direct measurement of the object's location in world coordinates. However, such techniques are vulnerable to environmental changes since the satellite imagery was captured, they can also present a significant computational burden, and do not provide clear bounds on error in geolocation due to mis-registration. More common, therefore, is the use of an estimate of the vehicle's pose derived from on-board sensors to project the tracked image coordinate as a ray in world coordinates that may be intersected with previous observations or a prior terrain map.

As the pose estimates provided by low-cost on-board sensors can be quite poor, the ray generated for a given image coordinate may be highly inaccurate. One approach to improving its accuracy is to directly improve the vehicle pose estimate upon which it is based, and there exists a sizeable body of work on attempting to reduce this error by augmenting these sensors with a visual localization algorithm. Full SLAM systems using Extended Kalman Filters (EKFs) were exhibited by Bryson and Sukkarieh[2] and Caballero et al.[3], in which the prior formulate their filter with bearing-only observations of tracked landmarks and the latter with 3D observations formed by projecting tracked landmarks and intersecting with a known ground plane. Madison et al.[4] demonstrate tracking multiple ground features over several tens of seconds to constrain the standard Structure-from-Motion 8-point algorithm, which solves for all the feature locations and vehicle poses. As an alternative to explicit landmark tracking, Andersen and Taylor[5] show that with a planar ground assumption, a homography-based visual odometry algorithm can be combined with GPS/INS sensors using an Unscented Kalman Filter (UKF) to improve the vehicle pose estimate. These visual localization approaches typically require stable visual tracking of the environment over hundreds of frames, which is difficult even in the easiest of scenarios given that SUAV and MAV video streams are often low quality, highly jerky, and provide a view frequently lacking in track-able features (such as when flying over homogeneous terrain including fields or forests). In addition to its use for geolocation, matching of live video to existing satellite imagery as performed by Conte and Doherty[6] and Sim et al.[7] can greatly improve UAV state estimates, but again at significant computational cost and vulnerability to changes in terrain appearance.

An alternative philosophy is to embrace the inevitable uncertainty in vehicle pose once an appropriate level of effort has been invested in improving it as much as is realistic. At this point, a series of uncertain object observations is being received, and the goal now becomes filtering these to yield the best estimate of the geolocation. This approach represents a minimization of error directly in the two or three dimensions of the object's location rather than solving for a 6 degree-of-freedom vehicle pose state using 2D image observations.

Madison et al.[4] show that multiple rays can be combined to find their best intersection. Given a prior model of the terrain, a single observation can be directly projected and intersected with the terrain model, as discussed by Gibbins et al.[8], to provide a 3D geolocation estimate. A filtered version of such a terrain intersection method is presented by Dobrokhodov et al.[9], which is based on a filter by Wang et al.[10]. Their filter can account for moving objects, however, when the object is known to be stationary, a more accurate version of this intersection method is to batch-combine observations from an orbit of an object as Barber et al.[11] show. They demonstrate a recursive least squares algorithm errors of 30 m are reduced down to below 5 m. However, these results appear to be applicable only when observation errors result in symmetrical biases uniformly distributed around the object's location that cancel-out over a symmetrical orbit. We present experimental data wherein a symmetrical orbit of an object produces raw observation errors that are not symmetrical and for which such a least squares solution does not produce a suitable answer.

While some of the aforementioned approaches give cursory mention of the use of a Kalman filter for object state estimation, they do not discuss how observation uncertainty is modeled given uncertainty in vehicle pose. Ross et al.[12] provide an example of explicitly using a 2D Kalman filter on object state receiving 2D estimates from ray intersections with a ground plane, however, the measurement covariances are entirely arbitrary (and highly optimistic).

To date, the geolocation filtering literature has not appropriately dealt with issues of large uncertainty in UAV pose resulting from low-grade sensor measurements, and current solutions lack precision in estimation which would otherwise be possible given careful algorithmic formulation. As has been observed in much of the relevant literature cited[4,12,13] and corroborated by our own experimental data, error in vehicle heading is often the dominant source of error. When combined with other sources of uncertainty in other attitude dimensions, GPS position error, camera calibration errors, and terrain model error, the correct shape of the likelihood distribution over object location is in fact a 3D crescent-like shape shown in Figure 1. The typical ellipse around an estimate of object location utilized by a Gaussian Kalman filter cannot adequately

represent this and must greatly under- or over-estimate the actual uncertainty. This paper studies such non-standard uncertainty distributions present in geolocation scenarios and shows that far greater accuracy can be achieved when the distribution is properly modeled.

III. Generating Geolocation Observations

In this work we consider in particular the category of UAVs that are typically up to several kilograms in mass and possessing a wingspan of one to three meters, and we performed our experiments on a commonly fielded commercial UAV providing relatively low-bandwidth pose information derived from the on-board sensors. The ideas presented here likely generalize to significantly larger or smaller UAVs for which minimizing the size and cost of the sensor payload are primary design criteria. We assume the presence of one or more cameras, which may commonly be installed on a reposition-able gimbal or fixed as fully or slightly downward-facing. In our case, both a downward-angled front and side camera are available, of which the side camera is used while orbiting an object.

Object observations are first made by designating a pixel in a video frame as representing the centroid of an object of interest. This pixel may be an arbitrary pixel directly highlighted by an operator or assumed to be the center of the frame when a user indication is provided that the camera points directly at the object as implemented by Quigley et al.[13]. In our case, an operator initially indicates the location of an object, which is then tracked automatically through succeeding video frames by a variation of either the adaptive mean-shift developed by Collins and Liu[14] or the Lucas-Kanade algorithm[15]. An example from the execution of the latter tracker is provided in Figure 2. Any similar algorithm that can maintain track of object's location within a video stream would be a suitable input for the method being described in this paper.

Assuming the typical pinhole camera model, the relationship between a object's 3D location in the camera reference frame $\mathbf{x}_{\text{obj}}^{\text{cam}}$ and its 2D image projection $\mathbf{x}_{\text{obj}}^{\text{img}}$ is via the 3x3 camera calibration matrix $\mathbf{M}$ by

$$\begin{bmatrix} \mathbf{x}_{\text{obj}}^{\text{img}} \\ 1 \end{bmatrix} \equiv \mathbf{M}\mathbf{x}_{\text{obj}}^{\text{cam}} \tag{1}$$

when expressed in homogeneous coordinates.

Inverting this relationship provides a ray parallel to the vector $\hat{\mathbf{v}}_{\text{obj}}^{\text{cam}}$ in the camera reference frame given by

$$\hat{\mathbf{v}}_{\text{obj}}^{\text{cam}} \equiv \mathbf{M}^{-1} \begin{bmatrix} \mathbf{x}_{\text{obj}}^{\text{img}} \\ 1 \end{bmatrix}, \tag{2}$$

again expressed in homogeneous coordinates.

The object's position in world coordinates $\mathbf{x}_{\text{obj}}^{\text{world}}$ must then satisfy

$$\mathbf{x}_{\text{obj}}^{\text{world}} = \mathbf{T}_{\text{UAV}}^{\text{world}} \mathbf{p}_{\text{cam}}^{\text{UAV}} + d\mathbf{R}_{\text{UAV}}^{\text{world}} \mathbf{R}_{\text{cam}}^{\text{UAV}} \hat{\mathbf{v}}_{\text{obj}}^{\text{cam}}, d \in \mathbb{R}, \tag{3}$$

where d is the distance to the ground object along the ray, $\mathbf{T}_i^j$ is the pose transformation taking coordinates in i to j, $\mathbf{R}_i^j$ is the corresponding rotation matrix, and $\mathbf{p}_i^j$ is the position of body i in the coordinate system of j. This defines a ray in space passing through both the camera and the object. Depending on the geolocation method used, some intersection of multiple such rays may be computed across multiple observations, or a single ray may be intersected with a model of the terrain, corresponding to solving for the value of d. If using the particularly simple model assuming a ground plane of constant height z_{terrain}, d is given explicitly by

$$d = \frac{z_{\text{terrain}} - z_{\mathbf{p}_{\text{cam}}^{\text{world}}}}{z_{\hat{\mathbf{v}}_{\text{obj}}^{\text{world}}}}. \tag{4}$$

It is worth taking a moment to note how severely even relatively small inaccuracies in the transformation matrices in the preceding equations may affect the accuracy of the geolocation result. Suppose the UAV is stable 100 m above flat horizontal terrain and using a camera pitched downward by 30° aimed directly at a object. The range to the object will be 200 m and the distance along the ground approximately 173 m. At these distances, a mere 5° error in heading will throw the estimate of the object off by approximately

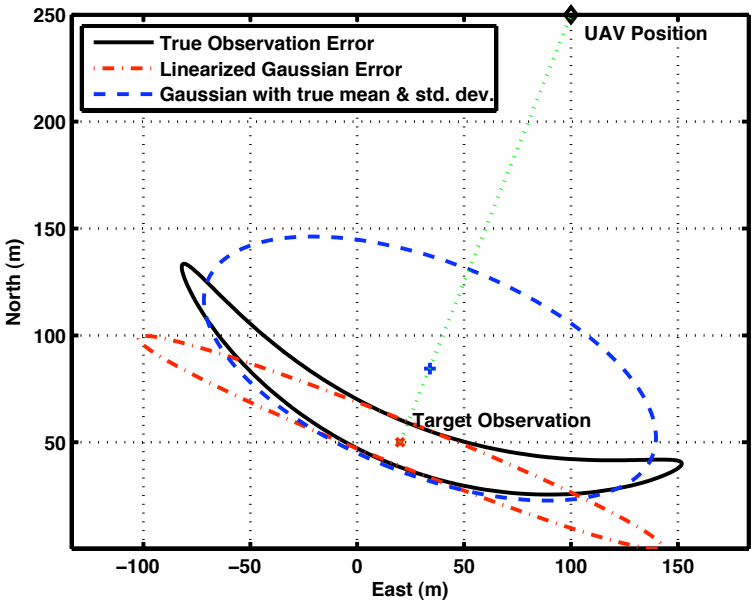


Figure 1. An illustration of the key issues pertaining to representation of geolocation measurement uncertainty. Level curves of the actual observation error distribution are crescent-shaped due to large heading uncertainty and not well approximated by a Gaussian, both in shape and likelihood mass. Linearizing about the true mean does not capture the general crescent shape, while a more conservative Gaussian is less informative and requires placement of the mean far from the actual one. Our method represents each observation as a discrete sampling of the uncertainty. This sampling naturally forms the true crescent-shape when projected onto the ground.

Figure 2. An example frame from the execution of our visual tracking algorithm run on a live UAV video stream, in this case tracking a car near a road. The centroid of the yellow corners is the current location of the object of interest.

$173sin(5°) \approx 15$ m, and a 5° error in pitch will contribute about $200sin(5°) \approx 17$ m. Yet, the small, low-cost sensors common on the types of UAVs we consider may be considerably worse.

IV. Representation of Observations

Given a series of observations providing rays passing through an object of interest, one obvious method of deriving the object's location is to simply intersect these rays. Ideally, given perfectly accurate sensing and a stationary object, these rays will intersect precisely. In reality, optimizing for the point in space that minimizes the distance to the rays is necessary. This method has the benefit of not being reliant upon intersecting rays with an unavailable or potentially inaccurate terrain model, however such a model can offer a powerful prior in helping to constrain the possible location of a object known to lie on the ground. Further, vehicle pose errors may result in rays whose intersection is far from the actual object or indeed which do not have a clearly best point of nearest intersection. Madison et al.[4] purport to provide a means to eliminate static sensor biases using a Kalman filter whose measurements are such rays, and such a method is ineffective when sensor errors do not take the form of a static or very slowly changing bias.

When intersecting the rays generated by each observation with a terrain model, the output is a series of points in world space. Given the large error in any one of these, filtering is required for an accurate geolocation result. The most straightforward means of combining these would of course be to simply average these points, however any error biases that are not symmetric in the world frame will result in a net error in the geolocation estimate. As Figure 3 depicts, such errors are in fact typical and serve to send the mean far from the object's actual position. An extension to this described by Barber et al.[11] performs nonlinear optimization to estimate constant offsets in the vehicle's pose estimate, but a constant offset does not come close to fully accounting for the errors present in practice. Indeed, for a representative dataset, in which vehicle heading error dominates, Figure 4(b) demonstrates that a batch method solving for the optimal constant heading offset does not increase geolocation accuracy significantly, as the actual optimal heading offset varies immensely over time due to ± 20 degree errors in on-board heading readings as shown in Figure 4(c). These measurements are taken from the magnetic compass of the vehicle. The heading can also be taken to be the current course of the vehicle, derived from the GPS track, but this also gives large varying errors in even moderate wind.

Explicitly modeling the uncertainty present in each distribution, as may be done using a Kalman filter, is a logical next step. The simplest such implementation might just use an arbitrary or "empirically tuned" measurement covariance in the world frame that is constant across observations, corresponding to a fixed uncertainty ellipse around each measurement. This fails to take into account the different effect that each sensor may have on error in the vehicle pose as well as how the object measurement error may vary with vehicle pose. Linearizing the image projection function at each time step permits Jacobian transformation of a covariance in sensor space to a covariance in world coordinates around a geolocation estimate, which improves performance, as it allows the ellipse to elongate along relevant axes, for instance if one dimension of the vehicle's attitude estimate is known to be worse than others. Nevertheless, as shown in Figure 1, an ellipse may provide an arbitrarily bad approximation of the actual measurement uncertainty, especially in the case of large vehicle heading uncertainty. Further, even in sensor space, a Gaussian error assumption may be unfounded if error is time-varying but systematic. Application of an Unscented Transform as is used in a UKF should improve inaccuracy in the measurement uncertainty due to linearization error, however the Gaussian error assumption is typically still made.

We present an approach that uses a discrete sampling for each observation. The sampling better approximates the actual uncertainty distribution in the object's location for a given measurement, making no assumption of constant biases or Gaussian errors in sensor inputs. We have implemented a grid-based method to filter the discrete samples and maintain a geolocation estimate which is robust to highly erroneous outlier measurements. Our method results in better geolocation accuracy when compared with other methods, even in the presence of large sensor inaccuracies.

V. Geolocation Filter

Rather than attempting to model the complex crescent-shaped uncertainty distributions of each geolocation measurement parametrically, we propose a discretized representation that captures the shape of this distribution, lying along the interior of each crescent. Towards this end, we first represent the solution space

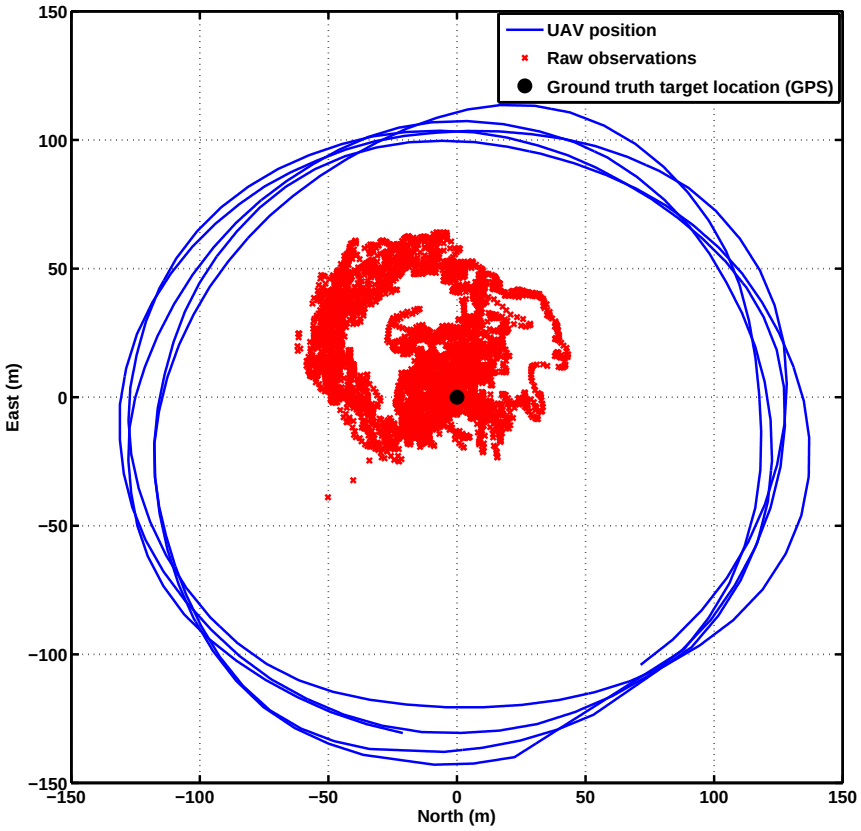
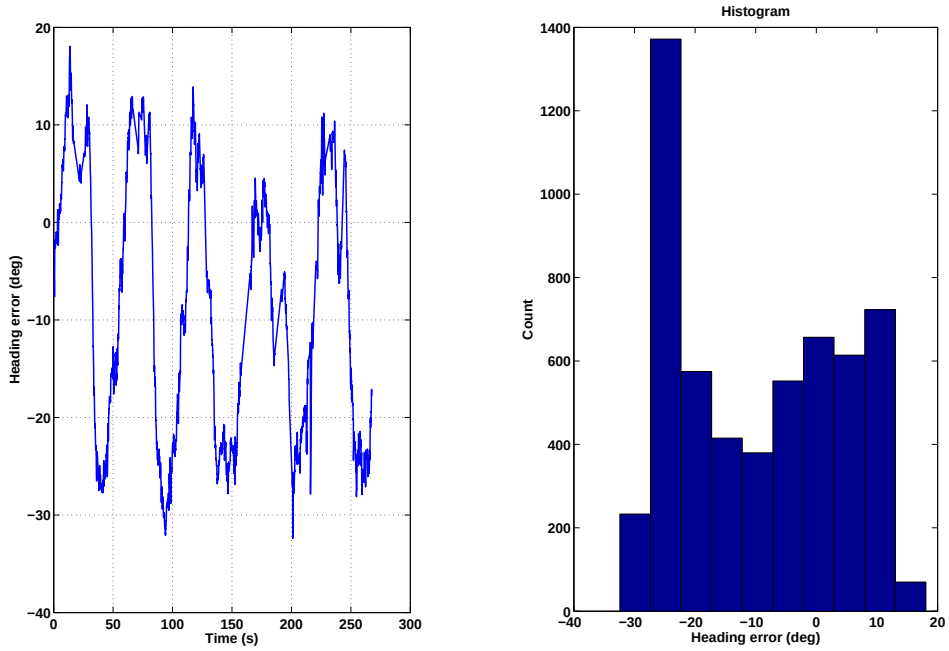
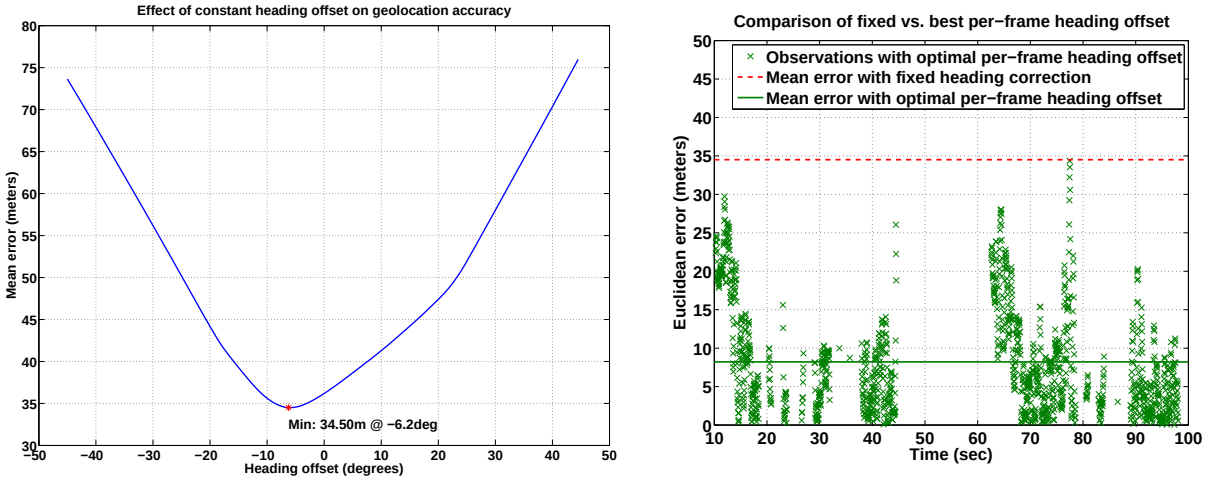


Figure 3. This graph shows the observations of an object from four complete orbits by a UAV. The observations are inaccurate primarily because of errors in the on-board sensors reporting poor heading estimates. These errors are not symmetric around the ground truth and will not naturally cancel-out with straight forward filters.



(a)



(b) (c)

Figure 4. Plots demonstrating the significant detrimental effect of UAV heading error on geolocation accuracy and the possible improvement in accuracy if just this one error source is eliminated. Plot (a) shows the actual error in heading (computed by subtracting reported from actual bearing to an object of known location), exposing heading error that is large, time-varying, and non-Gaussian in its distribution. Plot (b) shows only limited improvements in the geolocation accuracy can be achieved if a constant offset is applied to the entire dataset. While (c) shows the possible improvement in accuracy if the heading is corrected for each individual observation. This assumes that the true heading is known. It is not obvious how to obtain true heading, which is why this paper presents a filter that can accommodate large heading errors. Both (b) and (c) are taken from a dataset of a UAV orbiting an object once. Gaps in the data are from times where the object is not in view of the UAV's camera.

– the location of the object in world coordinates – as a discrete grid, in which each cell contains the relative likelihood of the object being in that cell.

The essential idea at this point is to sample the error distribution for each of the components of the vehicle pose and derive—after performing this across a number of observations—the most likely object location. Formally stated, let

$$\mathbf{x}_{\text{obj}}^{\text{world}} = G(\mathbf{s}_{\text{UAV}}, \mathbf{e}), \tag{5}$$

where the function $G(\cdot)$ takes the current state of the UAV, $\mathbf{s}_{\text{UAV}}$, and a specific offset in the UAV state, $\mathbf{e}$, to get a measurement for the location of the object in world coordinates, $\mathbf{x}_{\text{obj}}^{\text{world}}$. Executing $G(\cdot)$ to produce $\mathbf{x}_{\text{obj}}^{\text{world}}$ is the equivalent of solving for d in Equation 3. Our approach is to use $G(\cdot)$ to provide discrete geolocation hypotheses from a distribution of UAV state errors, $\mathbf{E}$, each comprised of many discrete state offsets, $\mathbf{e}$. The offsets that form $\mathbf{E}$ are sampled from the modeled error distribution of the UAV state. The error model comprises all dimensions of the vehicle pose, with the heading error being the dominant dimension in our case.

Next, let the pair of functions

$$(i, j) = \text{world2grid}(\mathbf{x}^{\text{world}}) \tag{6}$$

$$\overline{\mathbf{x}}^{\text{world}} = \text{grid2world}(i', j') \tag{7}$$

define the mapping from any point in world frame to the integer index (i, j) in the grid, and from any integer index (i', j') in the grid to world coordinates, respectively.

Given this, we may define the algorithm formally as shown in Algorithm 1. Starting with a blank grid representing the discretized world space comprised of $M \times N$ cells, we consider all states of the vehicle $\mathbf{S}_{UAV} = \{\mathbf{s}_{UAV}\}$ at which the object was observed during a flight trajectory. For each of these at which the object is identified at a given pixel, we loop through a discrete sampling of the space of possible error values $\mathbf{e}$ along the sensor or state dimensions constituting the error sources. For each of these, the terrain intersection function $G(\cdot)$ is applied to compute the world position of the object given this $\mathbf{e}$. We utilize principles of Kernel Density Estimation (KDE)[16] to estimate a continuous relative probability of object location by imposing a likelihood kernel function $K(\cdot)$ around each sampled world position. A two dimensional Gaussian kernel with standard deviation of one grid cell (representing the kernel bandwidth, here designated h) was selected in our implementation. The updates to the grid can be computed incrementally for each frame which make the filter efficient. To extract the estimate from the filter we compute a weighted mean, $\mathbf{x}_{\text{est}}$, over the grid.

Input: N, M, $\mathbf{S}_{\text{UAV}} = \{\mathbf{s}_{UAV}\}$, $\mathbf{E} = \{\mathbf{e}\}$, h
Output: Best geolocation estimate $\mathbf{x}_{\text{est}}$

grid$[1..M][1..N] = 0$
foreach $s_{\text{UAV}} \in S_{\text{UAV}}$ **do**
  foreach $\mathbf{e} \in \mathbf{E}$ **do**
    $\mathbf{x}' = G(\mathbf{s}_{\text{UAV}}, \mathbf{e})$
    for $i = 1$ *to* M **do**
      for $j = 1$ *to* N **do**
        $\mathbf{x} = \text{grid2world}(i, j)$
        $\text{grid}[i][j] = \text{grid}[i][j] + K_h(\mathbf{x}^{(k)} - \mathbf{x}'^{(k)})$
$\xi = \sum_{i=1}^{N} \sum_{j=1}^{M} (\text{grid}[i][j])$ `// Compute normalization factor`
$i_{\text{est}} = \sum_{i=1}^{N} \sum_{j=1}^{M} (i \times \text{grid}[i][j]/\xi)$
$j_{\text{est}} = \sum_{i=1}^{N} \sum_{j=1}^{M} (j \times \text{grid}[i][j]/\xi)$
$\mathbf{x}_{\text{est}} = \text{grid2world}(i_{\text{est}}, j_{\text{est}})$

Algorithm 1: Formal description of the object geolocation algorithm presented in this paper. Section V provides symbol definitions and an intuitive explanation of its operation.

Formation of the samples in the error distribution $\mathbf{E}$ depends on the vehicle uncertainty model. In our particular experiments we have noted that the heading of small UAVs is particularly bad, with a non-Gaussian distribution, as depicted in Figure 4. Thus, we incorporate heading uncertainty as a uniform

distribution over ±45 degrees. The errors of the other pose dimensions (roll, pitch, and position) are less profound and are noted from our flights to be closer to Gaussian distributions. The specific values we use are zero mean Gaussians with three standard deviations of 5 degrees for roll and pitch, and 7 meters for position.

We take 2000 samples for each observation, and project these into a grid which we have set to be 500m by 500m, with each cell being 5m by 5m. An example of a single observation from our flight tests can be seen in Figure 5. There is an obvious trade-off between the number of samples, the number of cells, and the computational load of the filter: more cells and more samples will increase filter accuracy, but this comes at the expense of computation.

To give the reader an illustration of the filter in operation, Figure 6 shows six different snapshots of the filter from actual data of the UAV orbiting an object, with these specific uncertainty distributions. In this figure, opaque green areas depict cells with high likelihood of object location, residing at the intersection of many observations. The final result of the filter from the aforementioned time-lapse is visualized in Figure 7. It is clear that errors in the raw observations are not symmetric about the actual object's location, making it difficult for many of the various alternative filters previously mentioned to converge anywhere near this, but the proposed filter finds a solution very close to the actual location.

This filter possesses many desirable traits. The first is its minimalist nature: a discrete sampling representation allows the capture of the precise shape of very complex measurement uncertainty distributions. Further, highly erroneous outlier observations do not adversely affect the filter, as long as the majority of all the observations are inliers. Lastly, filter convergence can be readily observed using visualizations such as the provided figures, allowing an operator immediate visual feedback as to convergence in the geolocation estimate.

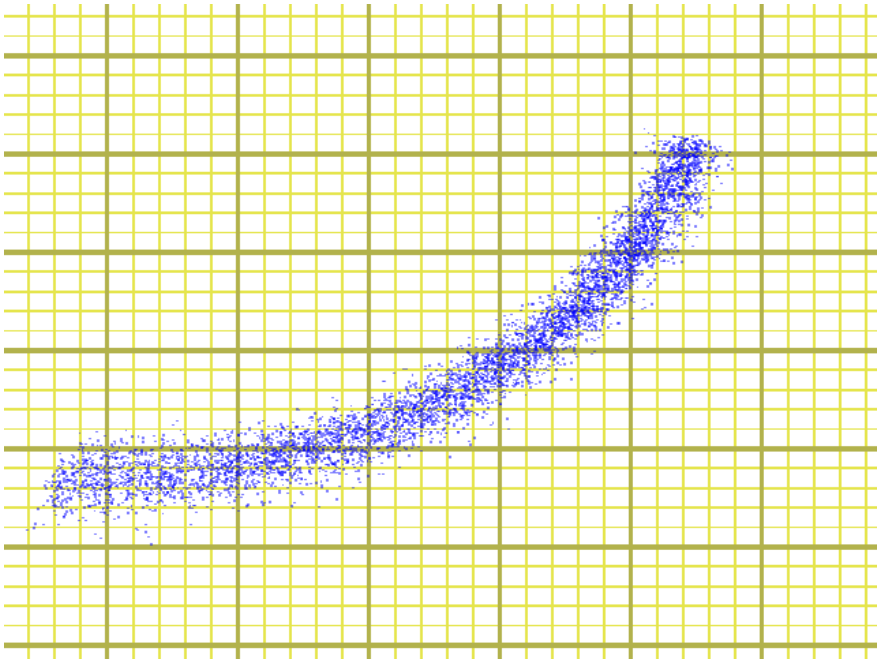


Figure 5. A figure illustrating the form of a single observation in our filter. We sample from the UAV uncertainty distribution and project the samples onto the ground plane. The samples are then passed into the grid-based filter, which maintains the geolocation estimate over many observations. A visualization of the filter can be seen in Figure 6. Scale: large grid cells are 50×50 meters.

VI. Experimental Results

The proposed geolocation filter has been tested during a number of live flights over several months of field trials, in wind conditions ranging from near still to as strong as the UAV's flying abilities would permit, and for which errors in both heading and altitude varied greatly. During operation, our visual tracking algorithm runs in real-time providing observations from full-frame-rate 30Hz video to the geolocation filter,

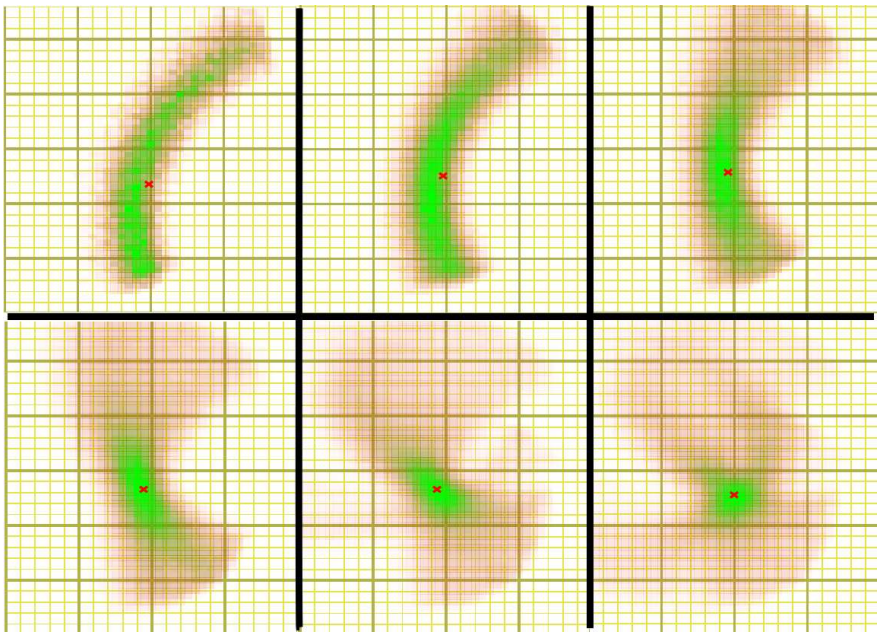


Figure 6. Visualization of the geolocation filter over time. There are eight snapshots in chronological order left to right on each row and then in order top to bottom. Here the filter applied to the case of a UAV orbiting and observing an object on the ground. The filter is a probabilistic 2D grid of cells (the two axes are the North and East location of the object). The cells with higher likelihood of the object being at that location are colored opaque green and those with less likelihood are transparent red. Scale: large grid cells are 50×50 meters.

which also runs in real-time accumulating observations of ground objects in an incremental fashion. As obtaining actual ground truth pose of a small UAV in large outdoor environments is difficult, making actual errors in its reported state challenging to quantify over long periods with high fidelity. Instead, a qualitative understanding of these errors is achieved by comparing reported and actual elevation while on the ground, as well as comparing reported heading against the bearing to known landmarks visible in imagery while the UAV is in the air.

The table also provides accuracy results for geolocating a static object on relatively flat terrain over a number of representative trials. Our proposed sampling filter is compared against the accuracy of two other common filtering methods. The first is simply the mean of all observations, and the second is an Extended Kalman Filter, which uses the previously described method of linearizing the terrain intersection function to compute Jacobians through which a covariance in sensor space is mapped to a covariance in world measurement space. As the object being geolocated is not moving, the Kalman filter is instantiated with a stationary process model.

Errors are calculated as Euclidean distance to the ground truth position of the object, as measured by a high-end GPS unit reporting uncertainty of below one meter.

Table 1. Details and errors recorded from flight tests

Flight #	Wind	Heading	Altitude	Mean	Kalman	Sampling
1	High	Moderate	Moderate	17.5	30.6m	10.6m
2	Moderate	High	Moderate	22.0m	28.1m	2.9m
3	Moderate	Very high	High	8.7m	15.3m	4.5m
4	Low	Moderate	Moderate	13.4m	19.7m	5.4m
5	High	Moderate	Moderate	29.6m	35.2m	8.0m
6	High	High	High	7.3m	15.4m	7.7m
7	Low	High	High	15.0m	22.2m	5.5m
Mean				16.2m	23.8m	6.4m

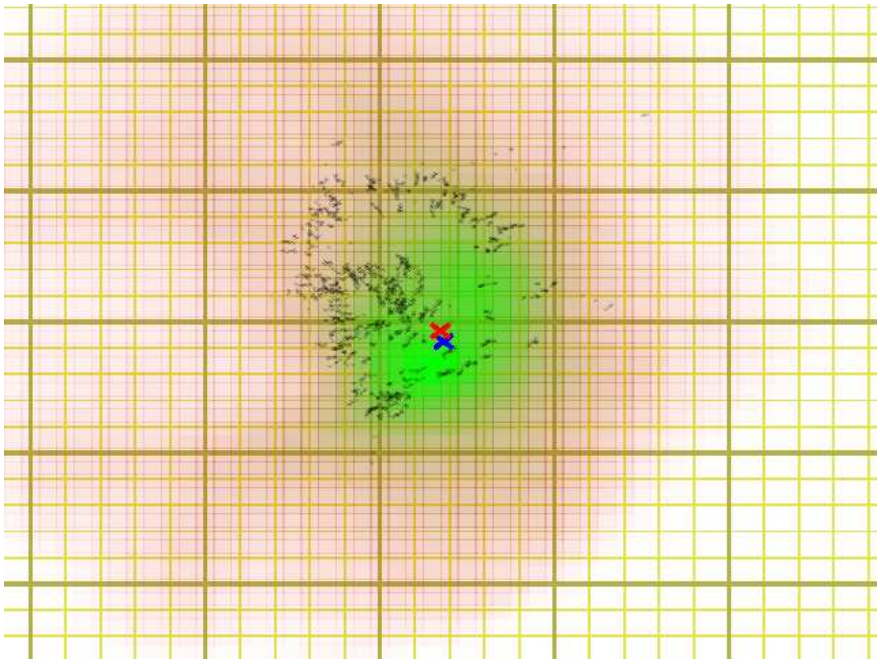


Figure 7. Visualization of the geolocation filter from the same flight as Figure 6. The small dots are the observations if they were naïvely projected without any uncertainty representation. The cells with higher likelihood of the object being at that location are colored opaque green and those with less likelihood are transparent red. The blue cross is the actual location of the object, and the red cross is the estimate extracted from the filter. Scale: large grid cells are 50×50 meters.

A particularly clear example of the strengths of the filter is the plot of the results from flight number 2 shown in Figure 8. In this trial, raw observations lay within a roughly 100 m square area centered some distance away from the object. The proposed filter outputs answers within 5 m of the correct location, while the comparison methods converge well away from this.

In other situations, such as flight number 1, results of which are shown in Figure 9, the UAV was buffeted quite badly by the wind and as such the attitude estimate of the vehicle was worse than in other flights. This results in a reduction of accuracy in geolocation.

VII. Conclusions and Future Work

This paper proposed and demonstrated an uncertainty representation that provides accurate geolocation of ground objects of interest in the presence of significant UAV heading uncertainty, a situation typically encountered in small UAV operations. Conventional linearized Gaussian Kalman filtering and fixed-offset optimization methods fall short in their representations of the true nature of the uncertainty distributions. The proposed geolocation uncertainty representation avoids the complexity associated with attempting to recover accurate UAV pose. Our approach focused on geolocating stationary objects using the limited sensing and computation available in fielded small commercial UAV systems.

We explicitly modeling the vehicle's state uncertainty, with the heading being a dominant source of uncertainty. We then take discrete samples from this uncertainty model and combine them in a grid filter. The filter has many desirable properties: it provides a natural framework for combining these complex non-linear observations, it is robust to gross outlier observations and it is computationally efficient.

The filter is shown to improve geolocation accuracy in seven distinct field flight tests with small commercial UAVs. The comparison with implementations of conventional approaches illustrated that our approach yields lower estimation errors. In some flights our approach even demonstrated several-fold improvements in accuracy: errors of a few meters as opposed to several tens of meters.

Anticipated extensions to this filter framework include schemes for representing object motion, providing a quantitative estimate confidence measure, and a hierarchical multi-resolution implementation that aims to eliminate nearly all practical drawbacks of using a discretized representation. Future work will also in-

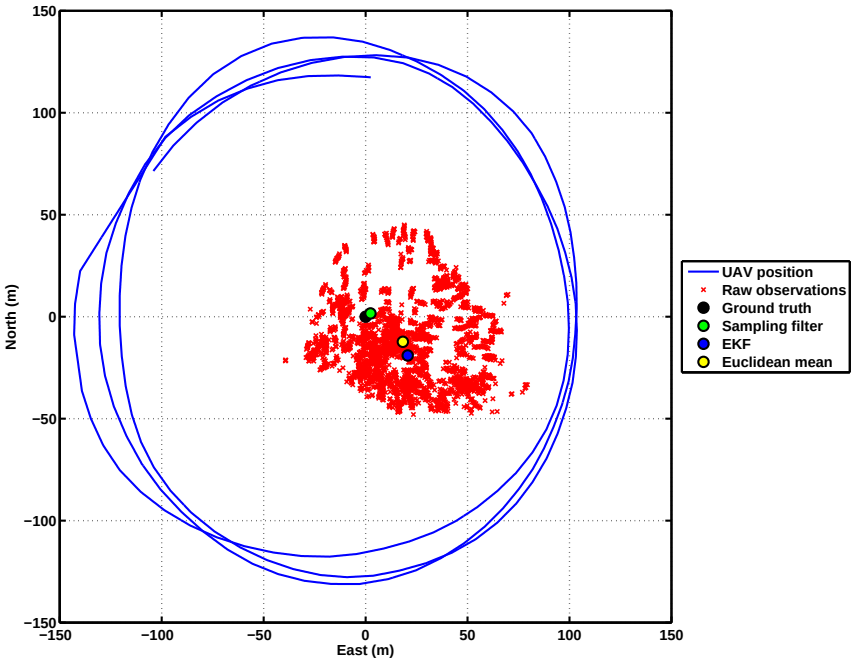


Figure 8. Graph showing the geolocation results of flight test number 2. The proposed filter is compared against a Kalman filter implementation and a Euclidean mean implementation. In this particular example, the dominant source of error is the heading with other sources contributing minimally, making our representation ideal.

vestigate the use of parametric representations with invariance to UAV error that naturally enable recursive tracking of moving objects. Additional improvements are anticipated through combining appropriate geolocation uncertainty representation with visual processing algorithms that increase UAV pose accuracy and bandwidth. Experiments to date involved tasking the UAV to orbit the best estimate of the object location as derived by the filter. Future development will implement UAV trajectory controllers that actively maneuver to improve geolocation accuracy and perform collaborative tracking by multiple UAVs.

VIII. Acknowledgements

This work is funded by the US Army ARDEC contract W15QKN-08-C-0029 under subcontract to iRobot Corporation. The authors wish to thank ARDEC and Carol Cheung and Mark Moseley from iRobot for their support.

References

[1]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.

[2]Bryson, M. and Sukkarieh, S., "Bearing-Only SLAM for an Airborne Vehicle," *Australasian Conference on Robotics and Automation*, 2005.

[3]Caballero, F., Merino, L., Ferruz, J., and Ollero, A., "Vision-Based Odometry and SLAM for Medium and High Altitude Flying UAVs," *J. Intell. Robotics Syst.*, Vol. 54, No. 1-3, 2009, pp. 137–161.

[4]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.

[5]Andersen, E. and Taylor, C., "Improving MAV pose estimation using visual information," *IEEE International Conference on Intelligent Robots and Systems (IROS)*, November 2007, pp. 3745–3750.

[6]Conte, G. and Doherty, P., "Vision-based unmanned aerial vehicle navigation using geo-referenced information," *EURASIP J. Adv. Signal Process*, Vol. 2009, 2009, pp. 1–18.

[7]Sim, D. G., Park, R. H., Kim, R. C., Lee, S. U., and Kim, I. C., "Integrated position estimation using aerial image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 1, 2002, pp. 1–18.

[8]Gibbins, D., Roberts, P., and Swierkowski, L., "A video geo-location and image enhancement tool for small unmanned

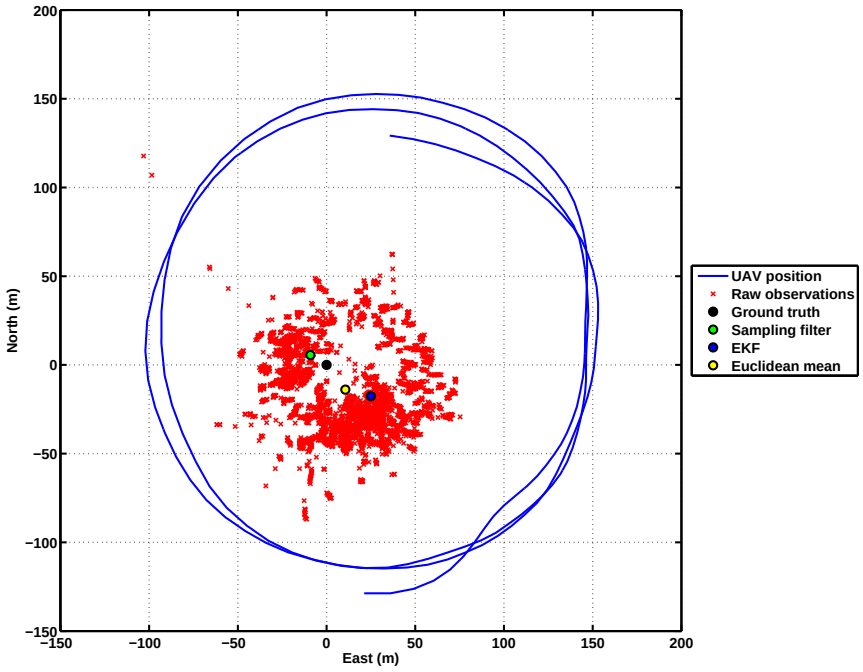


Figure 9. Graph showing the geolocation results of flight test number 1. On this flight test the wind was high and the UAV was buffeted quite severely. The attitude sensors of the vehicle reported poor estimates and therefore the geolocation accuracy was low. One possible solution to this case is to achieve better attitude estimates from a vision-based method.

air vehicles (UAVs)," *Intelligent Sensors, Sensor Networks and Information Processing Conference*, Dec. 2004, pp. 469–473.

[9]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.

[10]Wang, I. H., Dobrokhodov, V. N., Kaminer, I. I., and Jones, K. D., "On Vision-Based Target Tracking and Range Estimation for Small UAVs," *AIAA Conference on Guidance, Navigation, and Control*, August 2005.

[11]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.

[12]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.

[13]Quigley, M., Goodrich, M., Griffiths, S., Eldredge, A., and Beard, R., "Target Acquisition, Localization, and Surveillance Using a Fixed-Wing Mini-UAV and Gimbaled Camera," *IEEE International Conference on Robotics and Automation (ICRA)*, April 2005, pp. 2600–2605.

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

[15]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.

[16]Scott, D. W. and Sain, S. R., *Multi-Dimensional Density Estimation*, Elsevier, Amsterdam, 2004, pp. 229–263.