

UAV Vision: Feature Based Accurate Ground Target Localization Through Propagated Initializations and Interframe Homographies

Kyuseo Han, Chad Aeschliman, Johnny Park, Avinash C. Kak, Hyukseong Kwon, and Daniel J. Pack

Abstract—Our work presents solutions to two related vexing problems in feature-based localization of ground targets in Unmanned Aerial Vehicle (UAV) images: (i) A good initial guess at the pose estimate that would speed up the convergence to the final pose estimate for each image frame in a video sequence; and (ii) Time-bounded estimation of the position of the ground target. We address both these problems within the framework of the Iterative Closest Point (ICP) algorithm that now has a rich tradition of usage in computer vision and robotics applications. We solve the first of the two problems by frame-to-frame propagation of the computed pose estimates for the purpose of the initializations needed by ICP. The second problem is solved by terminating the iterative estimation process at the expiration of the available time for each image frame. We show that when frame-to-frame homography is factored into the iterative calculations, the accuracy of the position calculated at the time of bailing out of the iterations is nearly always sufficient for the goals of UAV vision.

I. INTRODUCTION

The goal of ground target localization is to determine the center-of-mass position of one or more ground targets in a world coordinate system using vision-enabled UAVs. Given input aerial images, ray-tracing is widely used owing to its limited time and computational requirements [1], [2], [3]. However, ray-tracing has poor accuracy in practice because of the difficulty in estimating the orientation of UAVs. More accurate target localization can be obtained by registering the aerial images to a geo-referenced image provided by a Geographic Information System (GIS) database [4], [5], [6].

In general, images of the same outdoor scene which are taken by different cameras and under different conditions can be substantially different. Because of this, feature based image matching is more suitable than direct image matching [7]. We use lines as our basic features for matching since line extraction is invariant to image resolution as well as modest changes in illumination and contrast. Also, lines capture the shape of man-made structures well and they can be stored and manipulated efficiently since they are defined by just their endpoints.

One common approach for line matching is to subsample the lines to obtain points and then apply the Iterative Closest Point (ICP) algorithm which iteratively solves for the transformation between two point sets. The ICP algorithm is robust against noise and missing information and is guaranteed to have decreasing error in every iteration and hence

Kyuseo Han, Chad Aeschliman, Johnny Park, and Avinash C. Kak are with School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907 han66@purdue.edu

Hyukseong Kwon and Daniel J. Pack are with Academy Center for UAS Research, US Air Force Academy, Colorado Springs, Colorado 80840 hyukseong.kwon@usafa.edu

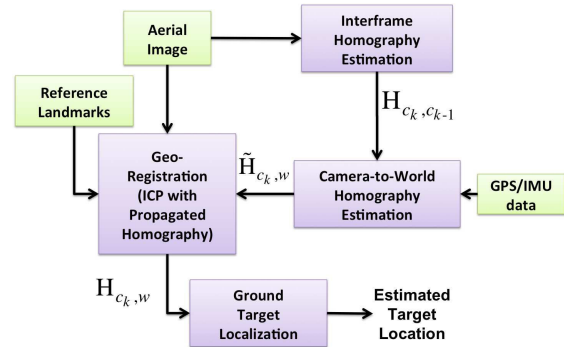


Fig. 1. Overview of the proposed ground target localization

a local minimum (but not necessarily the global minimum) is always reached [8]. Unfortunately, processing on UAVs is time-bounded, therefore convergence may not be possible within the allotted time.

To address the time-bounded problem, we bail out of ICP after a fixed number of iterations. This breaks the convergence guarantees of ICP. However, we show that by propagating information across subsequent frames we can still achieve desired convergence. The key observation is that the geo-registration solutions for two subsequent frames are related by the homography between the frames. By computing this homography, we can propagate the partial ICP results into the next frame and thus still achieve convergence as long as the error reduction in each frame is greater than the error introduced by the interframe homography. Our experimental results support our claim, with a substantial improvement in ground target localization after a few frames compared to backprojection.

A second issue with ICP is that the algorithm only guarantees convergence to a local minimum [9] and hence there is no guarantee that it will give the correct geo-registration. This problem with ICP can be solved by restarting the algorithm from a number of different initial positions and keeping the minimum cost solution [10]. Another solution is to run ICP several times using randomly selected subsets of the points and again keeping the minimum cost solution [11]. Neither solution is feasible for real-time processing on a UAV.

Fortunately, we can increase the likelihood that ICP converges to the correct geo-registration by simply using the ICP results from the previous frame to initialize ICP in the next frame – by the aforementioned homography between two subsequent frames. Each new frame captured by the

UAV reveals some new data which was not present in the previous frame. If the geo-registration in the previous frame was correct, then this new data will agree closely with the initialization and ICP will again converge to the correct solution. On the other hand, if the geo-registration in the previous frame was not correct, then the new data is unlikely to agree with the initialization and ICP will converge to a new solution. The instability of incorrect solutions and the stability of the correct solution makes it likely, though not guaranteed, that the correct solution will be achieved given enough frames. This intuition is borne out in our experimental results which show that ICP sometimes requires a few frames to converge to the correct registration, but once the correct registration is achieved it is maintained for as long as the interframe homographies can be correctly estimated.

Fig. 1 shows the high level structure of our proposed method for ground target localization. The remainder of the paper is organized as follows. Section III explains how lines are extracted from aerial images and how these lines can be used with ICP to geo-register an image. Section IV describes the estimation of the interframe homographies and how these can be used to initialize ICP. In Section V, we compare the experimental ground target localization error for three scenarios: GPS/IMU based ray tracing, geo-registration based on ICP initialized using the GPS/IMU, and geo-registration based on ICP initialized using the previous frame. Finally, we draw some conclusions in Section VI.

II. BACKGROUND

In vision-based methods, simple ray-tracing has been commonly used for target localization in UAVs due to the time and computational constraints on embedded processing. Redding et al [1] developed ground target localization based on vision from miniature air vehicles through a general localization technique using the UAV's position and orientation and the camera orientation. Monda et al [2] proposed a ground target localization method using a two-axis gimbaled camera in which they filter the ray-tracing results using both the extended Kalman filter (EKF) and the unscented Kalman filter (UKF).

Alternatively, image geo-registration has also been used for target localization. Conte et al [3] proposed a ground target localization method using direct image registration techniques based on edge extraction. The aerial edge image is registered to reference edge image using normalized cross-correlation of large square patches which may be too computationally expensive for onboard processing. Shiguemori et al [12] used texture features based on Gabor filters for describing landmarks in a supervised neural network for matching the features between captured aerial images and geo-referenced satellite images. Their proposed method, however, needs pre-processing for scaling and orientation of the aerial image using an on-board compass or altimeter. The landmarks are manually selected, e.g., railroad junctions, football fields, and overpasses. Ozguner et al [13] proposed a method based on geo-referencing aerial using Scale-invariant Feature Transform (SIFT) with their proposed

Robust Data Alignment (RDA) mechanism. They also use the UAV position and orientation generated from the UAV's IMU. Baer and Kolsch [14] matched captured aerial images to a simulated image generated from model data using the UAV's telemetry data.

ICP has typically been applied to UAVs and Unmanned Ground Vehicles (UGVs) using range sensors, such as LIDAR or sonar. Most of these studies have focused on vehicle state estimation through scan matching. Sappa et al [15] proposed a variant of the ICP algorithm based on edge points in images which used a subsampling technique to reduce the number of outliers and the computational requirements. Hsu et al [16] used ICP in a coarse-to-fine approach which matched two sets of 3D LIDAR data to detect occluded targets. Madhavan et al [17] adapted the weighted ICP for registering two different LIDAR datasets from UGVs. Also, they provided a hybrid registration consisting of feature-based matching followed by point-based matching. Their approach matches LIDAR data from UAVs to data from a UGV. In the context of indoor SLAM using micro UAVs, Sober and Johnson [18] used a laser sensor and sonar to augment an inertial measurement unit.

The main contribution of this paper is a new method for geo-registering images captured by a UAV with the goal of efficiently and accurately localizing ground targets. Our method is based on applying ICP to lines extracted from images in order to register the lines to a GIS database. Furthermore, we show that initializing ICP using past results and interframe homographies gives more robust and accurate estimation of the target location even if ICP is truncated after just a few iterations in each frame.

III. GEO-REGISTRATION USING LINE FEATURES AND ICP

As noted above, our strategy for localizing ground targets is to determine the geo-registration of each input image which has been previously shown to be an effective method [4], [5], [6]. The novelty in this paper lies in our use of line features along with ICP in order to obtain a quick and robust registration. Line features are largely invariant to illumination and contrast changes. Furthermore, by using lines we do not require the images to be of the same type, e.g., the reference GIS database could be made up of geo-referenced optical or IR images or even floor plans for man-made structures in the area. Our geo-registration algorithm is made up of two major steps: (1) extract parameterized lines from the aerial image and (2) match these lines to a set of reference lines using ICP.

A. Line Extraction in Aerial Images

We extract lines by first applying the Canny edge detector [19] and then organizing the low-level edge pixels into higher-level line features using the k-adjacent segment algorithm [20]. We then remove spurious lines – lines whose length is shorter than a prespecified threshold or lines that are connected by an acute angle – in order to produce a set of lines which are more likely to correspond to the artificial

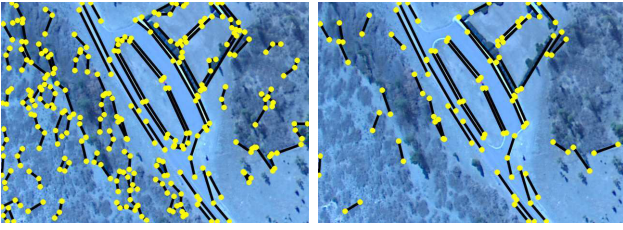


Fig. 2. An example of pruning spurious lines: initial extracted lines (left image) and refined lines after applying length and angle constraints (right image). The example shows the successful removal of spurious lines.

landmarks in the image (e.g., man-made structures such as buildings and roads). By pruning the lines in this way we reduce the computational requirements of the system and also increase the robustness of the matching process. An example of the extracted lines before and after pruning the spurious lines is shown in Fig. 2.

Before proceeding, we need to transform the data into the camera coordinate system. We map each endpoint of each line (x_i, y_i) to the corresponding camera coordinates (x_c, y_c)

$$\begin{bmatrix} x_c \\ y_c \end{bmatrix} = f_{undistort} \left(\begin{bmatrix} \frac{1}{f_x} & 0 & -\frac{c_x}{f_x} \\ 0 & \frac{1}{f_y} & -\frac{c_y}{f_y} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}; \kappa \right) \quad (1)$$

where f_x and f_y are the focal lengths of the camera, (c_x, c_y) is the principal point, and $f_{undistort}$ is a nonlinear function which removes the radial distortion in the image using a vector of distortion coefficients, κ .

B. Image Geo-Registration Via ICP

The iterative closest point (ICP) algorithm is a method for aligning two point clouds, first introduced by Besl and McKay [9]. Given points $\mathbf{p}, \mathbf{q} \in \mathbb{R}^2$, a set of data points $P = \{\mathbf{p}_1, \dots, \mathbf{p}_{n_p}\}$ and model points $Q = \{\mathbf{q}_1, \dots, \mathbf{q}_{n_q}\}$, the goal of ICP is to compute a rotation matrix \mathbf{R} and a translation vector $\mathbf{t} \in \mathbb{R}^2$ such that the transformed data points are best aligned with Q . The following is a summary of the algorithm:

- 1) **Initialization:** $k = 0$, $\mathbf{R} = \mathbf{I}_{2 \times 2}$, $\mathbf{t} = \mathbf{0}$.
- 2) **Compute the closest point:** For each point in $\hat{P} = \mathbf{R}P + \mathbf{t}$, compute the closest point in Q subject to a set of conditions (note that some points in \hat{P} may not be matched if no points in Q meet the conditions). The result is a set of m correspondences $\mathbf{p}_i \leftrightarrow \mathbf{q}_{c(i)}$.
- 3) **Compute the registration:** Given the set of corresponding pairs, compute a new \mathbf{R} and \mathbf{t} which minimize $J(\mathbf{R}, \mathbf{t}) = \frac{1}{m} \sum_{i=1}^m \|\mathbf{q}_{c(i)} - \mathbf{R}\mathbf{p}_i - \mathbf{t}\|^2$.
- 4) **Check for convergence:** Terminate if the reduction in cost is below a threshold or if a maximum number of iterations has been reached.

The final homography is given by $\mathbf{H}_{icp} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$.

In order to match line features extracted from aerial images to the GIS database, we must first apply two preprocessing steps to the data. The ICP algorithm is limited to estimating a

similarity transform. Therefore we transform the endpoints of the extracted lines based on an initial estimate of the camera to ground-plane homography $\tilde{\mathbf{H}}_{c_k, w}$ with the assumption that this estimate mostly accounts for any perspective distortion. Our method for determining $\tilde{\mathbf{H}}_{c_k, w}$ is described in the next section. Since the ICP algorithm is based on point matching, we sample each line at 1 meter intervals. This is done for both the transformed lines from the aerial image as well as the lines from the GIS database. The sample points from the aerial image become the set P in the above algorithm while the sample points from the GIS database become the set Q .

For the matching in step 2, we define the closest point to $\mathbf{p}_i \in P$ to be

$$\mathbf{q}_{c(i)} = \arg \min_{\mathbf{q}_j \in Q} \|\mathbf{R}\mathbf{p}_i + \mathbf{t} - \mathbf{q}_j\| \quad (2)$$

$$\text{subject to } \begin{cases} \|\mathbf{p}_i - \mathbf{q}_j\| \leq d_{max}, \\ \frac{|\mathbf{R}\mathbf{n}_{\mathbf{p}_i} \cdot \mathbf{n}_{\mathbf{q}_j}|}{\|\mathbf{n}_{\mathbf{p}_i}\| \|\mathbf{n}_{\mathbf{q}_j}\|} \geq \cos(\theta_{max}) \end{cases}$$

where \mathbf{n}_x is the normal vector to the line passing through the point \mathbf{x} . The maximum angle threshold θ_{max} ensures that only points from lines which roughly agree on orientation are matched. The maximum distance threshold d_{max} is intended to remove outliers during the matching process. Because of this, it is important to start with a large value for d_{max} and then progressively reduce it with each iteration. The initial maximum distance threshold d_{max} is set and updated by

$$d_{max} = \begin{cases} \max\{D_{th}, \mu_d + \lambda\sigma_d\}, & \text{first iteration} \\ (1 - \alpha)\eta_d + \alpha\xi_d, & \text{otherwise} \end{cases} \quad (3)$$

where D_{th} , λ , and α are user parameters, μ_d and σ_d are the mean and standard deviation of the matching distances from the final iteration of ICP in the previous frame, and η_d and ξ_d are the mean and max distances respectively from the preceding iteration.

By expressing d_{max} in this way, the distance threshold is always between the mean and max distance from the preceding iteration for any $\alpha \in [0, 1]$. Once the ICP algorithm converges or reaches the maximum number of iterations we compute the final estimate of the camera to ground-plane homography

$$\mathbf{H}_{c_k, w} = \mathbf{H}_{icp} \tilde{\mathbf{H}}_{c_k, w} \quad (4)$$

where $\tilde{\mathbf{H}}_{c_k, w}$ is an initial estimate of the camera to ground-plane homography. Once we have computed $\mathbf{H}_{c_k, w}$ it is straightforward to convert any \mathbf{x}_i in the image coordinate system (e.g., the image coordinates of a target) to the corresponding \mathbf{x}_w in the world coordinate system as follows:

$$\mathbf{x}_w = \mathbf{H}_{c_k, w} \mathbf{x}_c \quad (5)$$

where \mathbf{x}_c is the corresponding point of \mathbf{x}_i in the camera coordinate system.

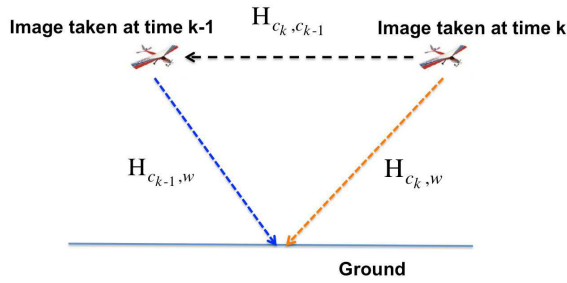


Fig. 3. Illustration of the camera to ground-plane homography of the current frame $H_{c_k, w}$, the camera to ground-plane homography of the previous frame $H_{c_{k-1}, w}$, and the interframe homography $H_{c_k, c_{k-1}}$

IV. INITIALIZING ICP USING INTERFRAME HOMOGRAPHIES

In the previous section, we discussed how ICP can be used to improve the initial estimate of the camera to ground-plane homography $\tilde{H}_{c_k, w}$. In this section we discuss how to obtain this estimate.

The simplest way to obtain the estimate $\tilde{H}_{c_k, w}$ is to use GPS/IMU information. Given the estimated position and orientation of the UAV, the desired estimate for the camera to ground-plane homography is computed by

$$\tilde{H}_{c_k, w} = H_{pos} H_R H_{c, g} \quad (6)$$

where H_{pos} is determined from the position of the UAV, H_R is determined from the orientation of the UAV, and $H_{c, g}$ accounts for the extrinsic calibration between the camera and GPS/IMU unit. While this method is fast, it often fails to give a good estimate due to the difficulty in accurately measuring the position and orientation of the UAV. For example, in our experiments the UAV is at an altitude of 220m. At this altitude, an error in the pitch or roll measurement of just 1° results in a localization error of around 4m if the UAV is directly overhead and much more than 4m if the UAV is viewing the target at an oblique angle. Difficulty in estimating the position of the UAV (particularly the height) also introduces a substantial amount of error.

A comparison of the red and green lines in Fig. 7 reveals that using ICP with $\tilde{H}_{c_k, w}$ determined from the GPS/IMU reduces the localization error by a factor of two compared to using ray-tracing. While this is a substantial improvement, much better results can be obtained by using a better initialization.

A. Using Interframe Homographies

In our method, we observe that the geo-registration solutions for two subsequent frames are related by the homography between the frames, i.e., the interframe homography. Note that this homography can typically be estimated very efficiently and accurately because the rotation and translation between subsequent frames is limited by the kinematic and dynamic constraints of UAVs. As illustrated in Fig. 3, the initial estimate of the camera to ground-plane homography

of the current frame is given by

$$\tilde{H}_{c_k, w} = H_{c_{k-1}, w} H_{c_k, c_{k-1}} \quad (7)$$

where $H_{c_{k-1}, w}$ and $H_{c_k, c_{k-1}}$ are the camera to ground-plane homography of the previous frame and the interframe homography from the current frame to the previous frame, respectively. Of course, this initialization can only be used when $H_{c_{k-1}, w}$ and $H_{c_k, c_{k-1}}$ are available, i.e., only after the first frame and only when $H_{c_k, c_{k-1}}$ can be accurately estimated. Otherwise we fall back on using the GPS/IMU to obtain $\tilde{H}_{c_k, w}$.

To estimate $H_{c_k, c_{k-1}}$, we first extract and match Speeded-Up Robust Feature (SURF) features [21] from each frame and then transform the coordinate of each feature point to the camera coordinate system. Finally, we apply Random Sample Consensus (RANSAC) to robustly estimate the homography in the presence of outliers. We evaluate the estimated homography via the number of inliers from RANSAC and the magnitude of the rotation and translation. If the number of inliers is below a threshold or the rotation and translation exceed a threshold then we discard $H_{c_k, c_{k-1}}$ and use the GPS/IMU to obtain $\tilde{H}_{c_k, w}$.

B. Analysis

Using Eq. (7) to initialize ICP can be understood as using the interframe homography to transfer the final ICP results from the previous frame to the current frame. This allows us to overcome two potential difficulties with ICP, the open-ended time requirements due to the fact that it is an iterative algorithm and the difficulty in avoiding local minima.

First, since real-time processing on a UAV is time-bounded, we limit the number of iterations of ICP to a small fixed number. Because of this, ICP is typically not able to converge if the initial error is high. However, by carrying these partial results into the next frame, in a sense ICP is able to continue the iterations. In this way, we still obtain convergence after several frames. Section V-B gives some experimental results that support the validity of this claim.

Using interframe homographies to initialize ICP also helps to reduce the problem of local minima. Given enough iterations, ICP is guaranteed to converge to a local minimum near the initialization point. In general, there is no guarantee that this local minimum will correspond to the correct geo-registration. However, if $H_{c_{k-1}, w}$ is a correct geo-registration, then using Eq. (7) to initialize ICP increases the likelihood of converging to the correct geo-registration (cf. Section V-C). On the other hand, if $H_{c_{k-1}, w}$ is not a correct geo-registration then the initialization obtained using Eq. (7) is unlikely to be near the correct local minimum and thus ICP will likely converge to an erroneous geo-registration. In practice, however, given enough frames and a rich enough scene we eventually obtain convergence to the correct geo-registration as shown by the blue line in Fig. 7.

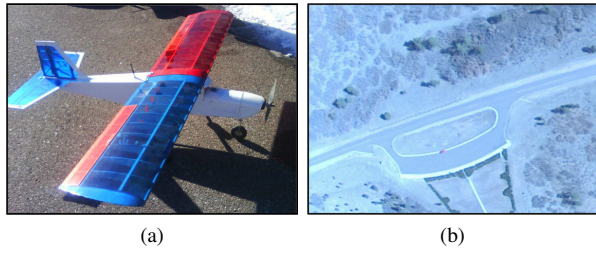


Fig. 4. Unmanned Aerial Vehicles in experiments: (a) UAV and (b) an example image captured in the experiment

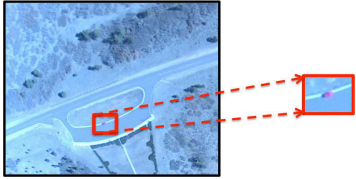


Fig. 5. The ground target marked in a red rectangle

V. EXPERIMENTS

A. Ground Target Localization

We conducted outdoor flight experiments to localize a static ground object on a road using a single UAV shown in Fig. 4. To generate the reference landmark lines, we first captured a satellite image of the applicable region from Microsoft Bing map and then manually extracted lines on the landmarks. The world geodetic system (WGS) coordinates of two end points of each line were obtained using the Microsoft Bing map API. These WGS coordinates were then mapped into a locally defined world coordinate system. The origin of the world coordinate system was set to the location of a known ground station with $+x$ to the east and $+y$ to the north. An example aerial image and the static ground target, a red car, are shown in Fig. 5. The ground truth WGS position of the ground target was manually measured by a GPS device prior to conducting the experiment.

Our dataset consists of 250 aerial images of the region containing the target. For each image, color based segmentation was used to detect the ground target with the center of mass used as the ground target location in the image coordinate system. The extracted lines from each aerial image were then registered to the reference lines using ICP. We set $\alpha = 0.95$, $D_{th} = 10\text{m}$, and $\lambda = 3$ in Eq. 3 for determining d_{max} in each iteration of ICP. The value of θ_{max} was fixed at 45 degrees for all iterations.

Fig. 6 shows corresponding pairs between query and model lines for selected iterations of ICP. In Fig. 6, the black square in the image indicates the ground truth target location, the black diamond shows the estimated target location by ray-tracing using the GPS/IMU, and the black circle shows the current estimated target location based on ICP.

Fig. 7 shows the ground target localization error over the whole dataset for three different methods while TABLE I summarizes this information. The three methods are: (1) conventional ray-tracing via GPS/IMU information; (2) ICP with the initial camera to ground-plane homography

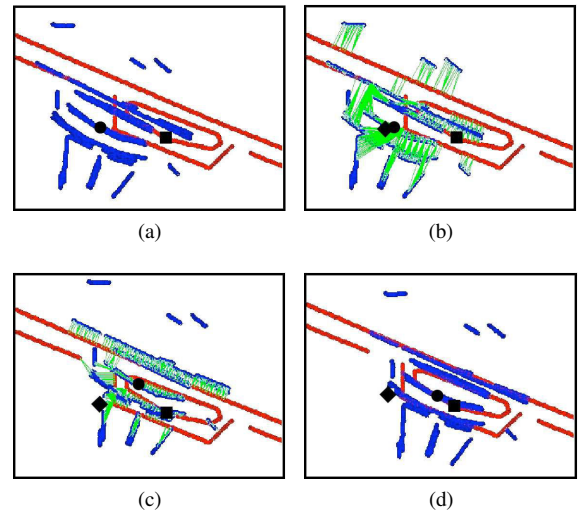


Fig. 6. ICP geo-registration example. The red lines are the reference lines while the blue lines were extracted from the aerial image. The green lines indicate corresponding points. Note that for this example, ICP was initialized using the GPS/IMU; (a) initial starting point of ICP, (b) iteration #1 of ICP, (c) iteration #20 of ICP, and (d) final result of ICP.

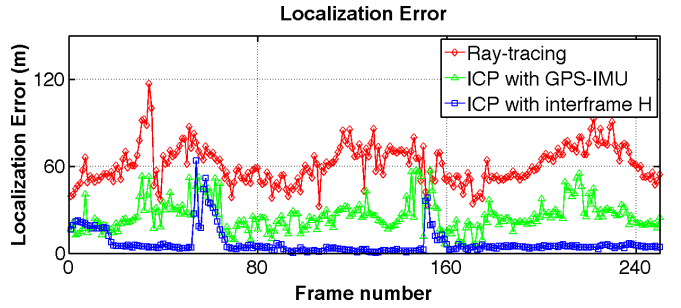


Fig. 7. The localization error by three different methods in which ICP was allowed to take an unlimited number of iterations.

computed from the GPS/IMU information; (3) the proposed method of ICP with the initial camera to ground-plane homography computed using interframe homographies. The spikes in the error for the proposed method are associated with frames in which the interframe homography could not be accurately measured and thus GPS/IMU initialization was used. The median localization error of our proposed method, 4.24m, was over five times better than ICP with GPS/IMU initialization and over fourteen times better than using ray tracing. Note that when there exist no prominent lines in the current frame, it is likely that the ICP will fail to estimate the correct homography. In this case, we set the initial estimate of camera to ground-to-plane homography $\tilde{H}_{C_k, w}$ as the $H_{C_k, w}$.

TABLE I
GROUND TARGET LOCALIZATION ERROR WITH UNLIMITED ICP ITERATIONS.

	Ray-Tracing	ICP with GPS/IMU	ICP with interframe homography
median error (m)	60.2964	24.4213	4.2373

TABLE II
AVERAGE COMPUTATIONAL REQUIREMENTS OF MAJOR STEPS IN OUR ALGORITHM.

	Line Extraction	Computing Interframe Homography	Each Iteration of ICP
Time (ms)	49.8	0.83	23.6

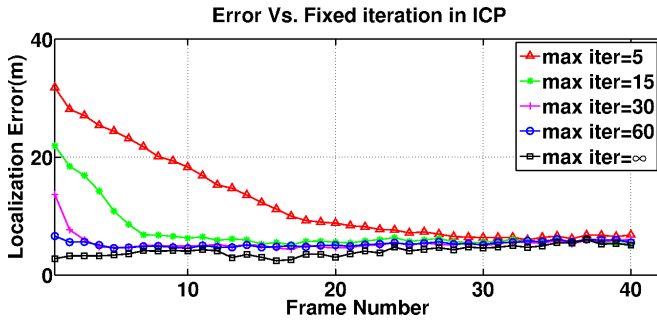


Fig. 8. The convergence of target localization error in subsequent inputs with fixed number of iteration in ICP

B. Time-Bounded ICP

For the results in the previous section, we allowed ICP to iterate until convergence. However, allowing an arbitrary number of iterations is not feasible for a real-time system. We now consider the performance of our proposed method when the number of iterations is limited to a fixed number. Fig. 8 compares the ground target localization errors for different limits on the number of iterations. The localization error of ICP with an arbitrary number of iterations is depicted by black rectangles in Fig. 8 and represents the best possible scenario. When we fix the maximum number of iterations, convergence is slowed and it takes a larger number of iterations to reach convergence. However, even with a small number of iterations the error is reduced in each frame and over a large number of frames approaches the best possible performance.

Fig. 9 shows the error reduction achieved by ICP in each frame when ICP is limited to 15 iterations. Notice that the initial error of each frame is close to the final error of the preceding frame. This is a result of using the interframe homographies to propagate the results from the previous frame into the current frame.

TABLE II shows the processing time required for the major components of our algorithm. In our system, new images are acquired at approximately 2 fps and so we have limited the maximum number of iterations to 15. Fig. 10 shows the ground target localization error for our 250 image dataset while TABLE III summarizes these results. Notice that the error increases substantially if we initialize ICP using the GPS/IMU but only increases a small amount if we use interframe homographies and previous results. This provides further evidence that our proposed initialization gives good performance even under time constraints.

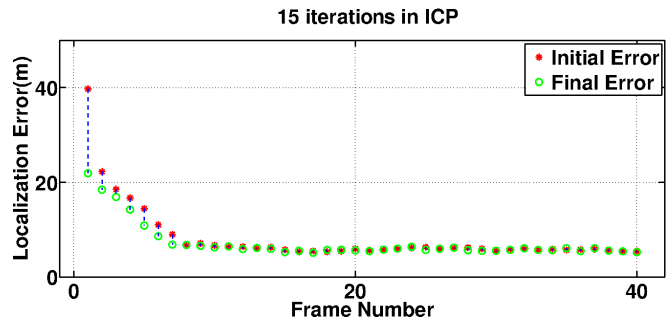


Fig. 9. Initial and final error of ICP initialized using interframe homographies.

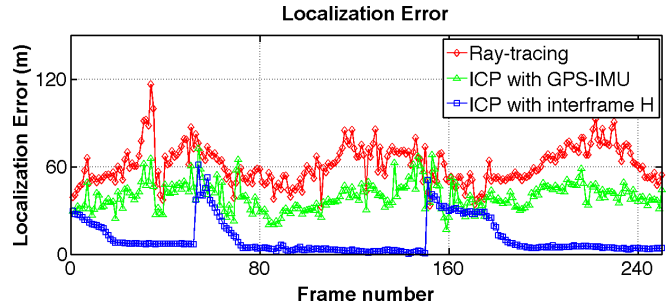


Fig. 10. The localization error by three different methods in which ICP was limited to 15 iterations.

C. Maintaining Correct Alignment

Fig. 7 and TABLE I make it clear that using interframe homographies and previous results greatly increases target localization accuracy compared to initialization with the GPS/IMU even if ICP is allowed to iterate to convergence in every frame. This is caused by the tendency of ICP to converge to a local minimum which may not correspond to the correct geo-registration.

Using the previous results to initialize ICP avoids this problem in two ways. First, since using ICP to correct the GPS/IMU initialization reduces the error by around 50%, the initial starting point obtained from previous results will typically be closer to the correct geo-registration compared to using just the GPS/IMU. This increases the likelihood that the ICP algorithm will converge to the correct geo-registration in subsequent frames.

Second, once a correct geo-registration is obtained, using these results to initialize the next frame greatly increases the chance that a correct geo-registration will be obtained again, i.e., once the correct solution is obtained for one frame it is maintained for as long as the interframe homographies

TABLE III
GROUND TARGET LOCALIZATION ERROR WITH ICP LIMITED TO 15 ITERATIONS.

	Ray-Tracing	ICP with GPS/IMU	ICP with interframe homography
median error (m)	60.2964	39.0408	5.1788

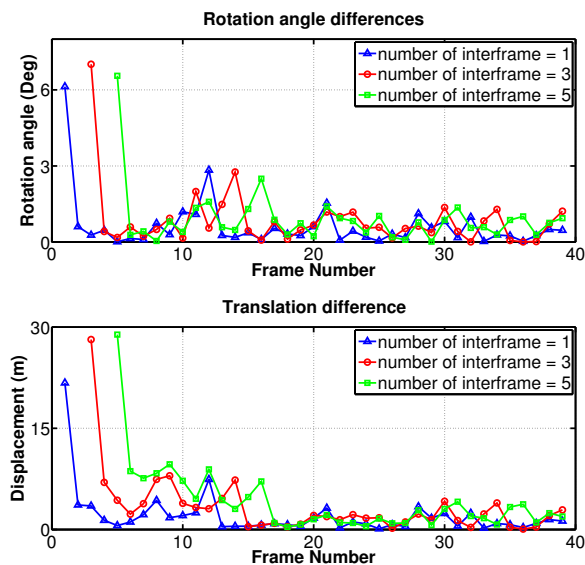


Fig. 11. Convergence evaluation for correct geo-registration

can be accurately estimated. To verify that ICP stabilizes to a consistent solution, we compared the predicted camera to ground-plane homography based on a sequence of 1, 3, or 5 interframe homographies with the actual solution obtained after running ICP. Fig. 11 shows the difference in rotation and displacement for a sequence of 40 frames. Initially there is a large amount of difference since the geo-registrations are substantially different. However, after 15-20 frames (which is after the result had converged to the correct geo-registration) the angular difference drops to around 1 degree and the displacement drops to less than 5m indicating that a stable solution has been reached.

VI. CONCLUSIONS

In this paper, we presented a new method for improving the accuracy of ground target localization based on geo-registration. We use line features to efficiently represent man-made structures in a GIS database. Extracted lines from aerial images are then registered to these lines via the ICP algorithm. In order to meet time constraints and to increase the likelihood of convergence to the correct geo-registration, we propagate the previous camera to ground-plane homography to the current frame using interframe homographies in order to initialize ICP. The experimental results show that this initialization improves the accuracy of ground target localization after a few frames even if we limit the number of iterations in ICP.

While the current system performs very well, it may be possible to improve it further by combining line features with other information such as appearance and texture. By using additional features, we should be able to obtain more accurate correspondences in ICP, making it faster and more accurate. We also plan to look into how data from multiple UAVs could be combined efficiently in order to improve accuracy and also to validate the ICP results.

ACKNOWLEDGMENT

This work was funded by AFOSR award number FA7000-10-2-0007.

REFERENCES

- [1] J. Redding, T. McLain, R. Beard, and C. Taylor, "Vision-based target localization from a fixed-wing miniature air vehicle," in *American Control Conference*, pp. 2862–2867, 2006.
- [2] M. J. Monda, C. A. Woolsey, and C. K. Reddy, "Ground target localization and tracking in a riverine environment from a uav with a gimbaled camera," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, August 2007.
- [3] G. Conte, M. Hempel, P. Rudol, D. Lundstrom, S. Duranti, M. Wzorek, and P. Doherty, "High accuracy ground target geo-location using autonomous micro aerial vehicle platform," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2008.
- [4] R. Kumar, S. Samarasekera, S. Hsu, and K. Hanna, "Registration of highly-oblique and zoomed in aerial video to reference imagery," in *15th Int. Conf. on Pattern Recognition*, 2000.
- [5] D. Hirvonen, B. Matei, R. Wildes, and S. Hsu, "Video to reference image alignment in the presence of sparse features and appearance change," in *IEEE Computer Vision and Pattern Recognition*, 2001.
- [6] D. Linh, R. Satzoda, S. Suchitra, and T. Srikanthan, "Improving robustness of real-time geo-registration process," in *IEEE Int. Conf. Neural Networks and Signal Processing*, 2008.
- [7] H. Li, B. Manjunath, and S. K. Mitra, "A contour-based approach to multisensor image registration," *IEEE Trans. on Image Processing*, vol. 4, pp. 320–334, 1995.
- [8] H. Pottmann, Q.-X. Huang, Y.-L. Yang, and S.-M. Hu, "Geometry and convergence analysis of algorithms for registration of 3d shapes," *Int J. Computer Vision*, vol. 67, pp. 277–296, 2006.
- [9] P. Besl and N. McKay, "A method for registration of 3-d shapes," *IEEE Trans. on Pattern Anal. and Machine Intel.*, vol. 14, pp. 239–256, February 1992.
- [10] D. Simon, *Fast and accurate shape-based registration*. PhD thesis, 1996.
- [11] T. Masuda, K. Sakaue, and N. Yokoya, "Registration and integration of multiple range images for 3-d model construction," in *Proc. IEEE Conf. on Comp. Vis. and Pat. Recog.*, 1996.
- [12] E. Shigemori, M. Pozzobon, and M. T. Monteiro, "Landmarks recognition for autonomous aerial navigation by neural networks and gabor transform," in *Image Processing: Algorithms and Systems V, SPIE-IS&T Electronic Imaging*, vol. 6497, 2007.
- [13] U. Ozguner, S. Jwa, J. Martin, and K. Redmill, "Robust data alignment and its application in an integrated multi-sensor surveillance system consisting of uavs and fixed platforms," Tech. Rep. AFRL-RY-WP-TR-2009-1031, Air force research laboratory sensors directorate, 2009.
- [14] W. Baer and M. Kolsch, "Shadow and feature recognition aids for rapid image geo-registration in uav vision system architecture," in *Signal Processing, Sensor Fusion, and Target Recognition XVII*, vol. 7336, SPIE, 2009.
- [15] A. Sappa, A. Restrepo-Specht, and M. Devy, "Range image registration by using an edge-based representation," in *the 9th International Symposium on Intelligent Robotic System*, 2001.
- [16] S. Hsu, S. Samarasekera, and R. Kumar, "Automatic registration and visualization of occluded targets using lidar data," in *SPIE Laser Radar Technology and Applications VIII*, vol. 5086, April 2003.
- [17] R. Madhavan, T. Hong, and E. Messina, "Temporal range registration for unmanned ground and aerial vehicles," *Journal of Intelligent and Robotic Systems*, no. 44, pp. 47–69, 2005.
- [18] D. M. S. Jr. and E. N. Johnson, "Indoor navigation for micro air vehicles," in *AIAA Infotech@Aerospace 2010*, April 2010.
- [19] J. Canny, "A computational approach to edge detection," *IEEE Trans. on Pattern Anal. and Machine Intel.*, vol. 8, no. 6, pp. 679–698, 1986.
- [20] V. Ferrai, L. Fevrier, F. Jurie, and C. Schmid, "Groups of adjacent contour segments for object detection," *IEEE Trans. on Pattern Anal. and Machine Intel.*, vol. 30, pp. 36–51, January 2008.
- [21] H. Bay, A. Ess, T. Tuyelaars, and L. Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2006.