Vision-based Target Localization from a Fixed-wing Miniature Air Vehicle

Joshua Redding Timothy W. McLain Randal W. Beard Clark Taylor

Abstract— **This paper presents a method for localizing a ground-based object when imaged from a small fixed-wing unmanned aerial vehicle (UAV). Using the pixel location of the target in an image, with measurements of UAV position and attitude, and camera pose angles, the target is localized in world coordinates. This paper presents a study of possible error sources and localization sensitivities to each source. The localization method has been implemented and experimental results are presented demonstrating the localization of a target to within 11 m of its known location.**

I. Introduction

Unmanned vehicles are prime candidates for tasks involving risk and repetition, or what the military calls the "dull, dirty and dangerous" [1]. The simplified goal of many of these activities is to image and locate a target for tracking, information-gathering or delivery purposes. Therefore, the ability to explicitly determine the location of a visible, ground-based object while only its general position is known a priori would prove beneficial to the completion of these tasks. The problem of estimating an object's exact location is called "localization". Many of the current approaches to this problem involve imaging a target from an unmanned blimp or rotor craft [2], [3], [4]. Due to their low-altitude, low-velocity flight capabilities, these aircraft allow significant simplification of the problem. However, blimps are not well suited for use in high winds or inclement weather, and the costs and complexities associated with rotor craft are high. It is therefore reasonable to explore localization methods involving more robust and less-expensive fixed-wing UAV platforms.

Fixed-wing UAVs, while lacking the ability to hover, present unique benefits such as adaptability to adverse weather, a shorter learning curve for the untrained operator and extreme durability against harsh environments. Also, the minimum airspeed requirement associated with a fixed-wing aircraft provides images from multiple vantage points, allowing for more robust localization. Man-packable, fixed-wing UAVs are small enough to be carried and operated by a single person while offering many of the same advantages as their larger counterparts. The development and deployment of such UAVs is comparatively inexpensive, making them expendable in some scenarios [5].

J. Redding is with the Intelligent and Autonomous Systems Group, Scientific Systems Company, Inc., Woburn, MA 01801, USA joshua@ssci.com

T. McLain is with the Department of Mechanical Engineering, Brigham Young University, Provo, UT 84602, USA mclain@byu.edu

R. Beard and C. Taylor are with the Department of Electrical and Computer Engineering, Brigham Young University, Provo, UT 84602, USA beard@ee.byu.edu, taylor@ee.byu.edu

The problem often exists in both military and search and rescue scenarios where an object needs to be found or photographed when only its general location is known. The goal of this research is to determine the exact location of this object or target in world/inertial coordinates using a gimballed camera on-board a small, fixed-wing UAV.

Initially finding and recognizing the target is accomplished by allowing the UAV to follow a prescribed search pattern covering the general area where the target is known to be while a color-segmentation routine searches for the highly-visible target within the camera's field of view. Once detected, localization begins immediately and continues while the camera is actively gimballed to keep the target in its image as the UAV adjusts its flight path to orbit the newly calculated estimates of the target's inertial location. These estimates stem from measurements of UAV location and attitude as well as camera pose angles and the target's pixel location in the image. Errors and uncertainties in measurements and estimates will degrade the estimate of target location. For this reason, a study of possible error sources and their propagation through the localization routine is also presented.

Vision-based localization is a well-studied problem. However, much of the research involves the self-localization of unmanned ground vehicles in controlled laboratory settings [6], [7], [8], [9]. This research extends the prior research by considering implementation on a small airborne platform operating in an unstructured environment. Closely related is the research of Chaimowicz, et al. [2], who present the results of an experiment using a tethered blimp to localize a stationary object on the ground. While hovering at an altitude of 18 m and equipped with a fixed camera, the blimp used a series of images containing both the target and a set of at least 6 known landmarks to estimate target location. Although this method led to very accurate localization and serves as a useful benchmark, it is limited. Chaimowicz also used GPS and IMU measurements from the blimp for localization, which resulted in estimates within 8.2 m of the target's GPS location.

Rysdyk [10] also conducted similar research that simulates maintaining a constant line-of-sight with a ground-based target from a fixed-wing UAV. He outlines camera gimbal control, but emphasizes UAV path-planning. Stolle [11] presents similar research, but with more details on camera control. Both Rysdyk and Stolle assume a priori knowledge of precise target location and that the on-board camera is located at the UAV center of mass.

This paper discusses research that extends the simulated

camera gimbal control done by Rysdyk and Stolle into hardware implementation on board a small, fixed-wing UAV flying at a nearly constant velocity of 13 m/s and at an altitude of 60 m. This paper presents a general approach to target localization from an airborne camera, provides an analysis of possible error sources, and demonstrates the effectiveness of the approach with experimental results.

II. Technical Approach

A simple projection camera model is shown in Figure 1. The point q is the projection of the point p^c_{obj} onto the image plane in pixels (ip), where p^c_{obj} denotes the location of an object p relative to the center of the camera. It is assumed that the location q is known in pixels and we desire to know it in meters, so that p^c_{obj} can be found using similar triangles. Trucco, et al [12] show that the change from pixels to meters in the image frame is accomplished by

$$\begin{aligned} x_{im} &= (-y_{ip} + 0_y)S_y \\ y_{im} &= (x_{ip} - 0_x)S_x \ , \end{aligned} \tag{1}$$

where 0_x and 0_y denote the necessary x and y offsets to bring the $(0, 0)$ point to the image center from the upper-left hand corner. S_x and S_y denote the image scalars for image y and x directions respectively due to the change of axes from ip to im coordinates.

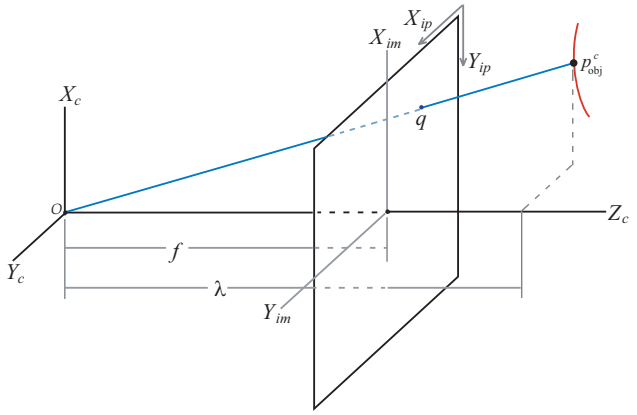

Fig. 1. Camera frames

If q is converted into meters, it can be scaled into the camera frame using the law of similar triangles and the camera focal length, f. However, since the distance λ from the camera center to p^c_{obj} is not known, scaling can only occur in two dimensions. For this reason, λ is usually extracted and the scaling process is combined with Equation (1) to form an expression for p^c_{obj} in terms of known pixel location q, as shown by

$$p^c_{\text{obj}} = \lambda C^{-1} q \ , \tag{2}$$

where C is a matrix containing the scaling information for transformation between the camera, c, and the pixel, ip, coordinate frames. This is shown by

$$\lambda q = \underbrace{\begin{bmatrix} 0 & f_x & 0_x \\ -f_y & 0 & 0_y \\ 0 & 0 & 1 \end{bmatrix}}_{C} p^c_{\text{obj}} \ , \tag{3}$$

where $\frac{f}{S_x} = f_x$ and $\frac{f}{S_y} = f_y$. As can be seen, an object can be successfully localized when λ is known. We will return to the issue of finding λ in section II-B. A few more coordinate frames are required when placing the camera in the sky, as seen in Figures 2 and 3.

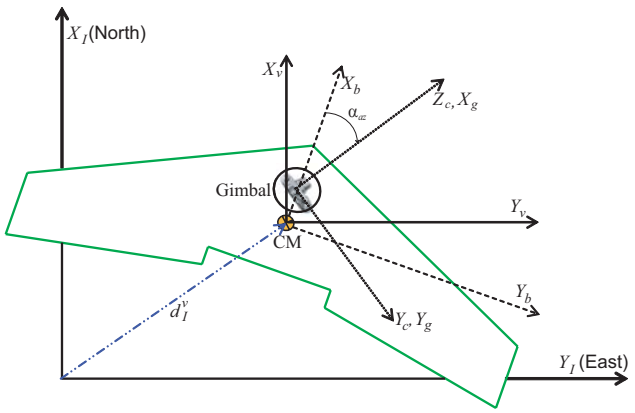

Fig. 2. Coordinate frames

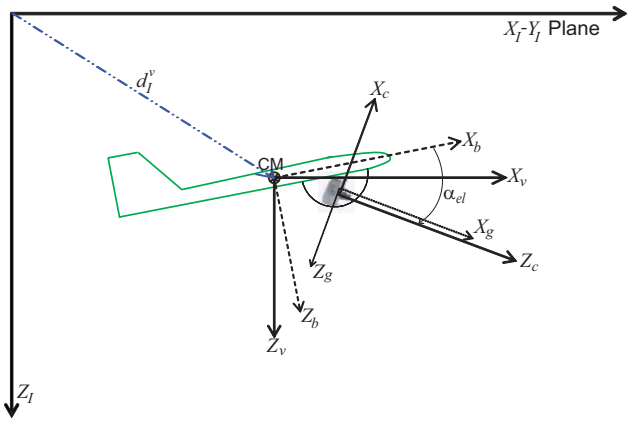

Fig. 3. Coordinate frames

All coordinate frames follow a right-hand rule. The camera frame c, has its origin at the camera's center, with the positive Z-axis, Z_c, representing the optical axis of the camera. The origin of the gimbal frame g, is the center of the two-axis gimbal. The UAV body frame is centered at the UAV center of mass, with the X-axis, X_b, out the nose of the aircraft and the Y-axis, Y_b, out the right wing. The UAV vehicle frame v, is identical to the inertial frame I, only translated so the UAV center of gravity is the origin. As it will be necessary to move between coordinate frames frequently, a set of homogeneous transformation matrices are now introduced.

A. Transformations

A homogeneous transformation matrix (HTM) combines both rotation and translation between coordinate frames into a single matrix. The structure of an arbitrary HTM, T_i^j, is shown by

$$T_i^j = \begin{bmatrix} & R & & -d_i^j \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

where R represents a (3×3) rotation matrix and d_i^j represents a (3×1) translation vector. Both rotation and translation occur from the i^{th} to the j^{th} coordinate frame. However, when written in this manner, the rotation occurs first, followed by the translation. Therefore the translation vector is resolved in the j^{th} coordinate frame and negated to avoid requiring a pre-multiplication by R_i^j. The function of each HTM is described in Table I, and this section will discuss each individually, identifying its elements and purpose.

TABLE I
HOMOGENEOUS TRANSFORMATION MATRICES.

HTM	Description
T_I^v	Transformation from Inertial to UAV Vehicle frame
T_v^b	Transformation from UAV Vehicle to UAV Body frame
T_b^g	Transformation from UAV Body to Gimbal frame
T_g^c	Transformation from Gimbal to Camera frame

1) Transformation T_I^v: Since the transformation from the inertial to the vehicle frame is actually a single translation, T_I^v will only depend on the UAV's GPS location and barometric altitude measurements shown by

$$\begin{aligned} T_I^v &= \begin{bmatrix} I & -d_I^v \\ 0 & 1 \end{bmatrix}, \text{where} \\ d_I^v &= \begin{bmatrix} x_{\text{UAV}} \\ y_{\text{UAV}} \\ -h_{\text{UAV}} \end{bmatrix}, \end{aligned} \tag{5}$$

where x_{UAV} and y_{UAV} represent the North and East location of the UAV as measured by the GPS, and h_{UAV} represents the UAV's altitude as measured by a barometric pressure sensor.

2) Transformation T_v^b: The transformation from the vehicle frame to the UAV body frame, T_v^b, consists of a single rotation based on measurements of Euler angles, shown by

$$\begin{aligned} T_v^b &= \begin{bmatrix} R_v^b & 0 \\ 0 & 1 \end{bmatrix}, \text{where} \\ R_v^b &= \begin{bmatrix} c_\theta c_\psi & c_\theta s_\psi & -s_\theta \\ s_\phi s_\theta c_\psi - c_\phi s_\psi & s_\phi s_\theta s_\psi + c_\phi c_\psi & s_\phi c_\theta \\ c_\phi s_\theta c_\psi + s_\phi s_\psi & c_\phi s_\theta s_\psi - s_\phi c_\psi & c_\phi c_\theta \end{bmatrix} \end{aligned} \tag{6}$$

where ϕ, θ and ψ represent the UAV's roll, pitch and heading angles in radians. Also, c_* and s_* abbreviate $\cos(*)$ and $\sin(*)$ respectively.

3) Transformation T_b^g: The transformation from the UAV body to the gimbal frame, T_b^g, will depend on the location of the UAV's center of mass with respect to the gimbal's rotation center. This vector, denoted by d_b^g, is resolved in the gimbal frame. T_b^g will also depend on the rotation that aligns the gimbal's coordinate frame with the UAV's body frame. This rotation is denoted R_b^g and requires measurements of the camera's azimuth and elevation angles α_{az} and α_{el} respectively, both of which are known. This transformation is shown by

$$\begin{aligned} T_b^g &= \begin{bmatrix} R_b^g & -d_b^g \\ 0 & 1 \end{bmatrix}, \text{where} \\ R_b^g &= R_{y,\alpha_{el}} R_{z,\alpha_{az}} \\ &= \begin{bmatrix} c_{el} & 0 & s_{el} \\ 0 & 1 & 0 \\ -s_{el} & 0 & c_{el} \end{bmatrix} \begin{bmatrix} c_{az} & s_{az} & 0 \\ -s_{az} & c_{az} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} c_{el}c_{az} & c_{el}s_{az} & s_{el} \\ -s_{el} & c_{az} & 0 \\ -s_{el}c_{az} & -s_{el}s_{az} & c_{el} \end{bmatrix}, \end{aligned} \tag{7}$$

where d_b^g denotes the vector from the gimbal center to the UAV center of mass, α_{az} denotes the azimuth angle of rotation about Z_g, and α_{el} the elevation angle of rotation about Y_g, after α_{az}.

4) Transformation T_g^c: T_g^c is the transformation from gimbal to camera reference frames. It will depend on the vector d_g^c, which describes the location of the gimbal's rotation center relative to the camera center and is resolved in the camera's coordinate frame. T_g^c also depends on a simple rotation R_g^c, which aligns the camera's coordinate frame with that of the gimbal. It is shown by

$$\begin{aligned} T_g^c &= \begin{bmatrix} R_g^c & -d_g^c \\ 0 & 1 \end{bmatrix}, \text{ where} \\ R_g^c &= \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \end{aligned} \tag{8}$$

since we chose $X_c = -Z_g$ and $Z_c = X_g$. Also, d_g^c denotes the vector from the camera center to the gimbal center, resolved in camera frame.

We now have four HTMs that are based on a priori calibrations and real-time measurements from on-board sensors. This means we can freely move between coordinate frames as data is collected during flight. We now can extend Equation (2) from the camera to the inertial frame, as shown by

$$p_{\text{obj}}^I = \lambda [C T_g^c T_b^g T_v^b T_I^v]^{-1} q \ , \tag{9}$$

where p_{obj}^I denotes the object location in the world, or inertial frame. Knowing all other parameters on the right-hand side of Equation (9), we are now ready to find the image depth, λ.

B. Image Depth

Image depth refers to the distance along the camera's optical axis, Z_c, to the object of interest in the image, and its value is usually unknown [13]. To estimate λ, the camera center is represented in inertial coordinates by p^I_{cc}, as shown in Figure 4 and defined as

$$p^I_{\text{cc}} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}^I_{\text{cc}} = [T^c_g T^g_b T^b_v T^v_I]^{-1} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}^c_{\text{cc}} , \qquad (10)$$

where the vector $[x\ y\ z\ 1]^{c\ T}_{\text{cc}}$ is equal to $[0\ 0\ 0\ 1]^T$, since it describes the location of the camera center in camera coordinates. Figure 4 also shows the location q, which was introduced in Figure 1.

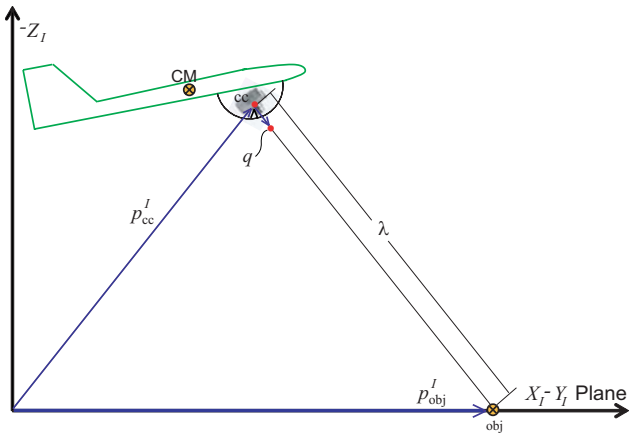

Fig. 4. Localization vectors

The location of q in the inertial frame is described by the vector $\bar{p}^I_{\text{obj}}$, which is depicted in Figure 5 and defined as

$$\bar{p}^I_{\text{obj}} = \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \\ 1 \end{bmatrix}^I_{\text{obj}} = [C T^c_g T^g_b T^b_v T^v_I]^{-1} q , \qquad (11)$$

where q contains the target's pixel location, shown by

$$q = [x_{ip}\ y_{ip}\ 1\ 1]^T . \qquad (12)$$

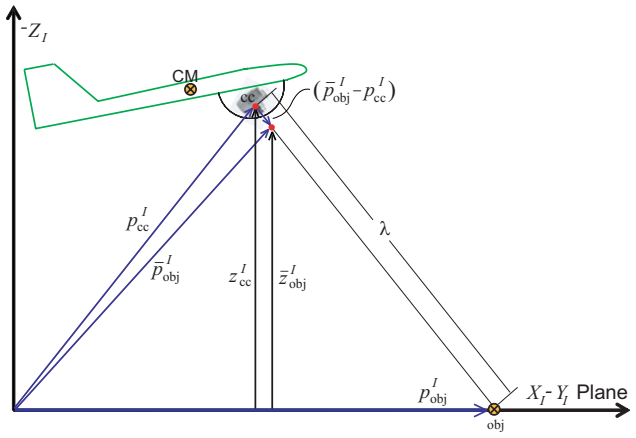

Fig. 5. Localization vectors

Referring again to Figure 5, and noting the implied assumption that the zero-altitude plane is defined where the UAV's altitude sensor was zeroed, the z components of $\bar{p}^I_{\text{obj}}$ and p^I_{cc} form the relationship

$$0 = z^I_{\text{cc}} + \lambda \left(\bar{z}^I_{\text{obj}} - z^I_{\text{cc}} \right) . \qquad (13)$$

The zero on the left-hand side of Equation (13) follows from the assumption that the target lies on this plane of zero altitude. This assumption is justified since future height-above-ground, or terrain map measurements can provide information to make the left-hand side of Equation (13) known, but non-zero. Since both z^I_{cc} and $\bar{z}^I_{\text{obj}}$ are known from Equations (10) and (11) respectively, Equation (13) can easily be solved for λ as

$$\lambda = \frac{-z^I_{\text{cc}}}{\left(\bar{z}^I_{\text{obj}} - z^I_{\text{cc}} \right)} . \qquad (14)$$

Since the inertial Z-axis, Z_I, is defined positive toward the center of the earth, z^I_{cc} will be negative for flight altitudes greater than the calibrated zero. Thus, Equation (14) yields a positive value for λ, as expected.

C. Target Location

Now that the depth of our current image is known, we can easily estimate the inertial location of the target in the image by

$$p^I_{\text{obj}} = \lambda [C T^c_g T^g_b T^b_v T^v_I]^{-1} q , \qquad (15)$$

or by continuing the method used to find λ, as shown by

$$p^I_{\text{obj}} = p^I_{\text{cc}} + \lambda \left(\bar{p}^I_{\text{obj}} - p^I_{\text{cc}} \right) . \qquad (16)$$

We see that the localization of a visible target was accomplished using only a camera and readily accessible UAV information. Noise and parameter uncertainty can significantly affect the quality of the localization estimate. Section III presents a study of error sources and their effects on target localization.

III. Error Analysis

Like all aircraft, UAVs are susceptible to outside influences including wind gusts, jet streams, variations in atmospheric pressures, air densities and temperatures. These phenomena, among others, add unwanted noise to aircraft sensors. This noise, combined with sensor errors and inaccuracies, contaminates each measurement of position, altitude, airspeed and heading as well as roll, pitch and yaw rates. The purpose of this section is to explore the main error sources in UAV and gimbal control and to study how each one individually affects the target localization result.

A. Error Sources

In the equation

$$p^I_{\text{obj}} = \lambda[CT^c_g T^g_b T^b_v T^v_I]^{-1} q, \tag{17}$$

each term introduces inaccuracies to the end result. Since λ is calculated from measurements of UAV altitude, its associated errors will be accounted for through altitude uncertainty. Errors in the camera calibration matrix, C, are comparatively small since they originate in the calibration routine itself and are therefore neglected. The transformation T^c_g depends on the location of the UAV center of mass with respect to the camera center, which can be measured to within millimeters and can also be ignored. T^g_b on the other hand, will depend on camera gimbal angles, which are not exactly known. The camera gimbal is controlled via commercially available hobby servos, which laboratory tests have shown to be accurate to less than half a degree and precise to less than one fifth of a degree. These measures are only valid when the servos are given sufficient time to reach their desired angles. The time required to do so is on the order of 5 ms/deg, which is sufficiently fast for typical changes in desired gimbal angles.

T^b_v introduces further inaccuracies through errors in UAV attitude estimation. Euler angles ϕ, θ, and ψ are estimated from gyro measurements of roll, pitch and yaw rates as well as accelerometer readings with reference to the gravity vector [14], [15]. Unfortunately, gyros tend to drift, causing accumulating errors in ψ. Estimates of ϕ and θ are generated by subtracting the gravity vector from accelerometer measurements, a technique which works well under static conditions. However, this subtraction yields degraded results when the UAV is experiencing accelerations common during flight. Through laboratory tests, ϕ, θ and ψ have been shown to be statically accurate to within five degrees, and dynamically accurate to within ten degrees.

The translation from inertial to vehicle frames, accomplished by T^v_I, adds inaccuracies that stem from both GPS measurements and barometric altitude readings. The major GPS inaccuracies are attributed to a variety of sources that combine to achieve an accuracy of roughly 10 m in the horizontal plane, and 25 m in the vertical plane [16]. For testing purposes, the known location of the target was measured using the GPS unit on the UAV. Since the bias portion of the GPS error equally effects measures of both UAV and target positions, it does not contribute to the error between the known location of the target and the estimated location of the target. Random errors in GPS measurements are average approximately 5 m in the horizontal plane. With the addition of an absolute pressure sensor, altitude inaccuracies are reduced to roughly 8 m [17].

TABLE II
Uncertainties, U_*

Source	± Value	Source	± Value
α_{az}	.5 deg	α_{el}	.5 deg
ϕ	5 deg	θ	5 deg
ψ	5 deg	x_{UAV}	5 m
y_{UAV}	5 m	h_{UAV}	8 m
x_{ip}	5 pixels	y_{ip}	5 pixels

The target pixel location, q, is also subject to uncertainties, including visual occlusions and lighting changes. Accounting for such, it is believed that q can be trusted within about 5 pixels in both x_{ip} and y_{ip}. Although actual uncertainties are not known, the values shown in Table II are the results of laboratory tests and it is assumed that they represent a 95% probability.

B. Sensitivity and Propagation

This section presents a study of localization sensitivity to uncertainties in measurements of UAV location and attitude as well as camera gimbal angles using the method of sequential perturbation [18]. The localization estimate of the target position can be expressed as a function of the UAV location and attitude and the gimbal azimuth and elevation angles according to

$$p^I_{\text{obj}} = \lambda[CT^c_g T^g_b T^b_v T^v_I]^{-1} q \tag{18}$$

$$= F\left(\alpha_{az}, \alpha_{el}, \phi, \theta, \psi, (x, y, h)_{\text{UAV}}\right). \tag{19}$$

Sequential perturbation is a numerical approach to estimate the propagation of uncertainties through to a result and is used when direct calculation of partial derivatives is not feasible, as is the case with Equation (19). Using sequential perturbation, sensitivities to errors in each variable are calculated under a nominal flight condition, in this case a large-orbit coordinated turn. These sensitivities are listed in Table III and show that errors in UAV roll angle and camera elevation angle most dramatically affect localization outcome. The fact that they are equally important is expected since during a localization flight the camera is panned to roughly ninety degrees, which aligns Y_g, the axis about which elevation occurs, with X_b, the axis about which aircraft roll occurs. When aligned in this manner, the localization algorithm cannot differentiate between changes in elevation angle and changes in UAV roll angle.

Using Table II in conjunction with Table III, the total expected localization error, Γ, can be computed [18] by

$$\Gamma = \sqrt{\sum_{i=1}^{N} \left(\frac{\partial F}{\partial i} U_i\right)^2}$$

$$= 14.9 \text{ m}, \tag{20}$$

TABLE III

NUMERICALLY APPROXIMATED PARTIAL DERIVATIVES

i	Parameter	$\partial F/\partial *$
1	α_{az}	1.1 m/deg
2	α_{el}	1.7 m/deg
3	ϕ	1.7 m/deg
4	θ	1.1 m/deg
5	ψ	0.8 m/deg
6	x_{UAV}	1.0 m/m
7	y_{UAV}	1.0 m/m
8	h_{UAV}	0.8 m/m
9	x_{ip}	0.15 m/pixel
10	y_{ip}	0.19 m/pixel

where i refers to each of the N parameters on which F is dependent, as listed in Table III. We therefore conclude that it is theoretically possible to locate a target within 15 m using computer vision from a fixed-wing UAV under nominal flight conditions of 60 m altitude and a large-radius coordinated turn.

IV. AVERAGING METHOD

Since each estimate of its location requires only one image of the target, it is theoretically possible to generate estimates at the frame rate of the camera. Although bandwidth constraints make this impossible, we can achieve several estimates per second, allowing for effective filtering to help reduce error. In this paper, we apply Recursive Least Squares (RLS) to filter the estimates.

Recursive Least Squares: Recursive Least Squares (RLS) is a simple method of recursively fitting a set of points to some function of choice by minimizing the sum of the squares of the offsets of the points. Typically, an RLS algorithm is used to fit a set of points to a characteristic line or quadratic, however, it can be also be used to find a characteristic point. In this case, the result of the RLS algorithm is identical to the result of a true average. The RLS algorithm implemented is detailed in Algorithm 1.

V. HARDWARE TESTBED

BYU has developed a reliable and robust platform for testing unmanned air vehicles [19], [20]. Figure 6 shows the key elements of the testbed. The first frame shows BYU's Kestrel autopilot which is equipped with a Rabbit 3400 29 MHz processor, rate gyros, accelerometers, absolute and differential pressure sensors. The autopilot measures $3.8 \times 5.1 \times 1.9$ cm and weighs 17 grams.

The second frame in Figure 6 shows the airframes used for the flight tests reported in this paper. The airframe is a 1.2 meter wingspan Zagi XS EPP foam flying wing, which was selected for its durability, ease of component installation, and flight characteristics. Embedded in the airframe are the Kestrel autopilot, batteries, a 1000 mW, 900 MHz radio modem, a GPS receiver, a video transmitter, and a small analog camera.

The third frame in Figure 6 shows the ground station components. A laptop runs the Virtual Cockpit software that interfaces through a communication box to the MAVs. An RC transmitter is used as a stand-by fail-safe mechanism to facilitate safe operations.

Algorithm 1 RLS Filter.

Input camera center location: $p^I_{\text{cc}} \leftarrow [x^I_{\text{cc}},\ y^I_{\text{cc}},\ z^I_{\text{cc}},\ 1]^T$
Input unscaled target location: $\bar{p}^I_{\text{obj}} \leftarrow [\bar{x}^I_{\text{obj}},\ \bar{y}^I_{\text{obj}},\ \bar{z}^I_{\text{obj}},\ 1]^T$
Input image depth estimate: λ

$\{$Pseudo-Code for $X\}$
Persistent P_N, A_N, b_N
$a_{N_1} \leftarrow I_{1\times 1}$
  $\{I_{1\times 1}$ refers to the (1×1) identity matrix$\}$
$b_{N_1} \leftarrow \left[\ x^I_{\text{cc}} + \lambda(\bar{x}^I_{\text{obj}} - x^I_{\text{cc}})\ \right]$
  $\{$The same equations apply for Y, only
    $b_{N_1} \leftarrow [y^I_{\text{cc}} + \lambda(\bar{y}^I_{\text{obj}} - y^I_{\text{cc}})]\}$

if isempty(A_N) **then**
  $A_N \leftarrow [a_{N_1}]$
  $b_N \leftarrow [b_{N_1}]$
  $P_N \leftarrow (A_N{}^T A_N)^{-1}$
  $X_{N_1} \leftarrow P_N A_N{}^T b_N$
else
  $P_{N_1} \leftarrow P_N - \frac{P_N a_{N_1}{}^T a_{N_1} P_N}{1 + a_{N_1} P_N a_{N_1}{}^T}$
  $A_{N_1} \leftarrow [A_N\ a_{N_1}]^T$
  $b_{N_1} \leftarrow [b_N\ b_{N_1}]^T$
  $X_{N_1} \leftarrow P_{N_1} A_{N_1}{}^T b_{N_1}$
end if

$P_N \leftarrow P_{N_1}$
$A_N \leftarrow A_{N_1}$
$b_N \leftarrow b_{N_1}$

return X_{N_1}

VI. EXPERIMENTAL RESULTS

The results of an initial hardware experiment are shown in Figure 7. The plot shows actual, estimated and filtered target locations from a UAV flying at 60 m altitude in a 50 m radius circular orbit around an initial guess of the target location. As can be seen in Figure 7, the filtered estimates of target location are nearly within the expected accuracy range of 15 m. At the start, the first few estimates are outside the predicted error radius. However, they quickly move into it as more estimates are made and the target's position is known with more confidence. The time history of the localization error is shown in Figure 8 demonstrating the rapid convergence of the RLS estimate to a quasi steady-state error of 10.9 m. In this case, RLS estimates converge within about 20 s, which is the time required to fly about one half of an orbit. Two main factors are believed to contribute to the steady-state error: attitude estimation errors and lack of synchronization between attitude and position telemetry and vision data. Future efforts will attempt to reduce these error sources to further improve localization capabilities.

Fig. 6. (a) Kestrel autopilot. (b) Zagi airframes. (c) Ground station components.

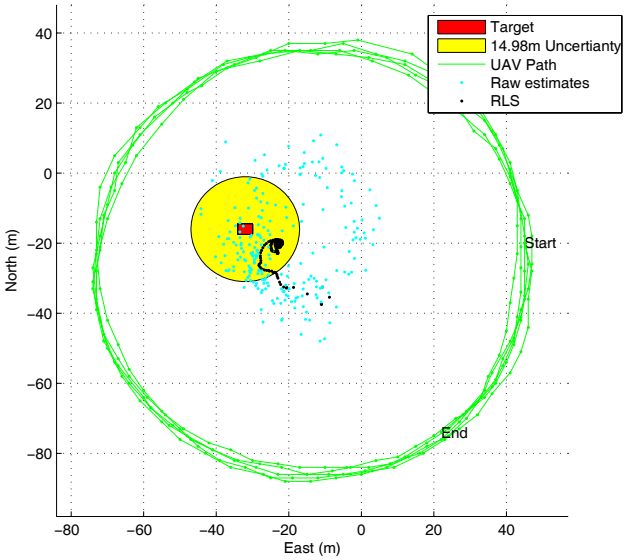

Fig. 7. Localization results

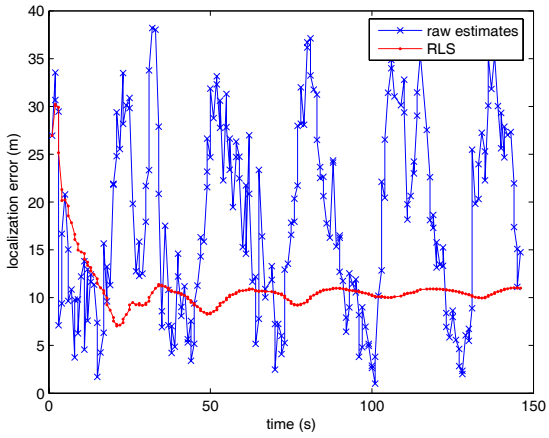

Fig. 8. Localization error

VII. Conclusions

This paper demonstrates the feasibility of vision-based target localization from a small, fixed-wing UAV. Results from hardware implementation show that the method produces satisfactory results, with excellent prospects for future improvement. Localization estimates could be improved by increasing the accuracy of attitude estimates, most notably the UAV roll angle.

Acknowledgments

This work was funded by AFOSR award number FA9550-04-1-0209 and the Utah State Centers of Excellence Program.

References

[1] Office of the Secretary of Defense, Ed., *Unmanned Aerial Vehicles Roadmap 2002-2027*. Washington DC, USA: United States Government, 2002.

[2] L. Chaimowicz, B. Grocholsky, J. F. Keller, V. Kumar, and C. J. Taylor, "Experiments in Multirobot Air-Ground Coordination," in *Proceedings of the 2004 International Conference on Robotics and Automation*, New Orleans, LA, April 2004, pp. 4053–4058.

[3] R. Vidal and S. Sastry, "Vision-Based Detection of Autonomous Vehicles for Pursuit-Evasion Games," in *IFAC World Congress on Automatic Control*, Barcelona, Spain, July 2002.

[4] R. Vidal, S. Rashid, C. Sharp, O. Shakernia, J. Kim, and S. Sastry, "Pursuit-Evasion Games with Unmanned Ground and Aerial Vehicles," in *Proceedings of the International Conference on Robotics and Automation*, Seoul, Korea, May 2001.

[5] A. Ollero and L. Merino, "Control and Perception Techniques for Aerial Robotics," in *7th IFAC Symposium on Robot Control (SYROCO 2003)*, Wroclaw, Poland, September 2003.

[6] P. Saeedi, D. G. Lowe, and P. D. Lawrence, "3D Localization and Tracking in Unknown Environments," in *Proceedings of the IEEE Conference on Robotics and Automation*, vol. 1, Sept 2003, pp. 1297–1303.

[7] S. G. Chroust and M. Vincze, "Fusion of Vision and Inertial Data for Motion and Structure Estimation," *Journal of Robotic Systems*, vol. 21, pp. 73–83, January 2003.

[8] H.-M. Gross, H.-J. Boehme, and T. Wilhelm, "A Contribution to Vision-based Localization, Tracking and Navigation Methods for an Interactive Mobile Service-robot," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, vol. 2, October 2001, pp. 672–677.

[9] T. Rofer and M. Jungel, "Vision-based Fast and Reactive Monte-Carlo Localization," in *Proceedings of the IEEE International Conference on Robotics and Automation*, September 2003.

[10] R. Rysdyk, "UAV Path Following for Constant Line-of-sight," in *2nd AIAA "Unmanned Unlimited" Systems, Technologies and Operations Aerospace, Land and Sea Conference*, September 2003.

[11] S. Stolle and R. Rysdyk, "Flight Path Following Guidance for Unmanned Air Vehicles with Pan-Tilt Camera for Target Observation," in *22nd Digital Avionics Systems Conference*, October 2003.

[12] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*. New Jersey, USA: Prentice-Hall, 2002.

[13] Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry, *An Invitation to 3-D Vision From Images to Geometric Models*. New York, USA: Springer-Verlag, 2003.

[14] D. B. Kingston, "Implementation Issues of Real-Time Trajectory Generation on Small UAVs," Master's thesis, Brigham Young University, Provo, Utah 84602, April 2004.

[15] D. B. Kingston and R. W. Beard, "Real-Time Attitude and Position Estimation for Small UAVs Using Low-Cost Sensors," in *AIAA 3rd Unmanned Unlimited Systems Conference and Workshop*, Chicago, Il, September 2004.

[16] Montana State University GPS Laboratory, "GPS Accuracy," 2004, http://www.montana.edu/ places/gps/lres357/slides/GPSaccuracy.ppt.

[17] Procerus Technologies, Inc., "Procerus UAV," 2005, http://www.procerusuav.com.

[18] R. S. Figliola and D. E. Beasley, *Theory and Design for Mechanical Measurements*. New York, USA: John Wiley and Sons, Inc., 2000.

[19] R. Beard, D. Kingston, M. Quigley, D. Snyder, R. Christiansen, W. Johnson, T. McLain, and M. Goodrich, "Autonomous vehicle technologies for small fixed wing UAVs," vol. 2, no. 1, pp. 92–108, January 2005.

[20] R. W. Beard, D. Lee, S. Thakoor, and S. Zornetzer, "A new approach to observation of descent and landing of future Mars mission using bioinspired technology innovations," vol. 2, no. 1, pp. 65–91, January 2005.