JOURNAL OF AEROSPACE COMPUTING, INFORMATION, AND COMMUNICATION
Vol. 2, September 2005

Vision-Aided Inertial Navigation for Flight Control

Allen D. Wu,* Eric N. Johnson,† and Alison A. Proctor‡
Georgia Institute of Technology, Atlanta, GA, 30332

Many onboard navigation systems use the Global Positioning System to bound the errors that result from integrating inertial sensors over time. Global Positioning System information, however, is not always accessible since it relies on external satellite signals. To this end, a vision sensor is explored as an alternative for inertial navigation in the context of an Extended Kalman Filter used in the closed-loop control of an unmanned aerial vehicle. The filter employs an onboard image processor that uses camera images to provide information about the size and position of a known target, thereby allowing the flight computer to derive the target's pose. Assuming that the position and orientation of the target are known a priori, vehicle position and attitude can be determined from the fusion of this information with inertial and heading measurements. Simulation and flight test results verify filter performance in the closed-loop control of an unmanned rotorcraft.

Nomenclature

$F_i = \{x_i, y_i, z_i\}$	inertial reference frame
$F_b = \{x_b, y_b, z_b\}$	body reference frame
$F_c = \{x_c, y_c, z_c\}$	camera reference frame
$\mathbf{q} = [q_1, q_2, q_3, q_4]^T$	attitude quaternions of F_b relative to F_i
$\mathbf{r} = [x, y, z]^T$	position vector of the vehicle measured from the datum
$\mathbf{R} = [X, Y, Z]^T$	relative position vector of the target's center from the vehicle
$\mathbf{t}$	position vector of the target's center measured from the datum
$\mathbf{n}$	vector normal to the plane of the target
A	area of target
f	camera focal length—relates pixels to angular displacement
FOV	camera field of view angle
α	pixel measurement of the square root of the target area in the camera image
u	pixel y-position of the target center in the camera image
v	pixel z-position of the target center in the camera image
$\mathbf{z} = [a, u, v]^T$	measurement vector
ψ_c	camera pan angle (rotation angle of F_c relative to F_b about the z_b-axis)
θ_c	camera tilt angle (rotation angle of F_c relative to F_b about the y_b-axis)
$\mathbf{L}_{j\kappa}$	direction cosine matrix converting from components in F_κ to components in F_j
$\mathbf{y}_\kappa$	vector $\mathbf{y}$ expressed as components in F_κ

Received 18 May 2005; revision received 29 July 2005; accepted for publication 9 May 2005.

* Graduate Research Assistant, Aerospace Engineering, Atlanta GA, Student Member AIAA, allen_wu@ae.gatech.edu.
† Lockheed Martin Assistant Professor of Avionics Integration, Aerospace Engineering, Atlanta GA, Member AIAA, Eric.Johnson@aerospace.gatech.edu.
‡ Graduate Research Assistant, Aerospace Engineering, Atlanta GA, Student Member AIAA, alison_proctor@ae.gatech.edu.

I. Introduction

UNMANNED aerial vehicles (UAVs) typically use a combination of the Global Positioning System (GPS) and inertial sensors for navigation and control. The GPS position data bounds the errors that accumulate from integrating linear acceleration and angular rate readings over time. However, this combination of sensors is not reliable enough for sustained autonomous flight in environments such as urban and indoor areas and at low altitudes where a GPS receiver antenna is prone to losing line-of-sight with satellites. Problems might also be encountered for vehicles operating in adversarial environments where access to the GPS network might possibly be jammed or otherwise denied.

Recent advances in the field of machine vision have shown great potential for vision sensors as an alternative to GPS in inertial navigation for obtaining aircraft position data. With a wealth of information available in each captured image, camera-based subsystems have a large margin for growth as onboard sensors. Furthermore, vision sensors allow for feature-based navigation through surrounding environments—a task impossible with GPS and inertial sensors alone. The research presented in this paper thus intends to apply recent advances in the field of image data processing to the classic problem of estimating the position and attitude of an aircraft by using a camera-based vision sensor for the role of positional correction in inertial navigation.

The notion of using machine vision as a secondary or primary means of aircraft navigation has seen a fair amount of progress over recent years. One of the first applications of vision-based sensors in avionics has been for assisting aircraft during landing.[1–4] Others have extended the capabilities of onboard cameras to include assisting UAVs in navigation during flight.[5–8] There has even been some success in controlling aircraft with vision-based sensors as the sole source of feedback.[9]

However, only recently have researchers begun seriously investigating the application of vision sensors in inertial navigation.[10–12] The problem of controlling position relative to a known feature point with only a camera and inertial sensors has been addressed.[13]

Researchers used closed-loop control to maneuver a robotic arm manipulator to a desired position relative to a point light source representing the feature point. Results from the experiment showed that an Extended Kalman Filter could not converge to reliable estimates with only camera pixel position information of the feature point. A modified filter was presented as an alternative solution.

The use of vision sensors in the inertial navigation of a ground robot has also been investigated.[14] Researchers have developed a system for a ground robot that uses the epipolar constraint to derive residuals based on measurements between two images and the vehicle motion as indicated by inertial sensors. Under the assumption of perfect orientation knowledge from flawless gyroscopes, the estimator corrects positional errors resulting from accelerometer drift.

The research presented in this paper builds upon previous work in the field of vision-based aircraft navigation by demonstrating the feasibility of vision-aided inertial navigation in the closed-loop control of an airborne system. This paper describes an Extended Kalman Filter (EKF) that uses information from camera images to correct the position and attitude estimates derived from inertial sensors for the navigation of an UAV. The experimental conditions of this research are first described, followed by a detailed description of the filter implementation, and finally simulation and flight-test results are presented in conclusion.

II. Experimental Setup

For this experiment, a rotorcraft UAV was selected as the test vehicle. The helicopter navigates towards a building using GPS, a magnetometer, and inertial sensors for navigation. Upon reaching the target building, the GPS is ignored. Vision-based position estimates, obtained from images of one of the building's windows, are then utilized in the EKF to bound the errors that result from integrating linear acceleration and angular rate readings over time.

A. Vehicle Description

The Georgia Institute of Technology GTMax helicopter, shown in Fig. 1, was selected as the test vehicle.[15] The GTMax is a modified Yamaha R-Max helicopter UAV that is equipped with dual flight computers, an inertial measurement unit (IMU), a differential GPS receiver, and a 3-axis magnetometer. An adaptive neural-network feedback controller uses EKF state estimates, obtained from the sensor array, to guide the helicopter appropriately. For this experiment, a commercially available CCD camera is included in addition to the normal flight hardware configuration, along with an image processor for interpreting visual images from the camera. The image processor

Fig. 1 The GTMax UAV helicopter.

is capable of detecting a dark square against a light background and provides the square root of the object's area and the position of the object's center in the camera image.[9]

B. Assumptions

It is assumed that the target is a stationary dark 2D square against a light background to simplify the requirements of computationally expensive image processing. The target's position, orientation, and area are also known beforehand. Regarding camera geometry, the assumptions are made that the distance from the camera to the vehicle center of mass is negligible and that the horizontal and vertical focal lengths are the same.

III. Position Estimation from Camera Images

The EKF estimates vehicle position using three measurements of a target obtained from camera images: the square root of the target's area (α) and the pixel y and z positions of the target center in the camera image (u and v respectively). Since the position and orientation of the target are known, the vision-based estimator needs only to determine the pose of the target to resolve the helicopter's position. This section first provides some fundamentals on relating 3D position to 2D images, and then describes the implementation of the EKF.

A. Relating 3D Position to 2D Images

A perspective projection model of a pinhole camera allows position in a 2D camera image to be inferred from 3D position as shown in Fig. 2.[9,16] The model projects an arbitrary point ($A,\ B,\ C$) to a pixel point ($b,\ c$) on the image plane (the camera image) according to the following relations:

$$b = f\frac{B}{A} \tag{1}$$

$$c = f\frac{C}{A}, \tag{2}$$

where f is the focal length of the camera. The focal length depends solely on the image width in pixels of the camera image (w) and the angle of the field of view (FOV), both of which are characteristics of the camera, according to

$$f = \frac{w}{2\tan\left(\frac{FOV}{2}\right)}. \tag{3}$$

Note that typically the scale factors for the horizontal and vertical axes are not identical, but they are assumed to be within acceptable tolerances for this experiment.

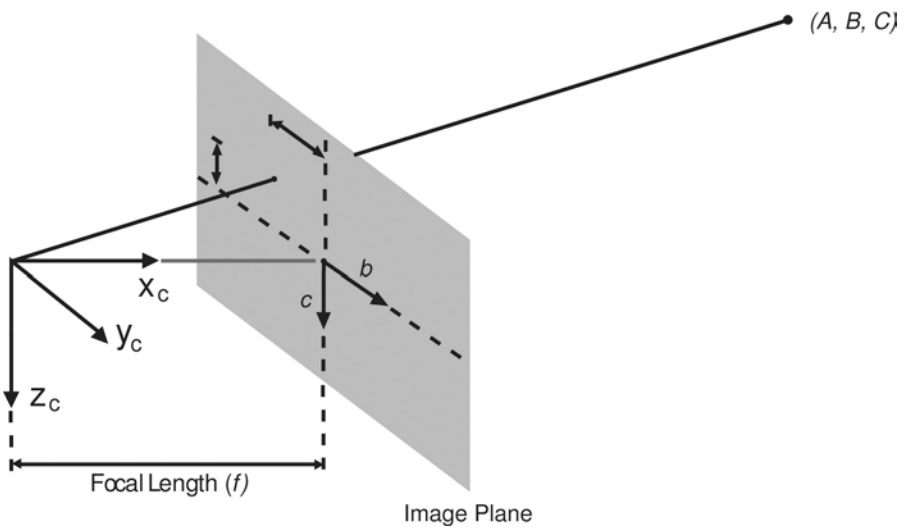

Fig. 2 Camera perspective projection model used for relating 3D position to position in 2D images. The point (A, B, C) is projected onto the camera image plane to the point (b, c).

B. Reference Frames

Three frames of reference are used in this paper. The inertial reference frame is a local inertial frame with its origin centered at the position of the differential GPS tower on the ground. The camera frame has its origin at the camera's principal point with the x_c axis along the camera's optical axis, whereas the body frame is the conventional aircraft body frame centered at the vehicle center of mass. Vector components in the different reference frames can be transformed using direction cosine matrix sequences as follows:

$$\mathbf{L}_{cb} = \begin{bmatrix} \cos\theta_c & 0 & -\sin\theta_c \\ 0 & 1 & 0 \\ \sin\theta_c & 0 & \cos\theta_c \end{bmatrix} \begin{bmatrix} \cos\psi_c & \sin\psi_c & 0 \\ -\sin\psi_c & \cos\psi_c & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

$$\mathbf{L}_{bi} = \begin{bmatrix} q_1^2 + q_2^2 - q_3^2 - q_4^2 & 2(q_2q_3 + q_1q_4) & 2(q_2q_4 - q_1q_3) \\ 2(q_2q_3 - q_1q_4) & q_1^2 - q_2^2 + q_3^2 - q_4^2 & 2(q_3q_4 + q_1q_2) \\ 2(q_2q_4 + q_1q_3) & 2(q_3q_4 - q_1q_2) & q_1^2 - q_2^2 - q_3^2 + q_4^2 \end{bmatrix} \tag{5}$$

$$\mathbf{L}_{ci} = \mathbf{L}_{cb}\mathbf{L}_{bi}. \tag{6}$$

$\mathbf{L}_{cb}$ is a rotation sequence that converts vectors from components in the body frame to components in the camera frame by using the pan (ψ_c) and tilt (θ_c) angles of the camera. Note, however, that the transformation from the body to the camera frame accounts for only the orientation differences between the two frames. The fact that the camera frame is centered at the camera location, whereas the body frame is centered at the vehicle center of mass, is neglected. $\mathbf{L}_{bi}$ is a standard rotation matrix from the body to the local inertial frame expressed in quaternions.

C. Extended Kalman Filter (EKF) Formulation

The vision-based EKF is a mixed continuous-discrete time filter. The process model uses a continuous-time linearized model of the helicopter dynamics to predict expected values for the state variables, whereas the measurement model runs at discrete time steps to correct the process model's estimates by using camera measurements.[17] The EKF measurement vector consists of the horizontal and vertical Cartesian coordinates (in pixels) of the target center in the camera image, and the square root of the target area in pixels. By comparing predicted values of the measurement

vector with actual measurements from the image processor, the EKF is able to maintain estimates of the helicopter's position and attitude.

1. Process Model

The process model for the EKF is a dynamic helicopter model that utilizes 16 states: attitude quaternions, position and velocity as components in the inertial reference frame, and accelerometer and gyroscope biases. In the process model phase of the EKF estimation algorithm, two main events occur. First, the state estimate $\hat{\mathbf{x}}$ is updated using the derivative of the state vector obtained directly from inertial measurements. Simultaneously, the covariance matrix $\mathbf{P}$ is updated according to the Differential Lyapunov Equation. The process model can be expressed as the following set of differential equations

$$\dot{\hat{\mathbf{x}}}(t) = f(\hat{\mathbf{x}}(t), t) \tag{7}$$

$$\dot{\mathbf{P}}(t) = \mathbf{AP} + \mathbf{PA}^T + \mathbf{Q} \tag{8}$$

where $\hat{\mathbf{x}}$ is the state estimate, $f(\hat{\mathbf{x}}(t), t)$ is the nonlinear helicopter model, $\mathbf{P}$ is the error covariance matrix, $\mathbf{A}$ is a matrix representing the linearized helicopter dynamics, and $\mathbf{Q}$ is a diagonal matrix representing the process noise inherent in the system model. Values for the $\mathbf{Q}$ matrix were initially set by using rough approximations derived from modeling assumptions, and then tuning the values based on data from subsequent flight tests.

2. Measurement Model

The helicopter generates expected values of the measurement vector by using the known position, attitude, and size of the window in conjunction with the helicopter state predicted by the dynamic process model. As shown in Fig. 3, the vector from the helicopter to the target window in the camera frame can be found by taking the difference between the known position vector to the target and the helicopter's predicted position expressed as components in the inertial frame. The result can then be converted from the inertial reference frame to the camera frame, giving

$$\mathbf{R}_c = \mathbf{L}_{ci}(\mathbf{t}_i - \mathbf{r}_i). \tag{9}$$

Previously described image plane relations provide the basis for the model of the measurement vector, $\mathbf{z} = [\alpha,\ u,\ v]^T$. Camera u and v measurements follow directly from (1) and (2). To find the square root of the target area

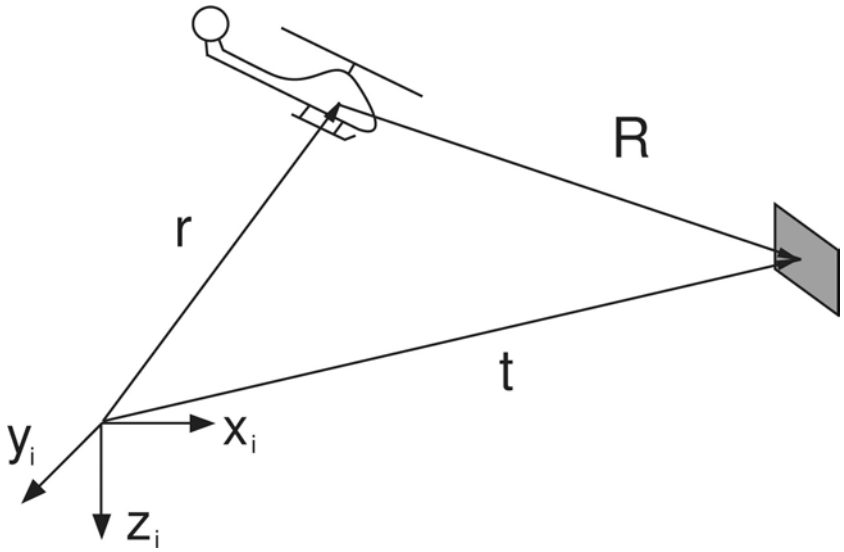

Fig. 3 Vectors used in describing the EKF measurement model.

in pixels, α, the actual window area is projected onto the image plane and then scaled to pixels according to

$$u = f\frac{Y}{X} \tag{10}$$

$$v = f\frac{Z}{X} \tag{11}$$

$$\alpha = \frac{f}{X}\sqrt{A\frac{\mathbf{n}\cdot\mathbf{R}}{\|\mathbf{n}\|\cdot\|\mathbf{R}\|}}. \tag{12}$$

$\mathbf{R}_c = [X,\ Y,\ Z]^T$ is the relative position vector of the target window from the helicopter expressed as components of the camera frame. The vector $\mathbf{n}$ is the normal vector of the target window derived from the window's known attitude. The given area of the window is represented by A and f is the camera focal length as described before.

The EKF makes use of the measurement model to update the result from the integration of the process model given in (7) and (8), at a rate of 10 Hz for both the camera and magnetometer readings. In this update phase of the state estimation, the Kalman gain is first computed. This gain is used as a weighting to fuse the actual camera measurements with the predicted measurements, given by the measurement model in (10)–(12), thus correcting the a priori process model estimate according to the following equations

$$\mathbf{K} = \mathbf{P}^-\mathbf{C}^T(\mathbf{C}\mathbf{P}^-\mathbf{C}^T + \mathbf{V})^{-1} \tag{13}$$

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}^- + \mathbf{K}[\mathbf{z} - h(\hat{\mathbf{x}}^-)] \tag{14}$$

$$\mathbf{P} = (\mathbf{I} - \mathbf{K}\mathbf{C})\mathbf{P}^- \tag{15}$$

where $\mathbf{K}$ is the Kalman gain, $\mathbf{V}$ is a diagonal matrix representing measurement noise in the camera sensor, $\mathbf{C}$ is the Jacobian of the measurement vector with respect to the state vector, also denoted by $\left(\frac{\partial \mathbf{z}}{\partial \mathbf{x}}\right)$, and $h(\hat{\mathbf{x}}^-)$ is the predicted measurement vector as given by the measurement model in (10)–(12). Minus superscripts in the above equations denote a priori values obtained from the process model equations. The results from (13)–(15) are used by the process model in the next time step to further propagate the state vector and the covariance matrix, and the procedure is repeated. Constant values for the measurement error covariance matrix $\mathbf{V}$ were roughly estimated based on the physical properties of the camera such as the focal length and dimensions of the CCD array. These parameters could be tuned using data from subsequent flight tests, but the inaccuracies from using an initial approximation were deemed acceptable for the purpose of demonstrating the filter's feasibility.

As can be seen from (13)–(15), the EKF requires the Jacobian matrix of the measurement vector with respect to the vehicle states of the helicopter, $\mathbf{C}$, to fuse the camera information with inertial data. Measurements from the camera are only affected by vehicle position and attitude, hence the partial derivatives with respect to the other state variables (vehicle velocity and accelerometer and gyroscope biases) are identically zero. The linearization of the measurement model with respect to vehicle position in the local inertial reference frame is obtained from (9)–(12) as

$$\frac{\partial \mathbf{z}}{\partial \mathbf{r}_i} = \left(\frac{\partial \mathbf{z}}{\partial \mathbf{R}_c}\right)\left(\frac{\partial \mathbf{R}_c}{\partial \mathbf{r}_i}\right) \tag{16}$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{R}_c} = \frac{f}{X^2}\begin{bmatrix}\sqrt{A} & 0 & 0\\ -Y & X & 0\\ -Z & 0 & X\end{bmatrix} \tag{17}$$

$$\frac{\partial \mathbf{R}_c}{\partial \mathbf{r}_i} = -\mathbf{L}_{ci}. \tag{18}$$

The linearized model of the measurement vector with respect to body attitude quaternions is computed in a similar fashion, giving

$$\frac{\partial \mathbf{z}}{\partial \mathbf{q}} = \left(\frac{\partial \mathbf{z}}{\partial \mathbf{R}_c}\right)\mathbf{L}_{cb}\left(\frac{\partial \mathbf{R}_b}{\partial \mathbf{q}}\right) \tag{19}$$

where $\mathbf{R}_b$ is the relative position vector of the target's center from the vehicle, expressed as components of the body frame.

It should be noted, however, that there exists a nontrivial latency in the visual information path due to the computational intensity of the image processing. The time that it takes to calculate the measurement vector from visual information, provided by the camera's framegrabber, is not negligible and needs to be taken into account. To mitigate this effect, a simple latency compensation scheme is employed. In the update phase of the EKF, the onboard computer uses state information that has been stored in memory for the latency period of time as opposed to using the most up to date state estimate. This prevents the EKF from comparing the current state estimate information with a camera snapshot that was actually taken some time ago. The system latency was estimated empirically from flight test data by finding the latency value that minimized the least squares error between the predicted camera measurements, found by applying the measurement model to recorded state data obtained using GPS, and the actual recorded camera measurements from the flight.

IV. Simulation Results

Simulation tests were used initially to verify the vision-based filter's performance. In the particular scenario presented in this section, the helicopter begins at a position 120 ft away from a 12.6 ft^2 window and uses only an IMU, a magnetometer, and a camera as sensors for navigation. The helicopter hovers and then performs a 7.5 ft step up in altitude, a 10 ft step parallel to the window, and a 10 ft step towards the window. The state estimates obtained from the vision-based EKF are used in the closed-loop control of the vehicle in this test.

A. Simulation Environment

Performance of the vision-based EKF was validated using the simulation environment shown in Fig. 4.[15] The simulation includes a nonlinear helicopter model in addition to sensor models for the IMU, the magnetometer, the vision sensor, and GPS. Synthetic images of the rendered environment are used as image processor inputs. The aircraft model has six rigid-body degrees of freedom and includes dynamics for the engine, fuel, landing gear, and rotor.

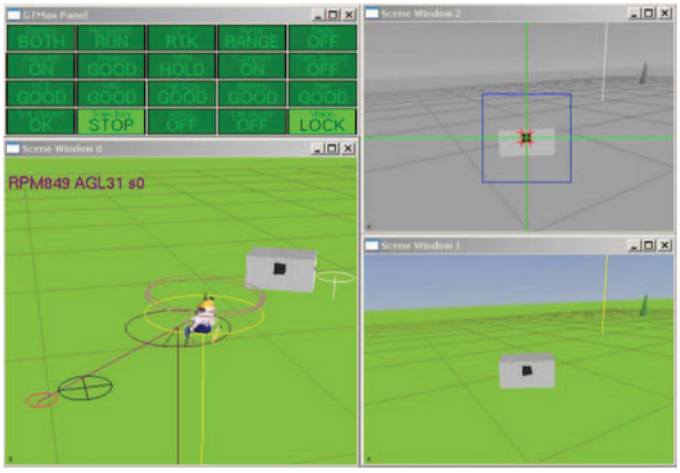

Fig. 4 An image from the simulation environment used for testing the EKF. The top right simulation window represents information from the onboard image processor. Two helicopters are actually shown in the simulation to represent the actual position of the helicopter and the estimated position as determined by the EKF. The yellow circle indicates the trajectory commanded to the helicopter's controller to maneuver towards the position commanded by the user, which is specified by the location of the purple circle. In the simulations that were performed, the helicopter representing the estimated position tended to maintain a closer proximity to the yellow trajectory ring than the helicopter representing the actual position.

Sensor models include errors, mounting location and orientation, and time delays in the transfer of data. Modeling errors and environmental disturbances, worse than those expected to be experienced during flight, are also injected into the simulation.

B. Simulation Data

Figure 6 compares the position estimates from the vision-based EKF with positional information obtained from GPS. Figure 7 shows the positional errors of the filter as calculated by the difference between the estimated and true positions given by GPS readings. These plots express all positions as components of the relative coordinate system defined in Fig. 5. The fact that the GPS receiver antenna is not located at the vehicle center of mass has been accounted for by including a bias in the raw GPS position data.

It can be seen from the data that the vision-based EKF is capable of maintaining a state estimate within a neighborhood of the actual position as represented by the GPS position measurements. This bound on the error in the presence of the commanded step inputs demonstrates the convergence of the filter. The maximum difference between the estimated and actual positions are 9.2 ft, 4.7 ft, and 1.9 ft in the x_w, y_w, and z_w directions respectively. Simulation tests revealed that the EKF is sensitive to variations in magnetometer readings because the selected measurement vector of $[\alpha,\ u,\ v]^T$, in conjunction with inertial measurements alone, is not sufficient to uniquely determine the position of the aircraft relative to the target. This occurs because the camera can obtain identical $[\alpha,\ u,\ v]^T$ readings for a multitude of positions and attitudes. Simulations revealed that the introduction of a bias into the raw magnetometer readings resulted in the helicopter settling into a new steady-state position offset from the desired position, and increasing the standard deviation of noise in the magnetometer caused the vehicle to oscillate along an arc around the target. Heading information from the magnetometer is required to fix the azimuth of the helicopter along this arc.

V. Flight Test Results

Figures 8 and 9 show results from a flight test in which the vision-based EKF is used in the closed-loop control of the GTMax helicopter. These plots compare estimates from the EKF with positional information obtained from GPS. As in the simulation results of the previous section, a bias has been included in the raw GPS position data to account for the fact that the GPS receiver antenna is not located at the vehicle center of mass. Flight tests were conducted using a mock square window setup with an area of 36 ft^2.

The flight test results demonstrate the capability of the EKF to converge and maintain a position estimate accurate to within a neighborhood of the actual position as represented by the GPS measurements, verifying the simulation results. The maximum difference between the estimated and actual positions are 9.6 ft, 8.3 ft, and 4.0 ft in the x_w, y_w, and z_w directions respectively. Biases in the x_w and y_w directions could potentially be caused by biases in the magnetometer readings as described in the discussion of the simulation results. Differences occur between the simulated and actual results because the simulation models the ideal situation where distortion and noise effects of the camera are neglected and the window is a perfect dark target against a white background. The EKF is less sensitive to noise in the camera images when the helicopter is closer to the target.

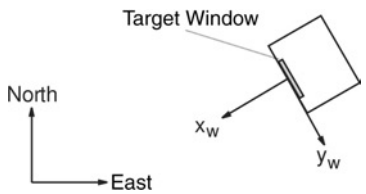

Fig. 5 The relative coordinate system used for plotting the simulation and flight test results.

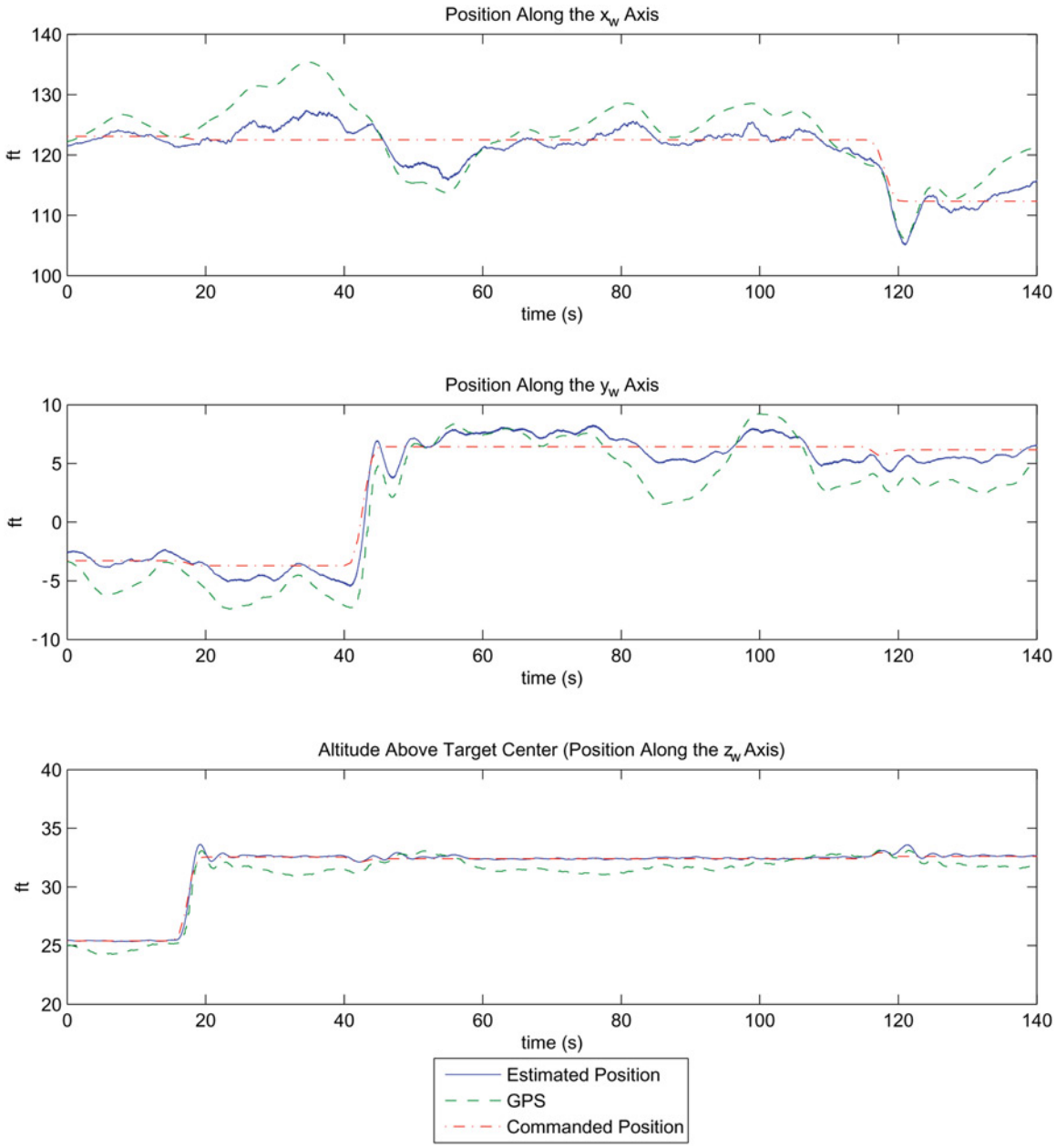

Fig. 6 Results from a *simulation* test comparing position estimates from the vision-based EKF with positional information obtained from GPS. Positions are expressed in coordinates relative to the center of the the target window (see Fig. 5). The helicopter starts at a location 120 ft from the target and is given a step input of 7.5 ft up in altitude, a 10 ft step parallel to the plane of the window, and finally a 10 ft step towards the window. The target's center is located at $[54, 69, -5]$ ft with a normal vector of $[-0.86, -0.50, 0.00]$ expressed in North, East, Down coordinates.

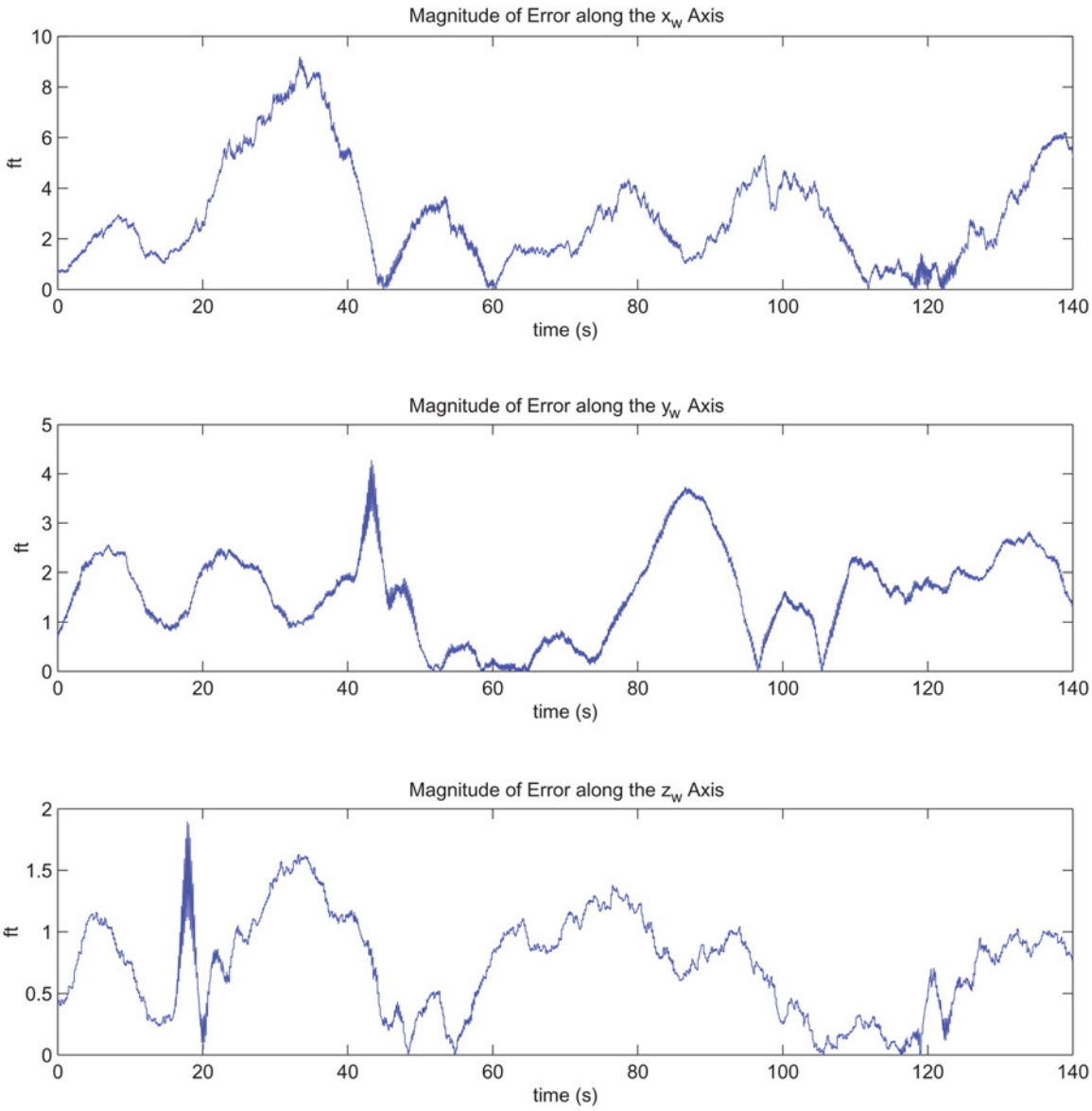

Fig. 7 Error plots from the *simulation* results. Error was calculated as the difference between the estimated position and GPS (a worst case representation of actual position). The maximum error magnitudes are 9.2 ft, 4.7 ft, and 1.9 ft in the x_w, y_w, and z_w directions respectively.

VI. Conclusion

This paper presented an EKF approach for the vision-aided inertial navigation of an UAV using only an IMU, camera, and magnetometer as sensors for navigation. Using camera images of a target whose position and attitude are given, the vehicle can determine position relative to the target, and therefore derive position in the world. Results from simulation and flight tests verified that utilizing a vision-based sensor in this manner for aiding inertial navigation successfully bounds the errors that result from integrating the inertial sensors over time.

With further development, vision sensors could become a viable alternative to GPS in inertial navigation for applications where the risks associated with a dependency on external signals is not acceptable. Firstly, the bound

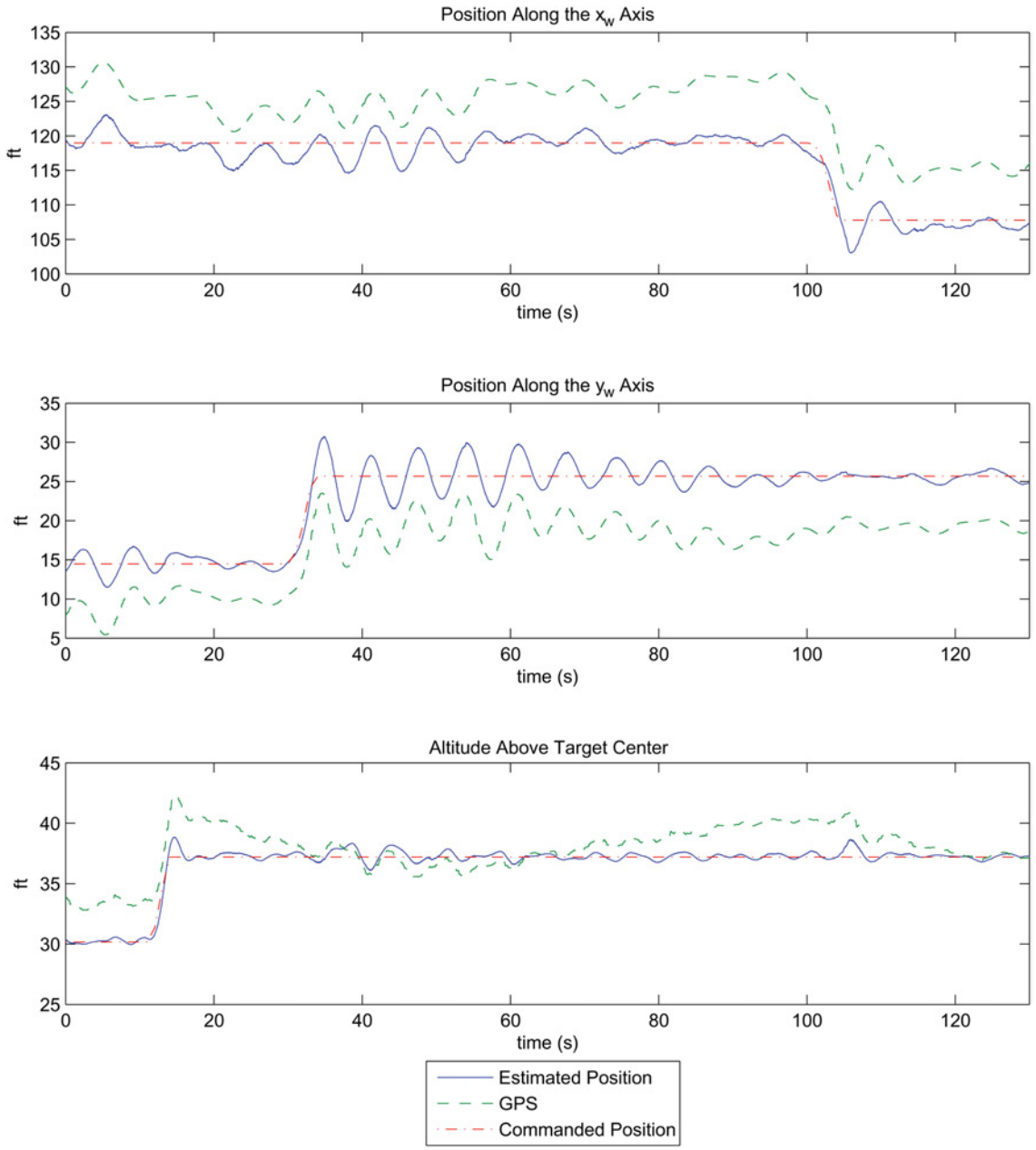

Fig. 8 Results from a *flight* test of the GTMax helicopter comparing position estimates from the vision-based EKF with positional information obtained from GPS. Positions are expressed in coordinates relative to the center of the the target window (see Fig. 5). The helicopter performs a 7.5 ft step up in altitude, a 10 ft step parallel to the window, and finally a 10 ft step towards the window. The target's center is located at [65.5, −100.5, 0.0] ft with a normal vector of [−0.64, 0.77, 0.00] expressed in North, East, Down coordinates.

on the position estimate from the filter could be tightened. This could be achieved by using more advanced cameras intended for metrology applications, as opposed to the commercial-grade camera used for this experiment, or by possibly expanding the capabilities of the image processor to include the ability to track multiple targets. Having the capability to track multiple targets simultaneously would also decrease the UAV's reliance on the magnetometer for

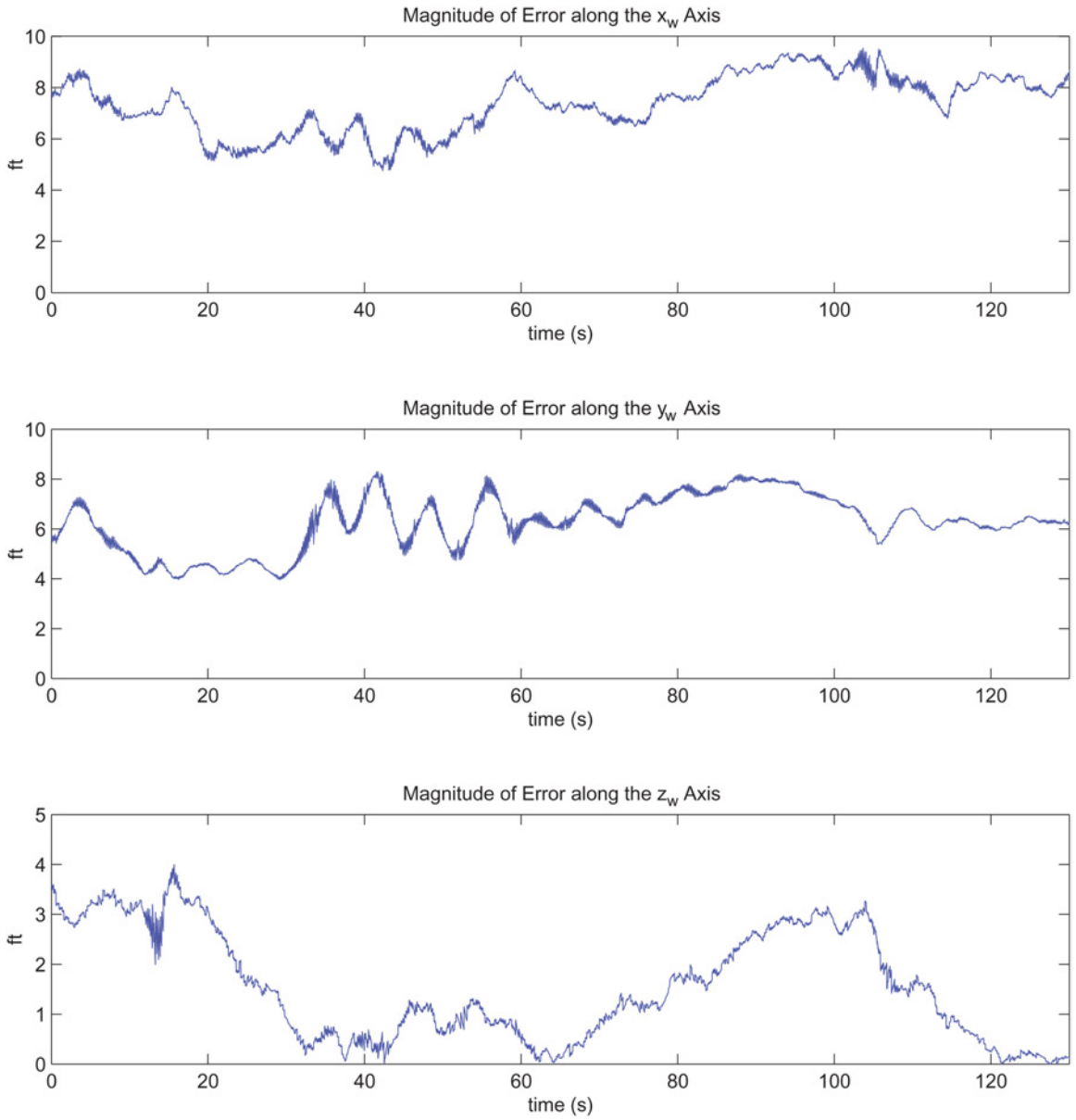

Fig. 9 Error plots from the *flight* test results. Error was calculated as the difference between the estimated position and GPS. The maximum error magnitudes are 9.6 ft, 8.3 ft, and 4.0 ft in the x_w, y_w, and z_w directions respectively.

heading information. In the long term, it is hoped that UAVs will be able to actively track and acquire new visual targets in real-time as they proceed along their flight paths so that vision-aided inertial navigation can be used for a variety of missions.

Acknowledgments

The authors would like to acknowledge contributions of the following people to the work presented in this paper: Henrik Christophersen, Jincheol Ha, Jeong Hur, Suresh Kannan, Adrian Koller, Alex Moodie, Wayne Pickell, and Nimrod Rooz. This work is supported by the Active-Vision Control Systems (AVCS) Multi-University Research

Initiative (MURI) Program under contract #F49620-03-1-0401 and by the Software Enabled Control (SEC) Program under contracts #33615-98-C-1341 and #33615-99-C-1500.

References

[1]Saripalli, S., Montogomery, J. F., and Sukhatme, G. S., "Visually Guided Landing of an Unmanned Aerial Vehicle," *Transactions on Robotics and Automation*, IEEE, Vol. 19, No. 3, June 2003, pp. 371–380.

[2]Shakernia, O., Vidal, R., Sharp, C. S., Ma, Y., and Sastry, S., "Multiple View Motion Estimation and Control for Landing an Unmanned Aerial Vehicle," *Proceedings of the International Conference on Robotics and Automation*, IEEE, Vol. 3, Washington, DC, May 2002, pp. 2793–2798.

[3]Chatterji, G. B., Menon, P. K., and Sridhar, B., "GPS/Machine Vision Navigation System for Aircraft," *Transactions on Aerospace and Electronic Systems*, IEEE, Vol. 33, No. 3, July 1997, pp. 1012–1025.

[4]Bosse, M., Karl, W. C., Castanon, D., and DeBitetto, P., "A Vision Augmented Navigation System," *Conference on Intelligent Transportation Systems*, IEEE, Nov. 1997, pp. 1028–1033.

[5]Amidi, O., "An Autonomous Vision-Guided Helicopter," Ph.D. Dissertation, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, 1996.

[6]Bernatz, A., and Thielecke, F., "Navigation of a Low Flying VTOL Aircraft With the Help of a Downwards Pointing Camera," *Guidance, Navigation, and Control Conference*, AIAA, Providence, RI, Aug. 2004.

[7]Webb, T. P., Prazenica, R. J. Kurdila, A. J., and Lind, R., "Vision-Based State Estimation for Autonomous Micro Air Vehicles," *Guidance, Navigation, and Control Conference*, AIAA Providence, RI, Aug. 2004.

[8]Roberts, J., Corke, P., and Buskey, G., "Low-Cost Flight Control System for a Small Autonomous Helicopter," Australasian Conference on Robotics and Automation, Auckland, New Zealand, 2002.

[9]Proctor, A. A., and Johnson E. N., "Vision-Only Aircraft Flight Control Methods and Test Results," *AIAA Guidance, Navigation, and Control Conference*, Providence, RI, Aug. 2004.

[10]Roberts, B. A., and Vallot, L. C., "Vision Aided Inertial Navigation," *Record of the Position Location and Navigation Symposium*, IEEE, March 1990, pp. 347–352.

[11]Rehbinder, H., and Ghosh, B. K., "Multi-Rate Fusion of Visual and Inertial Data," *International Conference on Multisensor Fusion and Integration for Intelligent Systems*, IEEE, Aug. 2001, pp. 97–102.

[12]Kaminer, I., Pascoal, A., Kang, W., "Integrated Vision/Inertial Navigation System Design Using Nonlinear Filtering," *Proceedings of the American Control Conference*, Vol. 3, San Diego, CA, 1999, pp. 1910–1914.

[13]Huster, A., "Relative Position Sensing by Fusing Monocular Vision and Inertial Rate Sensors," Ph.D. Dissertation, Department of Electrical Engineering, Stanford University, Stanford, CA, 2004.

[14]Diel, D., DeBitetto, P., and Teller, S., "Epipolar Constraints for Vision-Aided Inertial Navigation", *Workshop on Motion and Video Computing*, IEEE, Breckenridge, CO, 2005.

[15]Johnson, E. N., and Schrage, D. P., "System Integration and Operation of a Research Unmanned Aerial Vehicle," *Journal of Aerospace Computing, Information, and Communication*, Vol. 1, No. 1, 2004, pp. 5–18.

[16]Russell, S. J., and Norvig, P., *Artificial Intelligence: A Modern Approach*, Prentice Hall, New Jersey, 1995, pp. 725–742.

[17]Gelb, A. (ed.), *Applied Optimal Estimation*, The MIT Press, Cambridge, MA, 1974, pp. 102–110, 182–190.