

Vision-based Target Tracking with Adaptive Target State Estimator

Ramachandra J. Sattigeri¹ and Eric Johnson²
Georgia Institute of Technology, Atlanta, GA 30332-0150

Anthony J. Calise³
Georgia Institute of Technology, Atlanta, GA 30332-0150

and

Jincheol Ha⁴
Georgia Institute of Technology, Atlanta, GA 30332-0150

This paper presents an approach to vision-based target tracking with a neural network (NN) augmented Kalman filter as the adaptive target state estimator. The vision sensor onboard the follower (tracker) aircraft is a single camera. Real-time image processing implemented in the onboard flight computer is used to derive measurements of relative bearing (azimuth and elevation angles) and the maximum angle subtended by the target aircraft on the image plane. These measurements are used to update the NN augmented Kalman filter. This filter generates estimates of the target aircraft position, velocity and acceleration in inertial 3D space that are used in the guidance and flight control law to guide the follower aircraft relative to the target aircraft. Applications of the presented approach include vision-based autonomous formation flight, pursuit and autonomous aerial refueling. The NN augmenting the Kalman filter estimates the target acceleration and hence provides for robust state estimation in the presence of unmodeled target maneuvers. Vision-in-the-loop simulation results obtained in a 6DOF real-time simulation of vision-based autonomous formation flight are presented to illustrate the efficacy of the adaptive target state estimator design.

I. Introduction

As demonstrated in recent conflicts, Unmanned Aerial Vehicles (UAVs) are becoming an important component of our military force structure. UAVs, operating in close proximity to enemy forces, provide real-time information difficult to obtain from other sources, without risk to human pilots. Most modern UAVs contain multi-sensor based navigation systems, typically consisting of Global Positioning System (GPS) and Inertial Measurement Units (IMUs). These navigation systems can be complex and expensive, vulnerable to jamming and not suitable for use in small, disposable UAVs. Therefore, there is a need to develop reliable and affordable navigation alternatives¹. Vision sensors offer a cheap, light-weight and reliable alternative to GPS. Existing applications include a method for vision-aided inertial navigation, where the output of a vision sensor and an IMU are used in tandem to allow autonomous navigation of a rotorcraft UAV without the use of GPS updates². In another application, the GPS is replaced by a vision sensor (fixed camera) onboard a follower UAV to derive estimates of states relative to a target or leader UAV for the purpose of enabling autonomous formation flight³. A closely related application is vision-based autonomous aerial refueling⁴.

The problem of tracking maneuvering targets has been extensively researched over the past few decades. The

¹ Graduate Research Assistant, School of Aerospace Engineering, ram_sattigeri@gatech.edu, AIAA Member.

² Lockheed Martin Assistant Professor of Avionics Integration, School of Aerospace Engineering, Eric.Johnson@ae.gatech.edu, AIAA Member.

³ Professor, School of Aerospace Engineering, anthony.calise@ae.gatech.edu, Fellow AIAA.

⁴ Graduate Research Assistant, School of Aerospace Engineering, gtg203c@mail.gatech.edu.

primary objective of target tracking is to estimate the state trajectories of a moving object. One of the major challenges for target tracking arises from target motion uncertainty. This uncertainty refers to the fact that an accurate dynamic model of the target being tracked is generally not available to the tracker. In addition, any measurements of the target being tracked are corrupted by noise and time delays. A Kalman filter or one of its several variants is usually used as the target state estimator, but its performance may be seriously degraded unless the estimation error due to unknown target maneuvers and other uncertainties is compensated. Some of the most popular approaches to handling the unknown target maneuvers in the target state estimator design are based on the so called model-based filtering techniques⁵. The models may: 1) approximate the actually nonrandom target maneuver as a random process of certain properties, or 2) describe typical target trajectories by some representative motion models. Examples of the former include the simple white-noise acceleration model⁶, the slightly more sophisticated Markov process based Singer model⁷, and the more complex interacting multiple model technique⁸. When there is some *a priori* knowledge of the target maneuver, for example, if it is known that the target is maneuvering in a circle, circular motion models^{3,9} can be used for designing the target state estimator. However, in general for the model-based approaches to target state estimation, filter performance may not be satisfactory when the target maneuver does not comply with the model, and every approach can be defeated with a suitably chosen target maneuver. Neural Network (NN) based adaptive estimation and filtering techniques for state estimation design have been proposed to compensate for the modeling errors that arise due to nonlinearities and unmodeled dynamics^{10,11}. In these approaches, a nominal time-varying estimator is augmented with the output of an adaptive NN that is trained online with the residuals of the nominal estimator and with delayed values of available system measurements as inputs. Ref. [11] includes an approach to improve the effectiveness of the adaptation by deriving an additional error signal to train the NN.

The purpose of this paper is to present an approach to vision-based target tracking with a NN-based adaptive target state estimator. The approach will be demonstrated in an autonomous formation flight experiment with two UAVs in a leader-follower configuration. The vision sensor onboard the follower aircraft is a monocular fixed camera, which implies that the range to the leader aircraft is not available as a measurement and has to be estimated by a target state estimator, for example, an Extended Kalman Filter (EKF). Two approaches for estimating the range in an identical setup were presented in Ref. [3] and the second approach was successfully flight tested. In the flight test results reported, the target was maneuvering in the horizontal plane at a constant but unknown heading rate. The EKF in Ref. [3] incorporated a target maneuver model which was close to the actual target maneuver. Specifically, the EKF contained a state for the target centripetal acceleration which was assumed to be constant. When the target state estimator does not include an accurate target maneuver model, the state estimates are biased. These biased estimates can cause the follower aircraft to maneuver in a manner such that the target aircraft goes out of the field-of-view of the camera. In certain cases the estimates can diverge leading to dangerous divergence in the follower trajectory.

The contribution of this paper is in implementing the NN-based adaptive state estimator presented in Ref. [11] in the vision-based target tracking problem. The nominal target state estimator in this implementation is a linear, time-varying Kalman filter that models the target acceleration components in the North-East-Down (NED) coordinate frame as independent, zero-mean white noise processes. The NN estimates the unmodeled target acceleration and compensates the nominal filter. Vision-in-the-closed-loop simulation results obtained by integrating the adaptive estimator in the 6DOF Georgia Tech Unmanned Systems Testbed (GUST) real-time simulation software¹² are presented.

The organization of this paper is as follows. Section II briefly reviews the vision-based formation flight problem and the image processing used in this work. Section III reviews the formulation of the adaptive state estimator for the vision-based target tracking problem. Section IV presents the simulation results with some discussion. Section V presents the conclusions and states future research efforts.

II. Vision-based Formation Flight

The follower aircraft is tasked with the objective of maintaining a desired relative position from the leader aircraft. This problem requires simultaneous sensor data processing, state estimation and tracking control in the presence of unmodeled disturbances (leader acceleration) and measurement uncertainties. Figure 1 shows the block diagram of the follower aircraft with a vision-in-the-loop flight control system.

A. Image Processing

Although tracking in the presence of disturbances is a classical control issue, the problem at hand is very difficult and challenging due to the highly uncertain nature of the disturbance. Sensor data processing involves fast

converging computer vision algorithms that track the leader aircraft in the presence of background clutter and derive noisy measurements of the leader aircraft's position relative to itself^{4,13-15}. In this respect, the problem of vision-in-the-loop tracking differs from standard tracking problems in that the feedback signal is measured using vision sensors. The feedback signal is extracted by the computer vision algorithms and interpreted by a reasoning algorithm before being used by the control loop. The uncertainties arise from sensor noise and the assumptions embedded in the computer vision and reasoning algorithms, for example, the likelihood of various hypotheses.

The image processing algorithms employed in this paper are based on geometric active contours. Active contours or snakes are used to track various features of interest over time across several image frames as shown in Figure 2. Active contours were introduced in Ref. [16] and are deformable contours with the capability of conforming to objects in the image plane, making them ideal for segmentation, edge detection, shape modeling and visual tracking. Active contours fit naturally into control frameworks and have often been employed in conjunction with Kalman filtering¹⁴. Geometric active contours are implemented via level set methods which were based on implicit representations of the evolving contour¹⁷. The specific active contour model used in this application is based on a fast implementation of the Chan-Vese model¹⁸. The computational speed of these methods is over 30 frames/sec which enhances viability in real-time applications. Furthermore, these methods are robust to noise in the images. In the application, the image processing uses the geometry of the evolved contour to calculate the locations of the center and wing-tips of the target aircraft in the image plane. The location of the target center provides relative bearing (azimuth and elevation) measurements. The angle subtended by the line between the wing-tips subtends an angle on the image plane referred to as the subtended angle as shown in Figure 3. Further details on the theory and implementation of the image processing used in this paper are referred to Ref. [3, 18] and the references within.

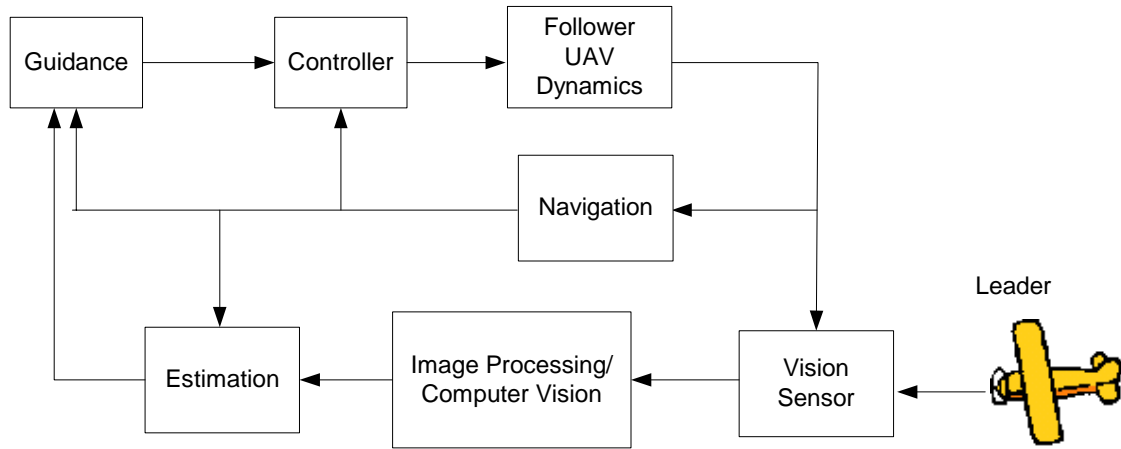


Figure 1. Vision-in-the-loop Flight Control System for Follower Aircraft



Figure 2. Sample image processing with background clutter a) Active Contour encircling Target, b) Feature Point Extraction from Image

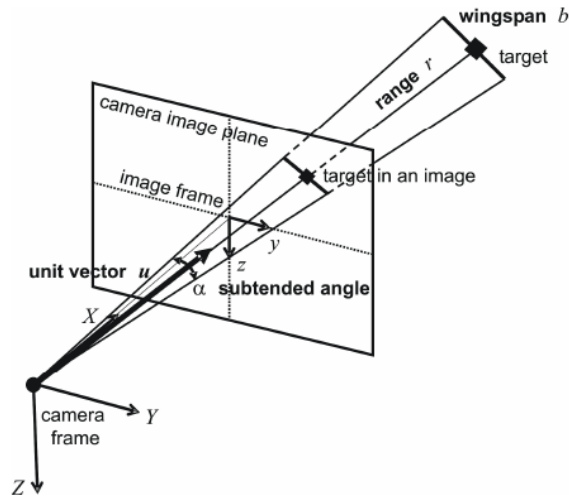


Figure 3. Image Plane Measurements of Target Aircraft ^{3,14}

A consequence of using a monocular fixed camera is that the range to the leader or target aircraft is not available as a direct measurement. So the measurements from the image processing algorithms have to be processed by a nonlinear filter, for example, an EKF, to generate estimates of the relative position, velocity and acceleration of the target with respect to the follower. These estimates are used to generate estimates of the target position, velocity and acceleration in 3D inertial space assuming the follower is equipped with an IMU for own-ship information. A critical source for uncertainty in the estimation process is the lack of knowledge of the target motion. Unmodeled target maneuvers lead to bias in the estimates of the range and line-of-sight (LOS) variables and in some cases can lead to divergence. The estimation process has to be made robust to all such uncertainties before being employed in closed-loop control and the next section summarizes one such adaptive approach.

III. Adaptive State Estimation

A. Problem Formulation

Consider the following *bounded* single-input-single-output (SISO) nonlinear system

$$\begin{aligned}\dot{\mathbf{x}} &= A\mathbf{x} + Bg(\mathbf{x}, \mathbf{z}), & \mathbf{x}(0) &= \mathbf{x}_0 \\ \dot{\mathbf{z}} &= \mathbf{f}_z(\mathbf{x}, \mathbf{z}), & \mathbf{z}(0) &= \mathbf{z}_0 \\ y &= C\mathbf{x}\end{aligned}\tag{1}$$

where $\mathbf{x} \in D_x \subseteq R^{n_x}$ and $\mathbf{z} \in D_z \subseteq R^{n_z}$ are the states of the system such that \mathbf{x} represents the modeled states and \mathbf{z} represents the unmodeled states, D_x and D_z are compact sets, $\mathbf{f}_z(\mathbf{x}, \mathbf{z}): R^{n_x} \times R^{n_z} \rightarrow R^{n_z}$ is an unknown, bounded function and represents the unmodeled dynamics, $g(\mathbf{x}, \mathbf{z}): R^{n_x} \times R^{n_z} \rightarrow R$ is an unknown, uniformly bounded and continuous function and represents the way in which the unmodeled dynamics is coupled to the system dynamics, for which n_z is also unknown but bounded, $y \in R$ represents the available measurement which is assumed to be bounded, the matrices (A, B, C) are known and the pair (A, C) is observable.

Remark 1: The function $g(\mathbf{x}, \mathbf{z})$ acts as the unknown system input or disturbance to the nominal linear system given by the matrices (A, B, C) .

Objective: Design a state estimator to estimate the states \mathbf{x} of the system in (1) with bounded estimation error in the presence of the unknown system input $g(\mathbf{x}, \mathbf{z})$.

As in Refs. [10,19], we consider the following NN parameterization of $g(\mathbf{x}, \mathbf{z})$

$$g(\mathbf{x}, \mathbf{z}) = \mathbf{W}^T \boldsymbol{\sigma}(\bar{\boldsymbol{\mu}}) + \varepsilon(\bar{\boldsymbol{\mu}}), \quad \|\mathbf{W}\|_F \leq W^*, \quad \|\varepsilon(\bar{\boldsymbol{\mu}})\| \leq \varepsilon^*, \quad \|\bar{\boldsymbol{\mu}}\| \leq \mu^*\tag{2}$$

$\forall (\mathbf{x}, \mathbf{z}) \in D_g \subset D_x \times D_z$, D_g is a compact set, $\boldsymbol{\sigma}(\bar{\boldsymbol{\mu}}) = [\sigma_1(\bar{\boldsymbol{\mu}}), \dots, \sigma_N(\bar{\boldsymbol{\mu}})]^T$ is a vector of sigmoidal functions $\sigma_i(\cdot)$, W^* and ε^* are the bounds on the Frobenius norms of the ideal, unknown weight vector \mathbf{W} and of the NN functional approximation error ε respectively, N is the number of neurons, and the input vector $\bar{\boldsymbol{\mu}} = \bar{\boldsymbol{\mu}}(y(t), d)$ is the vector of difference quotients of the output y

$$\bar{\boldsymbol{\mu}}(y(t), d) = \left[1, \Delta_d^{(0)} y^T(t) \quad \dots \quad \Delta_d^{(n-1)} y^T(t) \right]\tag{3}$$

where $\Delta_d^{(0)} y^T(t) = y^T(t)$, $\Delta_d^{(k)} y^T(t) = \frac{\Delta_d^{(k-1)} y^T(t) - \Delta_d^{(k-1)} y^T(t-d)}{d}$, $k = 1, 2, \dots$, and $d > 0$ is a time delay. The

sigmoidal functions are smooth and uniformly bounded, that is, $|\sigma_i(\bar{\boldsymbol{\mu}})| \leq 1$.

Remark 2: Equation (2) simply states that a continuous function of the states of the system (1) can be approximated to an arbitrary degree of accuracy by a NN over a compact domain with inputs that are a finite sample of the output history of the system.

Consider the following time-varying filter to estimate the states of the system in (1):

$$\begin{aligned}\dot{\hat{\mathbf{x}}}(t) &= A\hat{\mathbf{x}}(t) + K(t)(y(t) - \hat{y}(t)) + Bv_{ad}, & \hat{\mathbf{x}}(0) &= \hat{\mathbf{x}}_0 \\ \hat{y}(t) &= C\hat{\mathbf{x}}(t)\end{aligned}\tag{4}$$

where $K(t)$ is the Kalman gain obtained through the following set of matrix differential Riccati equations ²⁰

$$\begin{aligned}\dot{P}(t) &= AP(t) + P(t)A^T - P(t)C^T R^{-1} CP(t) + Q \\ K(t) &= P(t)C^T R^{-1}\end{aligned}\quad (5)$$

where $P(0) = P_0 > 0$, $Q = Q^T > 0$, $R = R^T > 0$ are design constants. The solution $P(t)$ of (5) is bounded, symmetric, positive definite and continuously differentiable. The output of the NN v_{ad} is given by

$$v_{ad} = \hat{W}(t)^T \sigma(\bar{\mu}) \quad (6)$$

where $\hat{W}(t)$ is the estimate of the weight vector W in (2) and v_{ad} is designed to approximate the bounded disturbance $g(x, z)$. The residual signal of the adaptive estimator $\tilde{y}(t) = y(t) - \hat{y}(t)$ is used to train the NN. A second training signal can be derived to improve the effectiveness of the NN adaptation performance. The approach to training the NN with multiple training signals is referred to as the composite adaptation approach ¹¹. Further details regarding the theory are referred to Refs. [10,11] and the references within. We complete this discussion by showing a block diagram of the composite adaptation based state estimator in Figure 4.

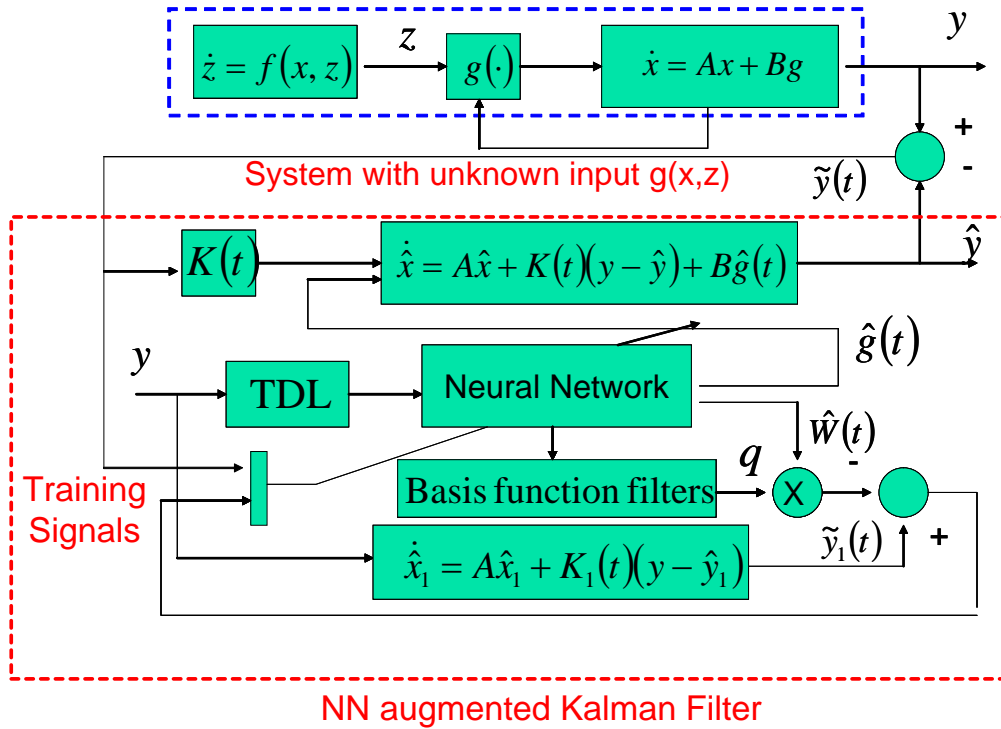


Figure 4. Composite Adaptation based Adaptive State Estimation

B. Application to Vision-based Target Tracking

Consider the relative LOS kinematics between a target and follower aircraft in the inertial Cartesian coordinate frame

$$\frac{d}{dt} \begin{bmatrix} R_X \\ \dot{R}_X \\ R_Y \\ \dot{R}_Y \\ R_Z \\ \dot{R}_Z \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_A \begin{bmatrix} R_X \\ \dot{R}_X \\ R_Y \\ \dot{R}_Y \\ R_Z \\ \dot{R}_Z \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_B \begin{bmatrix} a_{T_x} \\ a_{T_y} \\ a_{T_z} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -a_{F_x} \\ -a_{F_y} \\ -a_{F_z} \end{bmatrix} \quad (7)$$

where R_X, R_Y and R_Z are respectively the projections of the range vector from the follower to the target aircraft onto the inertial X, Y and Z axes, $[a_{T_x}, a_{T_y}, a_{T_z}]^T$ is the unknown target acceleration, and $[a_{F_x}, a_{F_y}, a_{F_z}]^T$ is the known follower acceleration. Eq. (7) is used in the design of the nominal time-varying Kalman filter where the target acceleration components along the inertial coordinate axes are modeled as independent, zero-mean white noise processes. We assume that there is a vision sensor and image processing onboard the follower aircraft that can measure the unit vector to the target aircraft in the camera fixed axes $[u_{x_m}, u_{y_m}, u_{z_m}]^T$ and the subtended angle α_m . The subscript 'm' indicates measurement. The unit vector measurements are transformed into the inertial coordinate (NED) frame $[u_{x_m}, u_{y_m}, u_{z_m}]^T$. The measurements to the NN augmented Kalman filter are constructed as follows:

$$\begin{aligned} R_{x_m} &= \frac{b}{2 \tan(\alpha_m/2)} u_{x_m} \\ R_{y_m} &= \frac{b}{2 \tan(\alpha_m/2)} u_{y_m} \\ R_{z_m} &= \frac{b}{2 \tan(\alpha_m/2)} u_{z_m} \end{aligned} \quad (8)$$

where the target wing-span b is assumed to be known. This conversion of the image plane noisy angle measurements into the measurements of R_X, R_Y and R_Z allows us to use the following linear, measurement model for the LOS kinematics:

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}}_Y = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}}_C \underbrace{\begin{bmatrix} R_X \\ \dot{R}_X \\ R_Y \\ \dot{R}_Y \\ R_Z \\ \dot{R}_Z \end{bmatrix}}_X + \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} \quad (9)$$

where v_x, v_y and v_z are now state-dependent measurement noise terms.

IV. Simulation Results

The adaptive state estimation algorithm presented in Section III was integrated with the image processing, guidance and flight control algorithms in the 6DOF GUST real-time simulation software¹². The image processing, a non-adaptive EKF as the target state estimator, the guidance and NN-based adaptive flight control algorithms have been separately validated. These were then flight-tested in closed-loop and completely autonomous vision-based formation flight test results between two UAVs in a leader-follower configuration were reported in Ref. [3].

Figure 5 shows a screenshot of the vision-based formation flight simulation. The ‘src’ screenshot is the frame-grabber window which is used by the image processing algorithm to capture the leader aircraft center (green crosshair) and wingtips (red crosshairs). The ‘Scene Window 1’ and ‘Scene Window 2’ screenshots depict the formation view from the top and from behind the follower aircraft respectively. The green circles in these screenshots depict the target estimator’s estimate of the leader position.

The leader aircraft is flying in a circle in the horizontal plane at a constant heading rate. The follower aircraft is tasked with maintaining specified separation distances along the x-, y- and z-axes of the follower body-fixed frame. The follower is first put into the desired formation using only GPS communicated data of the leader inertial position, velocity and acceleration. The leader GPS data is communicated at about 5 Hz and is filtered to produce leader state estimates at the rate required (~50 Hz) by the follower aircraft guidance and flight control algorithms. Once the leader aircraft is at the desired separation distance, the image processing and target state estimation algorithms are switched on. The update rate of the image processing in simulation ranges is ~ 10 Hz. The estimates of the leader position, velocity and acceleration from using the vision-based target state estimator are blended in with the corresponding GPS estimates to produce the leader state estimates that are used in the guidance and flight control algorithms for formation keeping.

The formation separation commands for the results shown below are given by $[dx,dy,dz]_{com} = [60,15,10]$ ft.

Figure 6 shows the flags pertaining to the image processing and adaptive estimation. The ‘IP flag’ indicates if the image processing (IP) has returned a measurement. The ‘useVision Fraction’ flag indicates the fraction of the vision-based leader position, velocity and acceleration estimates used in the guidance law relative to GPS estimates of the same. Figure 6 shows that the vision estimates are first blended in (about 50 %) at about 140 seconds, and the vision

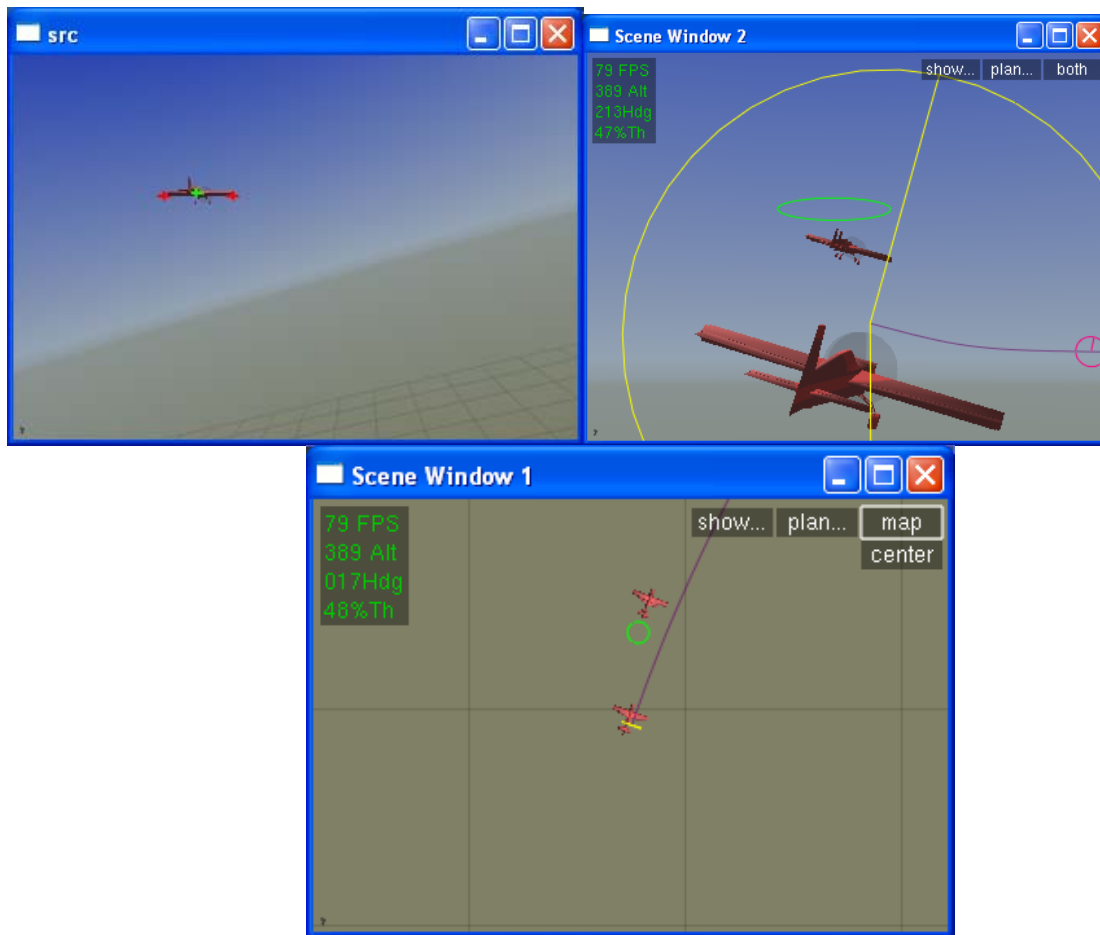


Figure 5. Screenshot of Vision-based Formation Flight in Real-Time Simulation

estimates are in complete control of the follower aircraft at about 290 seconds. The ‘NN switch’ flag indicates if the NN augmenting the nominal Kalman filter is switched on/off.

Figure 7 plots the IP relative position measurements along the inertial coordinate axes along with their corresponding GPS measurements. On the right hand side, the difference between these measurements is shown. The apparent chattering in the plot is due to the chattering of the GPS measurements. This plot shows that the image processing measurements are consistent with the corresponding GPS measurements.

Figure 8 plots the formation position command tracking performance. The magnitudes of the tracking errors are within acceptable bounds, but there is clearly room for improvement. Since the overall control architecture involves integration of several different components, it is difficult to pinpoint the exact steps to take to improve the performance. The accuracy of the tracking is definitely related to the update rate and accuracy of the image measurements relative to the GPS measurements shown in Figure 7. The image measurements chosen depend on the choice of the coordinates (rectangular or spherical) used to construct the target state estimator. At large separation distances, the subtended angle α_m becomes very small and the accuracy of the measurements in Eq. (8) is decreased. The target state estimator using the modified spherical coordinates in Ref. [3] is more robust to inaccuracies in α_m .

Figure 9-Figure 11 show the leader position, velocity and acceleration estimation performance. In Figure 11, the estimates of the leader acceleration are the outputs of the adaptive state estimator. The plots show reasonable estimation performance with the NN capturing the leader acceleration fairly accurately.

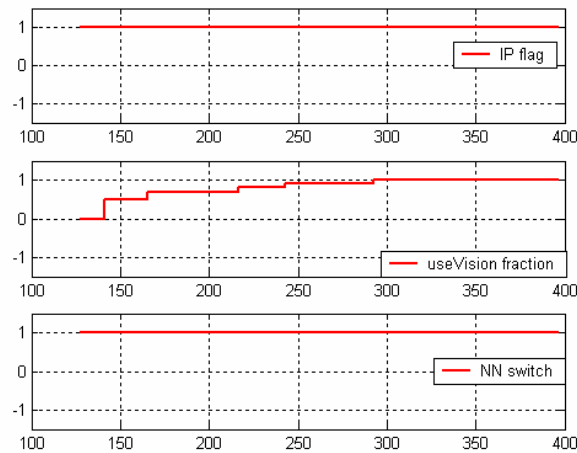


Figure 6. Image Processing (IP) flag, use Vision fraction, and NN on/off switch, with Adaptive Estimation

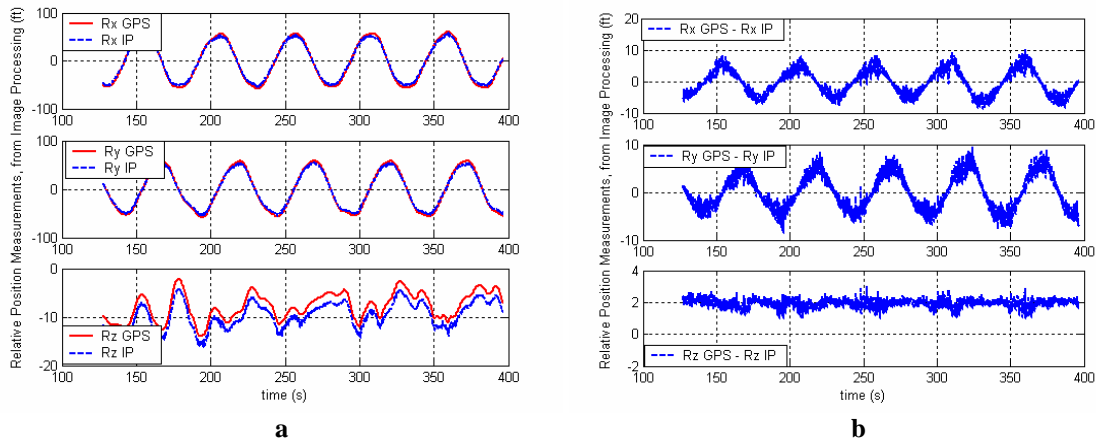


Figure 7. IP Relative Position Measurements and corresponding GPS measurements (ft)

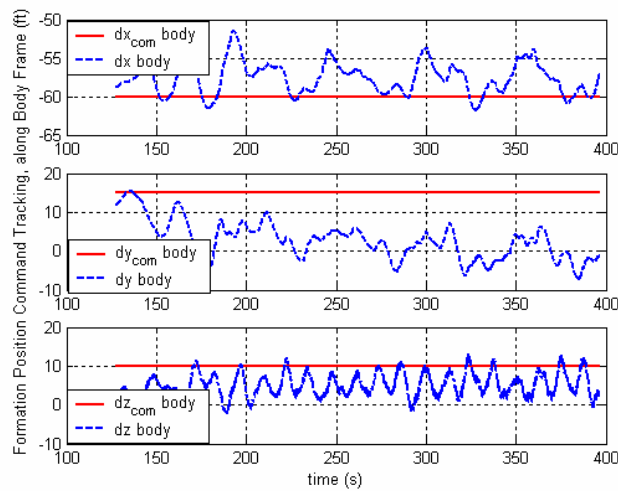


Figure 8. Formation Position Command Tracking, with blend of vision and GPS, with Adaptive Estimation

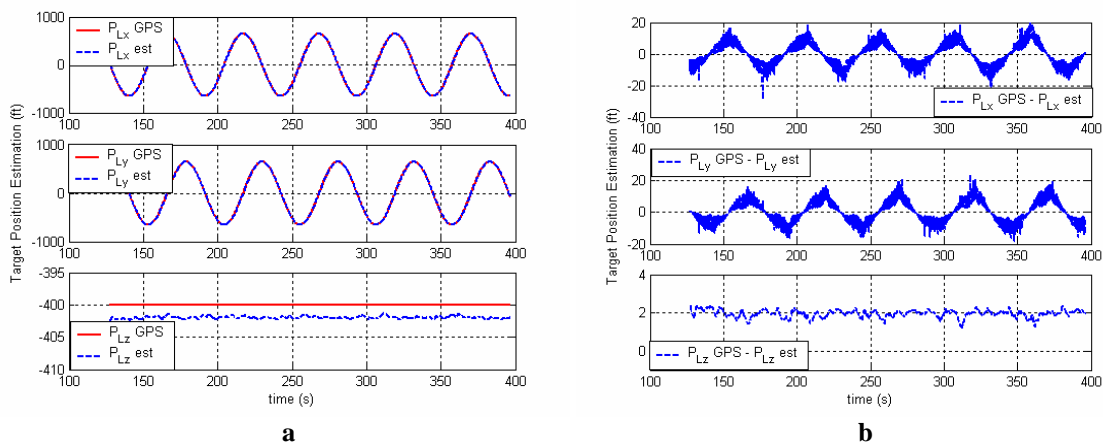


Figure 9. Leader Position Estimation Performance (ft), with Adaptive Estimation

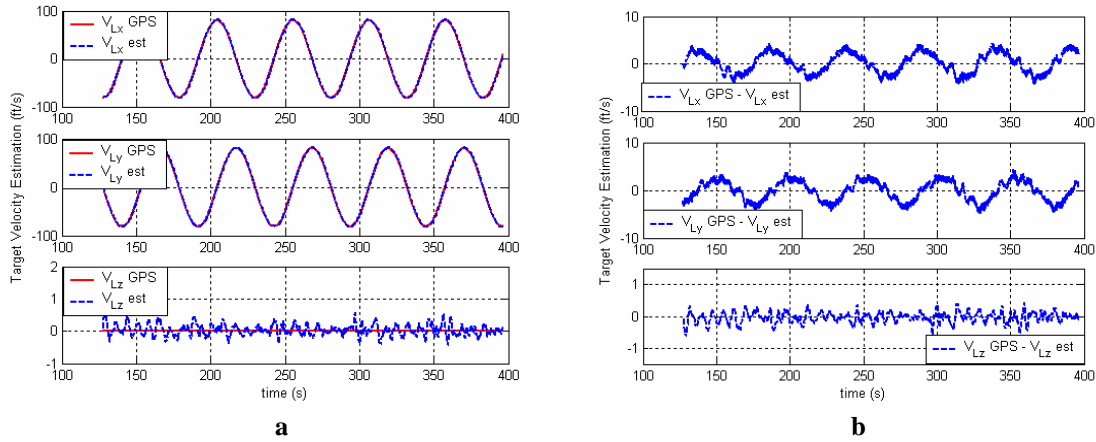


Figure 10. Leader Velocity Estimation Performance (ft/s), with Adaptive Estimation

Figure 12 and Figure 13 show the result if the NN augmenting the Kalman filter is switched off. As soon as the vision estimates are blended in (20% vision at about 148 seconds), the 'IP' flag is temporarily zero suggesting loss of vision track. As the percentage of vision estimates used increases (40% vision at about 175 seconds), the range diverges. Without adaptation in the estimation, the leader aircraft just drifts out the field-of-view causing loss of visual track and ultimately range divergence.

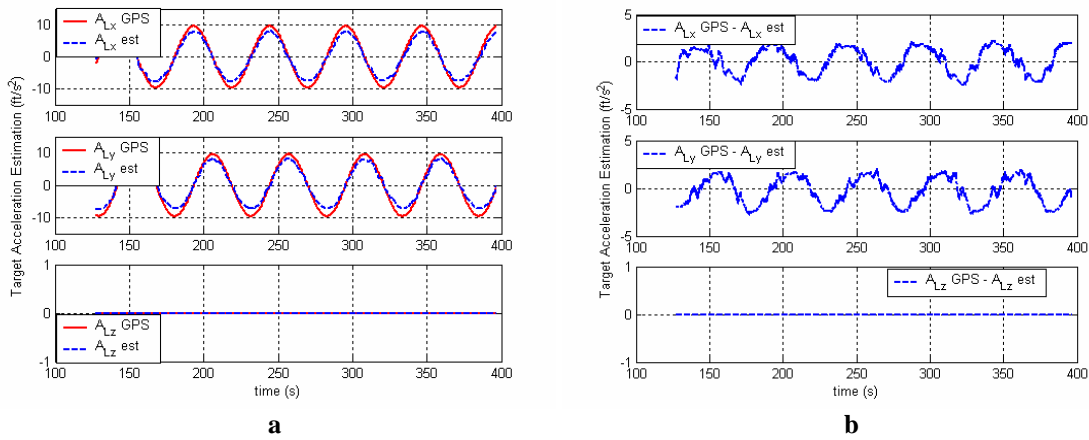


Figure 11. Leader Acceleration Estimation Performance (ft/s²), with Adaptive Estimation

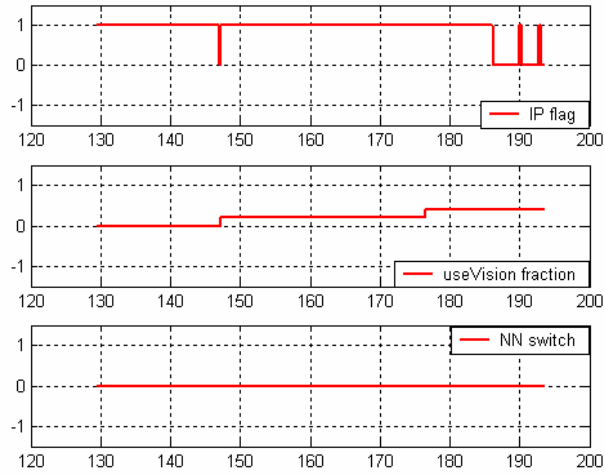


Figure 12. Image Processing (IP) flag, use Vision fraction, and NN on/off switch, without Adaptive Estimation

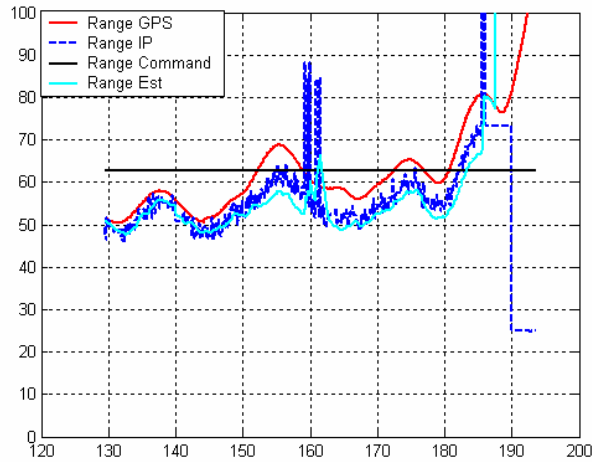


Figure 13. Range Estimation, without Adaptive Estimation

V. Conclusions

This paper discusses the implementation of a NN-augmented Kalman filter as an adaptive target state estimator in a vision-based target tracking and autonomous formation flight problem. The design of the adaptive target estimator reduces reliance on *a priori* knowledge of the target maneuver and/or of avoids construction of elaborate target maneuver models. The benefits of such a design are clearly illustrated via the vision-in-the-loop 6DOF simulation results. With adaptive estimation, the unknown target maneuver is fairly accurately captured by the output of the adaptive NN and vision-in-the-closed-loop formation flight is maintained. Without adaptation to compensate the nominal Kalman filter, the leader aircraft drifts out of the field-of-view of the follower onboard camera and vision formation cannot be maintained.

Using a single onboard vision sensor for relative navigation itself is an important contribution in that it significantly reduces the cost and complexity of airborne navigation systems that rely on GPS, INS and related communication systems. Future research efforts will focus on the flight testing of the presented adaptive estimation method in closed-loop vision-based formation flight, ground target tracking and obstacle avoidance applications.

Acknowledgments

This research has been sponsored under AFOSR contract F49620-03-1-0401. The authors thank Yoko Watanabe, Nimrod Rooz and Dr. Suresh Kannan at the School of Aerospace Engineering in Georgia Tech for their assistance in setting up the simulation.

References

- ¹ Office of Naval Research Broad Agency Announcement #ONR05-012, Navigation and Timekeeping Technology, web-site <http://www.onr.navy.mil/02/baa>
- ² A.D. Wu, E.N. Johnson and A. Proctor, "Vision-aided Inertial Navigation for Flight Control," *AIAA Guidance, Navigation and Control Conference*, San Francisco, CA, August 2005.
- ³ E. N. Johnson, A.J. Calise, Y. Watanabe, J. Ha and J. C. Neidhoefer, "Real-Time Vision-Based Relative Navigation," *AIAA Guidance, Navigation, and Control Conference*, Keystone, CO, August 2006.
- ⁴ J. Valasek, J. Kimmitt, D. Hughes, K. Gunnam and J. L. Junkins, "Vision based Sensor and Navigation System for Autonomous Aerial Refueling," *AIAA's 1st Technical Conference and Workshop on Unmanned Aerospace Vehicles*, Portsmouth, VA, May 2002.
- ⁵ X.R. Li and V.P. Jilkov, "Survey of Manuevering Target Tracking, Part I: Dynamic Models", *IEEE Transactions on Aerospace and Electronic Systems*, AES-39, pp. 1333 -1364, October 2003.
- ⁶ Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software*, Wiley-Interscience, New York, 2001.
- ⁷ R.A. Singer, "Estimating Optimal Tracking Filter Performance for Manned Maneuvering Targets", *IEEE Transactions on Aerospace and Electronic Systems*, AES-6, pp. 473-483, July 1970.
- ⁸ H.A.P. Blom and Y. Bar-Shalom, "The Interacting Multiple-Model Algorithm for Systems with Markovian Switching Coefficients", *IEEE Transactions on Automatic Control*, AC-33, pp. 780-783, August 1988.
- ⁹ J.A. Roecker and C.D. McGillem, "Target Tracking in Maneuver Centered Coordinates," *IEEE Transactions on Aerospace and Electronic Systems*, AES-25, pp. 836-843, November 1989.
- ¹⁰ V.K. Madyastha and A.J. Calise, "An Adaptive Filtering Approach to Target Tracking", *American Control Conference*, pp. 1269-1274, June 2005.
- ¹¹ R. Sattigeri and A.J. Calise, "Neural Network Augmented Kalman Filtering in the Presence of Unknown System Inputs," *AIAA Guidance, Navigation, and Control Conference*, Keystone, CO, August 2006.
- ¹² E.N. Johnson and D.P. Schrage, "System Integration and Operation of a Research Unmanned Aerial Vehicle," *AIAA Journal of Aerospace Computing, Communication, and Information*, 1 (1), pp. 5-18, January 2004
- ¹³ J. Ha, C. Alvino, G. Pryor, M. Niethammer, E. Johnson, and A. Tannenbaum, "Active Contours and Optical Flow for Automatic Tracking of Flying Vehicles," *American Control Conference*, Vol. 4, pp 3441-3446, 2004.
- ¹⁴ A. Betsler, P. Vela and A. Tannenbaum, "Automatic Tracking of Flying Vehicles using Geodesic Snakes and Kalman Filtering," *IEEE Conference on Decision and Control*, Atlantis Island, Bahamas, December 2004.
- ¹⁵ L. Pollini, R. Mati and M. Innocenti, "Experimental Evaluation of Vision Algorithms for Formation Flight and Aerial Refueling," *AIAA Modeling and Simulation Technologies Conference and Exhibit*, Providence, RI, August 2004.
- ¹⁶ M. Kass, A. Witkin and D. Terzopoulos, "Snakes: Active Contour Models", *International Journal of Computer Vision*, 1(4), pp. 321-331, 1987.
- ¹⁷ J. Sethian, *Level Set Methods and Fast Marching Methods*, Cambridge Press, 1996.
- ¹⁸ J. Ha, E.N. Johnson and A. Tannenbaum, "Real-Time Visual Tracking using Geometric Active Contours for the Navigation and Control of UAVs", *Proceedings of the American Control Conference*, 2007.
- ¹⁹ E. Lavretsky, N. Hovakimyan and A.J. Calise, "Upper Bounds for Approximation of Continuous Time Dynamics using Delayed Outputs and Feedforward Neural Networks", *IEEE Transactions on Automatic Control*, AC-48, pp. 1606-1610, September 2003.
- ²⁰ R. G. Brown and P.Y.C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, John Wiley and Sons Inc., 1992.