Vision-Based Tracking and Motion Estimation for Moving targets using Small UAVs

Vladimir N. Dobrokhodov, Isaac I. Kaminer *Member, IEEE*, Kevin D. Jones, and Reza Ghabcheloo

Abstract— **This paper addresses the development of a vision-based target tracking system for a small unmanned air vehicle. The algorithm performs autonomous tracking of a moving target, while simultaneously estimating GPS coordinates of the target. A low cost off the shelf system is utilized, with a modified radio controlled aircraft airframe, gas engine and servos. Tracking is enabled using a low-cost, miniature pan-tilt gimbal. The control algorithm provides rapid and sustained target acquisition and tracking capability. A target position estimator was designed and shown to provide reasonable targeting accuracy. The impact of target loss events on the control and estimation algorithms is analyzed in detail.**

I. Introduction

THE past two decades have witnessed a remarkable increase in the utilization of unmanned air vehicles (UAVs) both in the US and abroad. While many of the large UAV systems are quite capable, their cost is also very high. Consequently there is much interest in the development of small, low-cost platforms which can perform some of the tasks normally assigned to larger UAVs, for example vision-based target tracking.

This paper addresses the development of a vision based target tracking and position estimation system for a small UAV. This work is an extension of the results reported [1]. Therefore most of the details addressing the hardware design and software implementation have been omitted. In this paper the case of a moving target is studied.

The platform used to test the system is a modified RC aircraft with a miniature pan-tilt gimbaled camera built using COTS components (see Fig. 1). In a typical operational scenario, the system operator may select a target of interest using a joystick that steers the onboard camera. Once a target is selected, the UAV and the camera automatically track the target and provide an estimate of its position, velocity and heading. The target can be either stationary or moving.

Manuscript received September 23, 2005. This work was supported by the U.S. Government under the grants from USSOCOM and CDTEMS.

Vladimir N. Dobrokhodov is with the Mechanical and Astronautical Engineering department, Naval Postgraduate School, Monterey, CA 93943 USA (phone: 831-656-7714; fax: 831-656-2313; e-mail: vldobr@ nps.edu).

Isaac I. Kaminer is with the Mechanical and Astronautical Engineering department, Naval Postgraduate School, Monterey, CA 93943 USA (e-mail: kaminer@ nps.edu).

Kevin D. Jones is with the Mechanical and Astronautical Engineering department, Naval Postgraduate School, Monterey, CA 93943 USA (e-mail: jones@ nps.edu).

Reza Ghabcheloo is with Instituto de Sistemas e Robotica (ISR) at Instituto Superior Tecnico (IST), Lisbon, Portugal (mail reza@isr.ist.utl.pt).

To keep the airborne system cost low, much of the expensive equipment is left on the ground. The video is transmitted to the ground, where it is processed in real time. The centroid of the target in the camera frame is identified by an image processing algorithm and is used to drive the integrated UAV/gimbal control algorithm, which in turn steers the UAV and the gimbal to keep the target in the center of the camera frame.

Reliance on inexpensive off the shelf equipment as well as communication interrupts due to the RFI resulted in frequent loss of tracking by the system. Therefore, the key technical challenge was to design control and motion estimation algorithms that were robust in the presence of loss of tracking events. Therefore, the design and analysis of both the control and, particularly, the motion estimation algorithms have borrowed heavily from the theory of systems with brief instabilities [2] and the linear parametrically varying (LPV) systems [3]. In this paper the target-loss events were modeled as brief instabilities.

Fig. 1. Modified Telemaster UAV.

The paper is organized as follows. The design of the UAV control algorithm is discussed in Section II. The development of the target motion estimator is included in Section III. The results of flight experiments with moving targets are discussed in Section IV. The paper ends with some concluding remarks.

II. Control System Development

Consider Fig.2. Let ρ denote range from the UAV to the target, $\vec{V}_g$ - the UAV ground speed, $\vec{\lambda}_g$ - the line of sight (LOS) vector and $\vec{\lambda}_p$ - the vector perpendicular to $\vec{\lambda}_g$. Furthermore, let ε denote the angle between the LOS vector and the camera heading, λ - the LOS angle, ψ - the UAV

heading, ψ_h - the gimbal pan angle and η - the angle between the $\vec{V}_g$ and $\vec{\lambda}_p$ vectors.

In addition, suppose the target is moving with constant speed $\vec{V}_t$, and heading, ψ_t as shown in Fig.2.

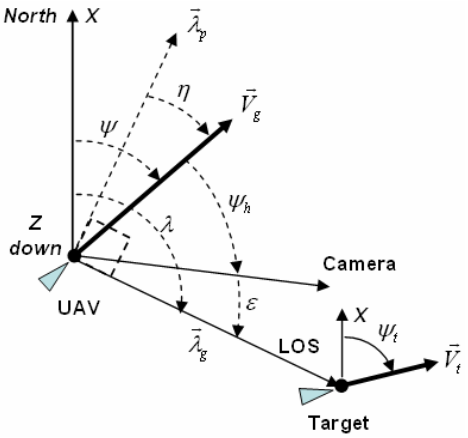


Fig. 2. Moving target tracking for the control law (4).

From Fig.2 it can be shown that tracking problem kinematics for a moving target are given by

$$\begin{aligned}\dot{\eta} &= -\frac{V_g\cos\eta - V_t\cos(\psi_t-(\psi-\eta))}{\rho}+\dot{\psi}\\ \dot{\varepsilon} &= \frac{V_g\cos\eta - V_t\cos(\psi_t-(\psi-\eta))}{\rho}-\dot{\psi}-\dot{\psi}_h \\ \dot{\rho} &= -V_g\sin\eta + V_t\sin(\psi_t-(\psi-\eta))\end{aligned} \quad (1)$$

The control objective is to drive ε and η to zero using the UAV turn rate $\dot{\psi}$ and pan rate $\dot{\psi}_h$ as control inputs. To this end the following control law is proposed

$$\begin{aligned}\dot{\psi} &= \frac{V_g}{\rho_d}\cos\eta - k_1\eta \\ \dot{\psi}_h &= k_1\eta + k_2\varepsilon\end{aligned}, \quad (2)$$

where ρ_d denotes a desired horizontal range to target to be selected by the operator. Define $\rho_e = \frac{1}{\rho_d} - \frac{1}{\rho}$ and $\bar{\rho} = \rho_e + \frac{1}{\rho_d}$. Then in can be shown that the feedback system consisting of (1) and (2) is given by (3):

$$\begin{aligned}\dot{\eta} &= -V_g\rho_e\cos\eta - k_1\eta - d\bar{\rho}\\ \dot{\varepsilon} &= V_g\rho_e\cos\eta - k_2\varepsilon - d\bar{\rho}\\ \dot{\rho}_e &= (\rho_e + 1/\rho_d)^2 V_g\sin\eta + d\end{aligned}, \quad (3)$$

where $|d| \le V_t$ represents a bounded disturbance. It can be shown that this system is uniformly ultimately bounded. The proof can be found in [4].

Results of a full scale nonlinear simulation (Fig.4) show that control law performs remarkably well when tracking a moving target while using information obtained from the onboard camera and the UAV velocity available from onboard GPS. Note, in the presence of target loss events the control system maintains the last turn rate command generated during target lock.

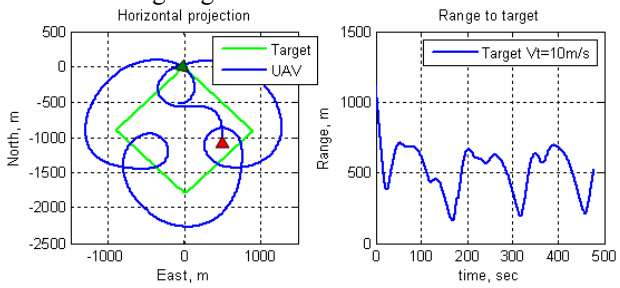


Fig. 4. UAV motion versus target motion.

III. Range Estimation

The standard approach to range estimation using a single camera involves triangulating between two consecutive points along the UAV path. At each point the stored measurements include the UAV position and the LOS angle of the onboard camera. Care must be taken that the baseline (distance between these points) is sufficiently large to guarantee low Dilution of Precision (DOP). Clearly, for a UAV tracking a target along the circular path this approach will result in a large wait time between each measurement.

In this paper we assume that UAV's altitude above target is known and use it as an additional measurement. To obtain this measurement we use the filter developed in [1] to get target's latitude and longitude. Target's altitude is then obtained from a geo-referenced database made available by the Perspective View Nascent Technologies (PVNT) [5] software package by providing it with target's estimated latitude and longitude.

Consider Fig.5, which depicts an aircraft equipped with a gimbaled optical camera pointing to the moving ground target. Let *{I}* denote an inertial reference frame, *{B}* a body-fixed frame that moves with the UAV, and *{C}* a gimbaled-camera frame that coincides with body frame origin and rotates with respect to *{B}*.

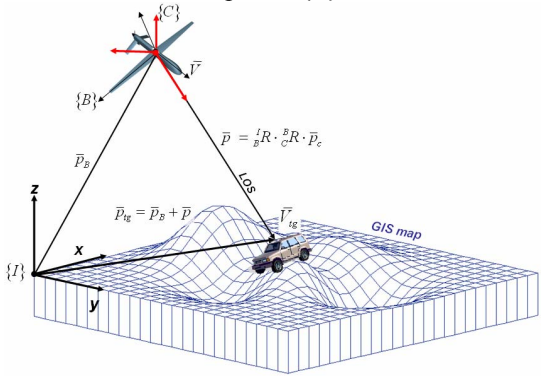


Fig. 5. UAV-Target relative kinematics.

Suppose that the target's inertial velocity ($\bar{V}_{tg}$) and heading ($\bar{\psi}_{tg}$) are fixed. Following the notations introduced in [2], let $\bar{p}_c = [x_c \; y_c \; z_c]^T$ denote the relative position of the center of {C} with respect to target, and let ${}^I_C R$, ${}^I_B R$ and ${}^B_C R$ denote coordinate transformations from {C} to {I}, from {B} to {I} and from {C} to {B} respectively. These transformations are available onboard and resulted from the IMU attitude measurements and the positional pan/tilt feedback of gimbaled camera.

From these definitions, it follows that

$$\bar{p}_{tg} = \bar{p}_B + {}^I_C R \cdot \bar{p}_c \;\Rightarrow$$
$$\frac{d}{dt}({}^I_C R \cdot \bar{p}_c) = -\frac{d}{dt}(\bar{p}_B) + \frac{d}{dt}(\bar{p}_{tg}) \quad . \tag{4}$$

Introducing $\bar{V}_{tg} = \frac{d}{dt}(\bar{p}_{tg})$ and $V = \frac{d}{dt}(\bar{p}_B)$ with an assumption of constant speed of the target provides the following process equations

$$\begin{cases} \frac{d}{dt}({}^I_C R \cdot \bar{p}_c) = -\bar{V} + \bar{V}_{tg} \\ \frac{d}{dt}\bar{V}_{tg} = 0 \end{cases} . \tag{5}$$

Here, measurements of the camera and its gimbal angles contribute to the $\bar{p}_c$ in first equation. In order to introduce these measurements to the process model, we assume that the camera readings are obtained using simple pinhole camera model of the form

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{f}{x_c}\begin{bmatrix} y_c \\ z_c \end{bmatrix} . \tag{6}$$

In this equation f is a focal length of the camera and $[u \; v]^T$ are the coordinates of the centroid of target image in a camera frame. Since the camera onboard is gimbaled (directly controlled through pan and tilt angular commands), the target is always located in front of the camera's image plane, i.e. $x_c > 0$. As discussed above in addition to measurements (6) we use UAV's altitude above target:

$$z = -{}^b x_c \sin\theta + {}^b y_c \sin\varphi \sin\theta + {}^b z_c \cos\varphi \cos\theta , \tag{7}$$

where φ, θ are the roll and pitch Euler angles that determine orientation of the camera with respect to {I}. This equation is a linear combination of the third row of ${}^I_C R$ and the $\bar{p}_c$ resolved in {B} ${}^b\bar{p}_c = [{}^b x_c \; {}^b y_c \; {}^b z_c]^T = {}^B_C R \cdot \bar{p}_c$. Let $y_m = [u \; v \; z]^T = g_{\varphi\theta}(p_c)$, then

$$g_{\varphi\theta}(p_c) = \begin{bmatrix} \frac{f}{x_c} \cdot \begin{bmatrix} y_c \\ z_c \end{bmatrix} \\ -{}^b x_c \sin\theta + {}^b y_c \sin\varphi \sin\theta + {}^b z_c \cos\varphi \cos\theta \end{bmatrix}$$

Therefore, the process model considered in this paper in the following form

$$\begin{cases} \frac{d}{dt}({}^I_C R \cdot \bar{p}_c) = -V + \bar{V}_{tg} \\ \frac{d}{dt}\bar{V}_{tg} = 0 \\ \bar{y}_m = g_{\varphi\theta}(\bar{p}_c) + \bar{w}_y \end{cases} , \tag{8}$$

where $\bar{y}_m$ denotes camera and altitude measurements corrupted by the process noise $\bar{w}_y$.

The practical problem now consists of determining the relative position and velocity of the moving target with respect to a UAV using IMU, GPS and tracking camera measurements complemented by the altitude above the target provided by the PVNT. In [2], a general structure of a nonlinear filter with guaranteed stability and performance characteristics that solves this problem in the presence of measurement noise was proposed, while in [3] these results were extended to include out-of frame events typical for vision-based applications.

During numerous flight tests [1] the image tracking software (see Section IV) lost target track on a regular basis. This prompted the following question: can the filtering solution maintain stability in the presence of target loss events? In fact, the ideas presented in [2] and [3] are used in this paper to derive a nonlinear filter that tracks a moving target using the process model (8) in the presence of out-of-frame events.

Following the development in [3], define an out-of frame event as a binary signal s: $[0,\infty) \to \{0, 1\}$

$$s = s(t) = \begin{cases} 0 - \text{out of frame event at time } t \\ 1 - \text{camera tracks the target at time } t \end{cases} .$$

For a given binary signal s and $t > \tau > 0$, let $T_s(\tau, t)$ denote the length of time in the interval (τ, t) that $s = 0$. Then formally

$$T_s(\tau, t) := \int_\tau^t (1 - s(l))dl .$$

The signal s is said to have brief target loss events if $T_s(\tau, t) \le T_0 + \alpha(t - \tau)$, $\forall t \ge \tau \ge 0$, for some $T_0 \ge 0$ and $\alpha \in [0,1]$.

Next, consider that the orientation of the camera frame installed onboard the UAV is constrained by a compact set

$$\Lambda_c = \left\{ \lambda : |\varphi| \le \varphi_{\max}, |\theta| \le \theta_{\max} \right\} \tag{9}$$

and that the relative position of the UAV relative to the

target and expressed in a camera frame is constrained to be in

$$P_c = \left\{ p_c = [x_c, y_c, z_c]^T : \begin{array}{r} x_{\min} \le x_c \le x_{\max} \\ y_{\min} \le y_c \le y_{\max} \\ z_{\min} \le z_c \le z_{\max} \end{array} \right\}. \quad (10)$$

Notice that the yaw angle ψ is not limited, $x_{\min} \ldots z_{\max}$ are chosen according to the geometry of the mission and the relative vehicles dynamics.

Filter is designed to provide the estimates $\hat{p}_c$ of p_c to be bounded by

$$\hat{P}_c = \left\{ \hat{\bar{p}}_c = [\hat{x}_c, \hat{y}_c, \hat{z}_c]^T : \begin{array}{r} |\hat{x}_c - x_c| \le x_{\max} - x_{\min} + dx \\ |\hat{y}_c - y_c| \le y_{\max} - y_{\min} + dy \\ |\hat{z}_c - z_c| \le z_{\max} - z_{\min} + dz \end{array} \right\}, \quad (11)$$

where *dx*, *dy* and *dz* are positive numbers, and $dx < x_{min}$.

The nonlinear filter used in this paper is given by (12) (see also Fig. 6.

$$\begin{cases} \frac{d}{dt}(\hat{\bar{p}}) = -\bar{V}_m + \hat{\bar{V}}_{tg} + s \cdot K_1 \cdot {}^I_C R \cdot H^{-1}(\hat{\bar{p}}_c) \cdot (g_{\varphi\theta}(\hat{\bar{p}}_c) - \bar{y}_m) \\ \frac{d}{dt}\hat{\bar{V}}_{tg} = s \cdot K_2 \cdot {}^I_C R \cdot H^{-1}(\hat{\bar{p}}_c) \cdot (g_{\varphi\theta}(\hat{\bar{p}}_c) - \bar{y}_m) \\ \hat{\bar{p}}_c = {}^I_C R \cdot \hat{\bar{p}} \end{cases} \quad (12)$$

where *s* defines the out-of-frame event and $H(p_c)$ is the Jacobean of nonlinear transformation $g_{\varphi\theta}(p_c)$ with respect to p_c

$$H(p_c) = \begin{bmatrix} -\frac{f y_c}{x_c^2} & \frac{f}{x_c} & 0 \\ -\frac{f z_c}{x_c^2} & 0 & \frac{f}{x_c} \\ -\sin\theta & \cos\theta \sin\varphi & \cos\theta \cos\varphi \end{bmatrix}. \quad (13)$$

It is easy to check that $\det(H) = f^2 \frac{z}{x_c^3}$ and therefore $H(p_c)$ is always invertible for all admissible values of $\bar{p}_c$, φ, θ except if altitude *z*=0.

The filtering solution (12) extends results proposed in [2] to include out-of-frame events. Theorem 1 in [3] can be used to prove regional stability of the filter (12) for the process model (8) with the regions P_c and $\hat{P}_c$ given by (10) and (11) in the presence of brief out-of-frame events characterized by the parameters T_0 and α. The proof follows directly from the one used in [3] and is therefore omitted.

Figure 6 shows implementation of the filter (12). When the out-of-frame event occurs, the filter integrates the velocity measurements to obtain an estimate of the relative position (dead reckoning). When target tracking is reestablished the integrators are reinitialized based on the real-time imagery.

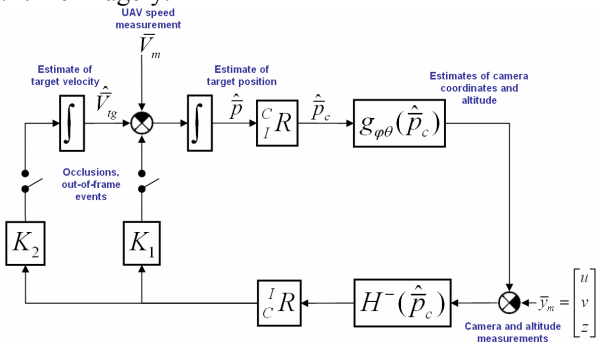


Fig. 6. Implementation of filter (12).

Next, the entire system including the control law (2) and the filter (12) was tested in a full scale 6DOF nonlinear simulation in the presence of wind and measurement noise. Scenario used for simulation assumed identification of a moving target and start of target tracking at 2.5 sec after beginning of flight, then initialization of position estimation filters at 26 sec when the object of interest is at 50° starboard. Between 2.5 and 26 seconds interval the UAV experiences transient of the control law that brings the UAV to a circular motion around the moving target. Target is moving with constant ground speed of 14 m/s and heading 45°. Based on the analysis of real measurements from numerous flight experiments the following sensors noise were applied to simulation: camera noise for both channels with 0° mean and 2.5° variance, measurements of altitude above the target with 0m mean and 20m variance (here we assumed worst case scenario when only GPS measurements are available and target is moving on a flat ground at known altitude MSL).

The result of this simulation for the ideal case when no out-of-frame events occur ($\alpha = 0$) are presented next. Figure 7 shows 3D and plane projections of the target, UAV trajectories and the projection of the estimated target position obtained with filter (12). The filter is initialized with the horizontal coordinates of the UAV but with the altitude of the target.

Analysis shows that except for the very short convergence interval the estimated target position closely follows true motion of the target. Figure 8 represents the filtering results for position, speed and heading estimation errors. It can be seen that in an ideal scenario with $\alpha = 0$ the convergence time for the positional error (see Fig.8.a – shows convergence to 10 m) does not exceed 5.5 seconds and 11 seconds for both speed and heading (see Fig.8.b – shows convergence to 5 m/s and 5°).

Analysis of the same experiment with a variable target loss parameter α is presented in Fig.9. The metrics used to evaluate performance of the filter as α increases were

chosen to be the speed of convergence parameters. Specifically, these were defined to be the 1st time instance pass which the estimate stays within 10% of the true value. Here represents the position metric and V_{conv} – the velocity metric.

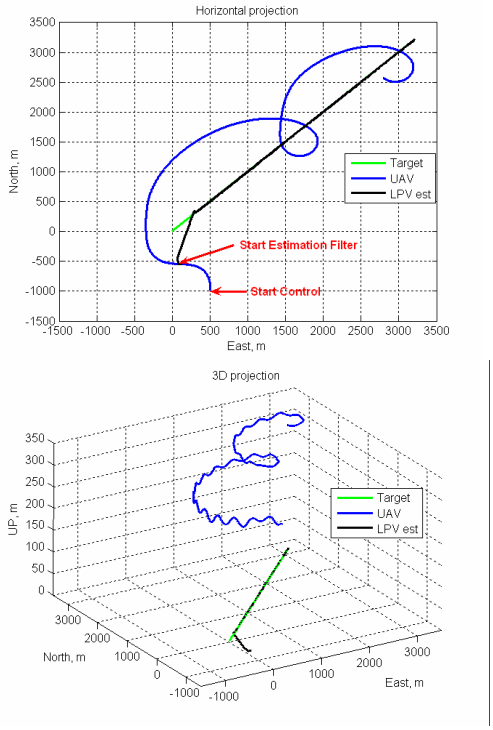
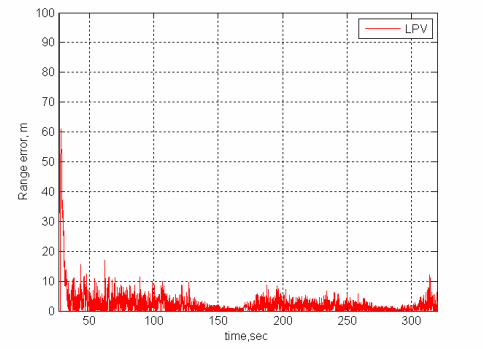
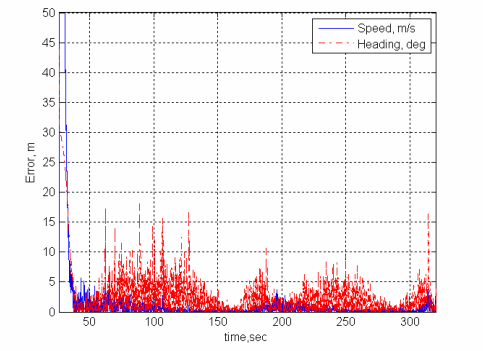


Fig. 7. 3D and 2D projections of relative motion.



a. Position error.



a. Velocity and Heading errors.

Fig. 8. Convergence results for filter (12).

The analysis shows that the filter exhibits stable convergence times for both position and velocity estimates in the presence of out-of-frame events characterized by α as high as 0.45 (the target is lost 45% of the time). The positional convergence time P_{conv} for the nonlinear filter (NLF) reported earlier [1] is also included in Fig. 9. It is obvious that the filter (12) outperforms the NLF filter for the entire range of values of α considered.

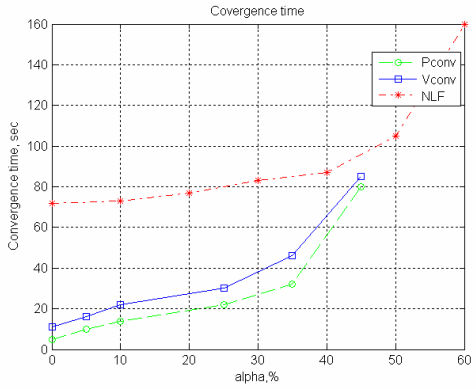


Fig. 9. Convergence time vs. variable α .

IV. Flight Test Results

The flight test setup for to test the filter (12) is almost identical to the one described in [1] and is shown in Fig.10. A customized Senior Telemaster model aircraft was used to house the gimbaled camera, wireless video and serial links as well as Piccolo autopilot [6] with its dedicated control link. The image obtained by the onboard camera was broadcast on a 2.4 GHz link and processed on the ground by off-the-shelf PerceptiVU image processing software [7].

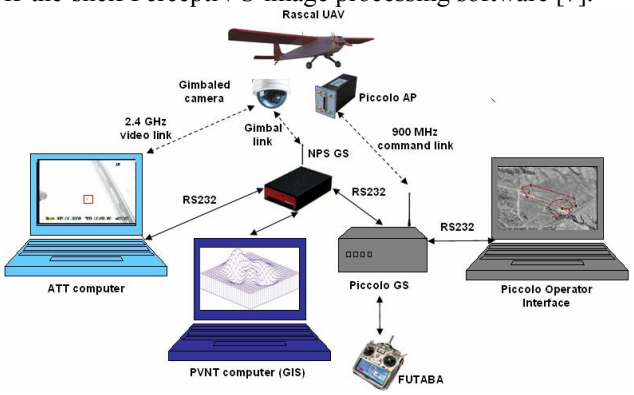


Fig. 10. Flight test setup.

PerceptiVU allows the user to select and lock on a target displayed on a ground station screen. In the configuration used in this experiment, PerceptiVU provides coordinates of the centroid of the target selected by the user. These coordinates were then employed by the control and filtering algorithms introduced in previous sections that were implemented on the NPS ground station (GS).

Multiple flight tests of the complete system were conducted in February-May and August-September of 2005. This time, rather than being fixed the target (white minivan) was moving along side of the runway with fixed speed of 4-

5 m/s and heading 296° (parallel to the runway) (see Fig.11). In order to evaluate the system performance the position, direction and speed of the target were continuously tracked by a GPS receiver.

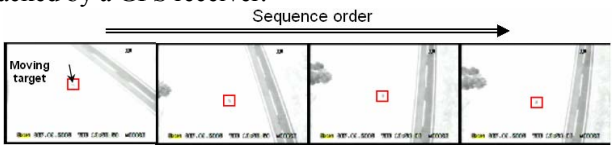


Fig. 11. An example of visual tracking

Results of the tracking are summarized in Figs.12 and Fig.13. For the sake of comparison they represent implementation of two estimation algorithms. Figure 12 includes a 3D plot of the UAV trajectory at the top as well as the estimates of the target position at the bottom. The UAV trajectory is color coded to display the time intervals where the target track was lost. Due to low speed of the target, the control law maintains a circular motion with the turn radius of about 200m and a slowly moving center as predicted by the analysis presented in Section II.

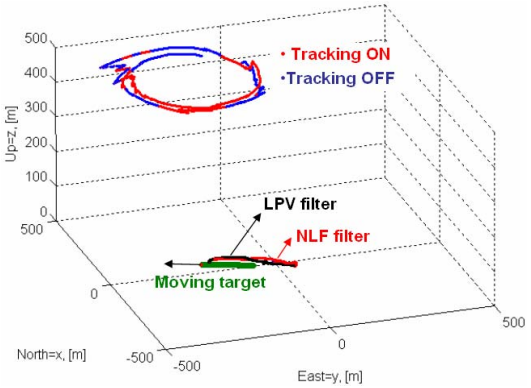


Fig. 12. Flight test result of tracking a moving target

Figure 13 shows range and Fig.14 velocity estimation errors. Superimposed on the position estimation error plot is the time history of the tracking loss events.

As can be seen from Fig. 13 the filter (12) performs significantly better than the NLF filter, while the velocity estimation error obtained with the filter (12) does not exceed 0.5 m/s.

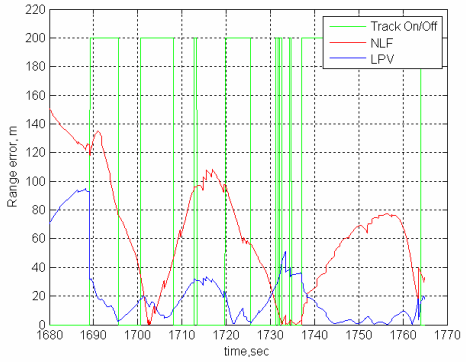


Fig. 13. Flight test range estimation errors for two algorithms.

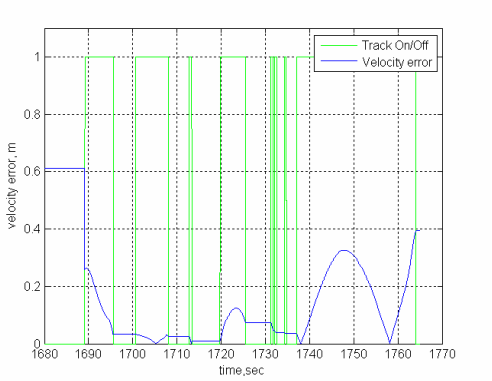


Fig. 14. Flight test velocity estimation error.

V. Conclusions

A system capable of tracking a moving target and estimating its position and velocity was developed. Straightforward nonlinear analysis was used to motivate a simple control system for integrated control of a UAV and of an onboard gimbaled camera. The control system was shown to perform well in both nonlinear simulation and in flight tests. Furthermore, a nonlinear LPV filter for target motion estimation was introduced. The filter performance was analyzed in the presence of target loss events. It was shown that the filter exhibited graceful degradation of performance in the presence of these events. The flight test results for moving target supported this conclusion. Future work will address improving performance of the target tracking and motion estimation algorithms by decreasing convergence times, reducing occurrence of target loss events and of their impact on the filter performance.

References

[1] Whang I.H., Dobrokhodov V.N., Kaminer I.I., Jones K.D., "On Vision-Based Target Tracking and Range Estimation for Small UAVs," Proceedings of AIAA Guidance, Navigation, and Control Conference, San Francisco, CA, August 15-18, 2005.

[2] Olivera P., Pascoal A., Kaminer I.I, "A nonlinear Vision Based Tracking System for Coordinated Control of Marine Vehicles," *Proceedings of the 10th Mediterranean Conference on Control and Automation – MED 2002,* Lisbon, Portugal, July 9-12, 2002.

[3] Hespanha J., Yakimenko O., Kaminer I., Pascoal A., "Linear Parametrically Varying Systems with Brief Instabilities: An Application to Integrated Vision / IMU Navigation", *IEEE Transactions on Aerospace and Electronic Systems Technology*, July 2004.

[4] V. Dobrokhodov, I. Kaminer, K. Jones and R. Ghabcheloo, "Vision Based Tracking and Motion Estimation for Moving Targets using Small UAVs," NPS Internal Report, January 2006.

[5] Baer, W., "UAV Target Mensuration Experiment Using Synthetic Images from High Resolution Terrain Databases at Camp Roberts," 72nd MORSS, 10-12 June 2004, NPS Monterey, CA ,WG 25 T&E, web site at http://www.trac.nps.navy.mil/PVNT/

[6] Piccolo/Piccolo Plus autopilots - A highly integrated autopilots for small UAVs, Cloud Cap Technology, Inc., http://cloudcaptech.com/.

[7] The PerceptiVU Target Tracking Software manual, PerceptiVU Inc, www. PerceptiVU.com .