

Vision-based Target Tracking with a Small UAV: Optimization-based Control Strategies

Steven A. P. Quintero^{a,1,*}, João P. Hespanha^{a,1}

^a*Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, U.S.A.*

Abstract

This paper considers the problem of a small, fixed-wing UAV equipped with a gimbaled camera autonomously tracking an unpredictable moving ground vehicle. Thus, the UAV must maintain close proximity to the ground target and simultaneously keep the target in its camera's visibility region. To achieve this objective robustly, two novel optimization-based control strategies are developed. The first assumes an evasive target motion while the second assumes a stochastic target motion. The resulting optimal control policies have been successfully flight tested, thereby demonstrating the efficacy of both approaches in a real-world implementation and highlighting the advantages of one approach over the other.

Keywords: unmanned aerial vehicle, target tracking, motion planning, autonomous vehicle

1. Introduction

The purpose of this work is to detail the design of two different control policies that enable a small, fixed-wing unmanned aerial vehicle (UAV), equipped with a pan-tilt gimbaled camera, to autonomously track a moving ground vehicle (target). The specific control objective is best described by Saunders in §4.1 of [1], where he defines *vision-based target tracking* as “maintaining a target in the field-of-view of an onboard camera with sufficient resolution for visual detection.”

The class of UAV platforms under consideration, are small, fixed-wing aircraft that are battery-powered and either hand or catapult launched. Specific to the UAV used in the flights experiments is a mechanical limitation of the pan-tilt gimbal mechanism that requires the UAV to keep the target towards its left-hand side for visibility. Nonetheless, by adjusting the cost function of the dynamic optimization, this work can be adapted to the fixed-camera scenario that is common on smaller platforms such as Micro Air Vehicles (MAVs). The sensor visibility constraint coupled with uncertain target motion and underactuated UAV dynamics compose the control challenge for which two novel solutions are presented.

1.1. Contributions

Two different styles of optimization-based control policies are developed to enable a small UAV to maintain visibility and proximity to target in spite of sensor blind spots,

underactuated dynamics, and evasive or stochastic non-holonomic target motion. The first is a game theoretic approach that assumes evasive target motion. Hence, the problem is formulated as a two-player, zero-sum game with perfect state feedback and simultaneous play. The second is a stochastic optimal control approach that assumes stochastic target motion. Accordingly, in this approach, the problem is treated in the framework of Markov Decision Processes (MDPs). In both problem formulations, the following are key features of the control design:

1. The UAV and the target are modeled by fourth-order discrete time dynamics, including simplified roll (bank) angle dynamics with the desired roll angle as the control input.
2. The UAV minimizes an expected cumulative cost over a finite horizon.
3. The cost function favors good viewing geometry, i.e., visibility and proximity to the target, with modest pan-tilt gimbal angles.
4. The dynamic optimization is solved offline via dynamic programming.

Both approaches incorporate roll dynamics because the roll dynamics can be on the same time scale as the heading dynamics, even for small (hand-launched) UAVs. Accordingly, this work directly addresses the phase lag in the heading angle introduced by a comparatively slow roll rate that would otherwise be detrimental to the UAV's tracking performance. Additionally, for small aircraft, the range of permissible airspeeds may very limited, as noted in [2], while frequent changes in airspeed may be either undesirable for fuel economy or unattainable for underactuated aircraft. Thus, both control approaches assume a constant airspeed and treat the desired roll angle of the

*Corresponding author

Email addresses: quintero@ece.ucsb.edu (Steven A. P. Quintero), hespanha@ece.ucsb.edu (João P. Hespanha)

¹Authors are with UCSB's Center for Control, Dynamical Systems, and Computation

aircraft as the sole control input that affects the horizontal plant dynamics. The target is modeled as a nonholonomic vehicle that turns and accelerates.

In order to determine control policies (decision rules for the desired roll angle) that facilitate good viewing geometry, a cost function is introduced to penalize extreme pan-tilt angles as well as distance from the target. Dynamic programming is used to compute (offline) the optimal control policies that minimize the expected cumulative cost over a finite planning horizon. The control policies are effectively lookup tables for any given UAV/target configuration, and hence they are well suited for embedded implementations onboard small UAVs. These control policies have been successfully flight tested on hardware in the field, thereby verifying their robustness to unpredictable target motion, unmodeled system dynamics, and environmental disturbances. Lastly, although steady wind is not directly addressed in the problem formulation, high fidelity simulations were performed that both verify and quantify the policies' inherent robustness to light and even moderate winds as well.

1.2. Related Work

Significant attention has been given to the target tracking problem in the past decade. Research groups have approached this problem from several different vantage points, and hence notable work from these perspectives is now highlighted. One line of research proposes designing periodic reference flight trajectories that enable the UAV to maintain close proximity to the target as it tracks the reference trajectories via waypoint navigation [3] or good helmsman steering [4]. Although one reference trajectory is typically not suitable for all target speeds, one can optimally switch between them based on UAV-to-target speed in order to minimize the maximum deviation from the target [5]. A particularly unique line of work on target tracking is that of oscillatory control of a fixed-speed UAV. In this approach, one controls the amplitude and phase of a sinusoidal heading-rate input to a kinematic unicycle such that the average velocity along the direction of motion equals that of the ground target, which is assumed to be piecewise constant [6, 7]. None of the preceding works, however, consider any limitations imposed by miniature vision sensors that are common on small, inexpensive UAVs.

Perhaps the greatest amount of research in the general area of target tracking is devoted to solving the specific problem of *standoff tracking*. The control objective for this problem is to have a UAV orbit a moving target at a fixed, planar standoff distance. Frew et al. use a Lyapunov guidance vector field (LGVF) approach to achieve the said objective [8]. Summers builds upon the work of Frew and further assumes the target and wind speeds are unknown. Using the LGVF approach and adaptive estimates of the combined wind and target velocity, Summers determines control laws that provably achieve the desired standoff tracking objective along with a conservative up-

per bound on the maximum combined wind/target speed that can be tolerated due to airspeed limitations [9].

Whereas the previous two lines of work did not consider sensor limitations, there is a line of work, based on nonlinear feedback control of the UAV's heading rate, wherein vision-sensor requirements are addressed. Dobrokhodov et al. develop control laws for controlling both a UAV and its camera gimbal [10]. The authors design nonlinear control laws to align the gimbal pan angle with the target line-of-sight (LOS) vector and the UAV heading with the vector tangent to the LOS vector; however, only uniform ultimate boundedness is proved. Li et al. advance the previous work by reformulating the control objective, adapting the original guidance law, and proving asymptotic stability of the resultant closed-loop, non-autonomous system [11]. The authors of [11] further adapt this newly designed control law to achieve asymptotic stability for the case of time-varying target velocity, although it comes at the high cost of requiring airspeed control as well as data that is nontrivially acquired, namely the target's turn rate and acceleration. Saunders and Beard consider using a fixed-camera MAV to perform vision-based target tracking [12]. By devising appropriate nonlinear feedback control laws, they are able to minimize the standoff distance to a constant-velocity target, while simultaneously respecting field of view (FOV) and max roll angle constraints.

Anderson and Milutinović present an innovative approach to the standoff tracking problem by solving the problem using stochastic optimal control [13]. Modeling the target as a Brownian particle (and the UAV as a deterministic Dubins vehicle), the authors employ specialized value iteration techniques to minimize the expected cost of the total squared distance error discounted over an infinite horizon. As no penalty is imposed on the control value, the resulting optimal control policy is a bang-bang turn-rate controller that is highly robust to unpredictable target motion. However, the discontinuous turn rate and absence of sensor limitations render the control policy infeasible in a real-world implementation.

Others have also employed optimal control to address the general target tracking problem, wherein the optimization criterion is mean-squared tracking error. Ponda et al. consider the problem of optimizing trajectories for a single UAV equipped with a bearings-only sensor to estimate and track both fixed and moving targets [14]. By performing a constrained optimization that minimizes the trace of the Cramer-Rao Lower Bound at each discrete time step, they show that the UAV tends to spiral towards the target in order to increase the angular separation between measurements while simultaneously reducing its distance to the target. While Ponda's approach is myopic, i.e., no lookahead, and controls are based on the true target position, Miller et al. propose a non-myopic solution that selects the control input based on the probability distribution of the target state, where the distribution is updated by a Kalman filter that assumes a nearly constant velocity target model [15]. Moreover, Miller poses the target

tracking problem as a partially observable Markov decision process (POMDP) and presents a new approximate solution, as nontrivial POMDP problems are intractable to solve exactly [16].

While the present paper focuses on target tracking with a single UAV, there is also notable work devoted to coordinating two or more UAVs to perform cooperative target geolocation (localization) and tracking. Such work generally aims at achieving coordinated standoff tracking, wherein two UAVs coordinate the phase of their fixed-radius orbits about the target in order to improve geolocation estimates and/or obtain a diversity of viewing angles. For example, in [17], Oh et al. propose a tangent vector field guidance strategy for coordinated standoff tracking with multiple UAVs wherein they employ sliding mode control with adaptive terms in order to mitigate the effects of disturbances and unmodelled dynamics. With estimates of the target state from a decentralized information filter, the UAVs achieve a desired angular separation in a decentralized fashion under various information / communication architectures by changing either airspeed or the nominal orbit radius. The problem has been approached from alternative vantage points in [18–20]. While such work typically uses nonlinear feedback control, other have used optimal control for the flight coordination of two UAVs. For example, Stachura et al. develop an online receding horizon controller in [21] to maximize the total geolocation information in the presence of communication packet losses, and in [22] the authors determine optimal control policies offline that minimize the joint geolocation error over a long planning horizon.

Although the present paper focuses on tracking a single target with one UAV, we note that tracking multiple ground targets with multiple UAVs has received considerable attention as well. In [23], Oh et al. use Frew’s LGVF approach to achieve coordinated standoff tracking of *clusters* of targets using multiple UAV teams. Within each group, the nominal orbit radius is set to ensure that each UAV maintains visibility of all targets in the corresponding cluster, as each agent has a limited FOV. Furthermore, the UAV groups follow a set of rules to discard stray target vehicles as well as ensure successful target handoff between teams. In [15], Miller et al. treat the problem of tracking multiple targets with multiple UAVs as a POMDP, wherein heuristics are employed in the approximate solution to overcome the limitations of short planning horizons in the presence of occlusions. Additional work on the subject of multi-target tracking can be found in the references of the aforementioned works.

None of the preceding works have considered a target that performs evasive maneuvers to escape the camera’s FOV, yet similar problems have been addressed long ago in the context of differential games [24]. In particular, Dobbie characterized the surveillance-evasion game in which a variable-speed pursuer with limited turn radius strives to keep a constant-speed evader within a specified surveillance region [25]. Lewin and Breakwell extend this work

to a similar surveillance-evasion game wherein the evader strives to escape in minimum time, if possible [26]. While the ground target may not be evasive, treating the problem in this fashion will produce a UAV control policy robust to unpredictable changes in target velocity.

In all of the preceding works, at least one or more assumptions are made that impose practical limitations. Namely, the works mentioned thus far assume at least one of the following:

1. Input dynamics are first order, which implies that roll dynamics have been ignored.
2. Changing airspeed quickly / reliably is both acceptable and attainable
3. Target travels at a constant velocity.
4. Sensor is omnidirectional.
5. Sinusoidal/orbital trajectories are optimal, including those resulting from standoff tracking.

The work presented here removes all of these assumptions to promote a practical, robust solution that can be readily adapted to other similar target tracking applications that may have different dynamics and hardware constraints. The policies also possess an inherent robustness that allow them to even track an unpredictable target in the presence of light to moderate steady winds.

1.3. Paper Outline

The remainder of the paper is organized as follows. Sections 2 and 3 detail the game theoretic and stochastic optimal control approaches to vision-based target tracking, respectively. These sections discuss the specific UAV and target dynamical models, the common cost function, and the individual dynamic programming solutions. Section 4 describes the experimental hardware setup and also presents the flight test results for each control approach. Furthermore, this section also provides a quantitative comparison of the two approaches and draws conclusions concerning the preferred control approach. Section 4 concludes by studying the effects of wind on the performance of the policies in a high fidelity simulation environment to quantify practical upper limits on the wind speeds that can be tolerated. Finally, Section 5 provides conclusions of the overall work and discusses venues for future work.

2. Game Theoretic Control Design

This section details the game theoretic approach to vision-based target tracking. The key motivations for this approach are to remedy the usual constant-velocity target assumption seen in much of the literature and also to account for sensor visibility limitations. This is done by assuming the target performs evasive maneuvers, i.e., it strives to enter the sensor blind spots of the UAV according to some control policy optimized to play against that of the UAV. Accordingly, the problem is posed as a multi-stage, two-player, zero-sum game with simultaneous play

and solved with tools from noncooperative game theory. The two main elements of a game are the actions available to the players and their associated cost. Thus, the players' actions at each stage are first described, along with their respective dynamics. The cost function of the viewing geometry is presented next and is the same as that used in the stochastic approach. Lastly, this section presents the formal problem statement and the dynamic programming solution that generates a control policy for each player.

2.1. Game Dynamics

While the majority of work on target tracking uses continuous time motion models, this work treats the optimization in discrete time. Thus, each player is initially modeled by fourth-order continuous-time dynamics, and then a T -second zero-order hold (ZOH) is applied to both sets of dynamical equations to arrive at the discrete-time dynamics of the overall system.

The UAV is assumed to have an autopilot that regulates pitch, airspeed, altitude, and commanded roll angle via internal feedback loops, typically using Proportional-Integral-Derivative (PID) control. Once every T seconds, the roll-angle setpoint is updated according to the control policy for the game. The aircraft is further assumed to fly at a fixed airspeed and constant altitude above the ground plane. Hence, the UAV's state comprises its planar position $(x_a, y_a) \in \mathbb{R}^2$, heading $\psi_a \in \mathbb{S}^1$, and roll (bank) angle $\phi \in \mathbb{S}^1$. Thus, the UAV state is defined as $\xi := (x_a, y_a, \psi_a, \phi)$.

The UAV is modeled as a planar kinematic unicycle with second-order rotational dynamics, i.e.,

$$\dot{\xi} = \frac{d}{dt} \begin{pmatrix} x_a \\ y_a \\ \psi_a \\ \phi \end{pmatrix} = \begin{pmatrix} s \cos \psi_a \\ s \sin \psi_a \\ -(\alpha_g/s) \tan \phi \\ -\alpha_\phi(\phi - r) \end{pmatrix}, \quad (1)$$

where $\alpha_g > 0$ is the acceleration due to gravity and s denotes the airspeed, which is equivalent to the aircraft's groundspeed since we presently assume zero wind. Nonzero wind is addressed in Section 4.1. Furthermore, $1/\alpha_\phi > 0$ is the time constant corresponding to the autopilot control loop that regulates the actual roll ϕ to the roll-angle setpoint r . Most work in this area assumes that α_ϕ is large enough so that there is a separation of time scales between the heading dynamics $\dot{\psi}_a$ and the controlled roll dynamics $\dot{\phi}$, and consequently, the roll dynamics can be ignored. However, when this assumption does not hold, the resultant phase lag introduced into the system can prove detrimental to target tracking performance. This is the case for small UAVs, like the one used in the experimental work of the present paper, and hence such a simplifying assumption was avoided.

Applying a T -second ZOH to the UAV subsystem produces a discrete-time system $\xi^+ = \mathbf{f}_1(\xi, r)$ that is obtained by solving the system of differential equations (1). The roll

command r is assumed to belong to the set

$$C = \{0, \pm\Delta, \pm 2\Delta\}, \quad (2)$$

where $\Delta > 0$ is the maximum allowable change in the roll setpoint from one ZOH period to the next. Denoting the roll command over the next ZOH period by r' , this work further stipulates that $r' \in U(r)$, where

$$U(r) := \{r, r \pm \Delta\} \cap C \quad (3)$$

is the *roll-angle action space*. This allows roll commands to change by at most Δ and avoids sharp changes in roll that would be detrimental to image processing algorithms in the target tracking task [2].

The target is assumed to be a nonholonomic vehicle that travels in the ground plane and has the ability to turn and accelerate. Its state comprises its planar position $(x_g, y_g) \in \mathbb{R}^2$, heading $\psi_g \in \mathbb{S}^1$, and speed $v \in \mathbb{R}_{\geq 0}$ and is hence defined as $\eta := (x_g, y_g, \psi_g, v)$. The target's dynamics are those of a planar kinematic unicycle, i.e.,

$$\dot{\eta} = \frac{d}{dt} \begin{pmatrix} x_g \\ y_g \\ \psi_g \\ v \end{pmatrix} = \begin{pmatrix} v \cos \psi_g \\ v \sin \psi_g \\ \omega \\ a \end{pmatrix}, \quad (4)$$

where ω and a are the turn-rate and acceleration control inputs, respectively.

Applying a T -second ZOH to the target subsystem produces a discrete-time system $\eta^+ = \mathbf{f}_2(\eta, \mathbf{d})$ that is obtained by solving (4) with $\mathbf{d} := (\omega, a)$. To describe the target's action space, the set of admissible target speeds W is first introduced, along with the target's maximum speed \bar{v} , which is assumed to be less than the UAV's airspeed. Denoting the target's maximum acceleration by \bar{a} , the target's acceleration a is assumed to belong to the following set:

$$D_a(v) := \begin{cases} \{0, \bar{a}\}, & v = 0 \\ \{-\bar{a}, 0, \bar{a}\}, & v \in W \setminus \{0, \bar{v}\} \\ \{-\bar{a}, 0\}, & v = \bar{v}. \end{cases} \quad (5)$$

Furthermore, with $W = \{0, \bar{a}T, 2\bar{a}T, \dots, \bar{v}\}$, the set W is invariant in the sense that $v \in W$ and $a \in D_a(v)$ implies $v^+ \in W$, as $v^+ = aT + v$. This property not only enforces speed bounds, but also improves the accuracy of the solution to the dynamic optimization. Denoting the target's maximum turn rate by $\bar{\omega}$, the target's turn-rate ω is assumed to take on values in the following set:

$$D_\omega(v) := \begin{cases} \{0\} & , v \in \{0, \bar{v}\} \\ \{-\bar{\omega}, 0, \bar{\omega}\} & , v \in W \setminus \{0, \bar{v}\} \end{cases}. \quad (6)$$

This restriction on the turn rate implies that the target vehicle cannot turn while stopped nor while traveling at its maximum speed \bar{v} . Otherwise, it has the ability to turn

left, go straight, or turn right using its maximum turn rate. The target's overall action space is defined as

$$D(v) := D_\omega(v) \times D_a(v), \quad (7)$$

and hence $\mathbf{d} \in D(v)$. Accordingly, depending on its current speed v , the target has anywhere from 2 – 9 action pairs from which to choose at a given stage of the game.

The overall state of the game, $\zeta \in \mathcal{Z}$, is defined as

$$\mathcal{Z} := \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1 \times \mathbb{R}_{\geq 0}, \quad (8)$$

which combines the UAV and target states. The first 3 components of ζ are relative quantities in a target-centric coordinate frame, and the rest are absolute. The relative planar position in the target-centric coordinate frame is

$$\begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} = \begin{bmatrix} \cos \psi_g & \sin \psi_g \\ -\sin \psi_g & \cos \psi_g \end{bmatrix} \begin{bmatrix} x_a - x_g \\ y_a - y_g \end{bmatrix} \quad (9)$$

while the remaining states are as follows:

$$(\zeta_3, \zeta_4, \zeta_5) = (\psi_a - \psi_g, \phi, v). \quad (10)$$

The overall dynamics of the game, $\zeta^+ = \mathbf{f}(\zeta, r, \mathbf{d})$, are given implicitly by $\mathbf{f}_1(\xi, r)$ and $\mathbf{f}_2(\eta, \mathbf{d})$ and the preceding transformations of the states in (9) and (10).

2.2. Cost Objective

Small, inexpensive UAVs performing vision-based target tracking commonly carry miniature pan-tilt gimbal mechanisms that have limited sensing regions similar to the one depicted in Figure 1. The most prominent feature of this diagram is that there is a large blind spot extending from the right side of the UAV to its back, and hence a UAV with this particular field of regard would have to keep the target to its left for visibility.

Typical tilt angle limitations for a miniature pan-tilt gimbal mechanism are illustrated in Figure 2. While the mechanical tilt angle limitations shown in Figure 2 do not create blind spots in the down-looking direction, there are still reasons to avoid extreme tilt angles. In particular, a tilt angle close to 0° usually means that the UAV airframe is visible to the camera, which can block visibility of the target and/or generate false detections in image processing software. On the other hand, a tilt angle close to 90° results in unpredictable movement of the gimbal, as this represents a singularity point in the gimbal geometry, i.e., the pan angle is not unique [2].

Based on these sensing limitations, which this work treats as soft constraints, the game objective for the UAV will be to maintain a good viewing geometry with respect to the ground target while the target's objective is the opposite. This means that the UAV wants to be as close to the target as possible (to maximize the resolution of the target in the image plane and reduce the error in the vision-based measurements of the target's position [22]), while simultaneously avoiding extreme gimbal angles. This can be captured by a cost objective for the game to be

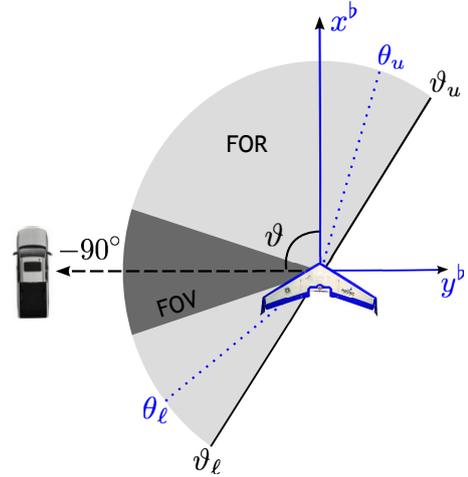


Figure 1: The camera's instantaneous *field of view* (FOV) and total *field of regard* (FOR) are indicated by the dark and light gray regions, respectively. The FOR is the total area visible to the camera as the gimbal is panned from its lower mechanical limit θ_ℓ to its upper mechanical limit θ_u . The “b” superscript on each axis denotes the UAV's local North-East-Down body frame. The azimuth angle of the line-of-sight vector to the target in this body frame is indicated by ϑ , and if it lies within the upper and lower FOR bounds, ϑ_u and ϑ_ℓ , respectively, then the target is in the UAV's field of regard. Although the camera's pan angle is not explicitly shown, it is assumed to equal the azimuth angle ϑ when $\vartheta \in [\theta_\ell, \theta_u]$ and otherwise be saturated at either θ_ℓ or θ_u .

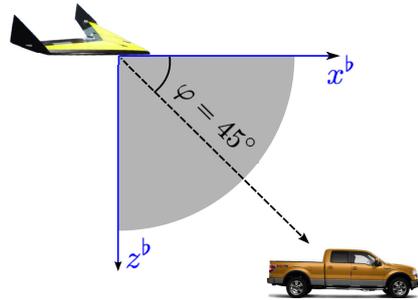


Figure 2: The range of tilt angles is indicated by the shaded region. The elevation angle of the line-of-sight vector to the target is denoted by φ and is measured in a positive sense from the (x^b, y^b) plane of the aircraft. For simplicity, the camera's tilt angle is taken to be the same as φ .

minimized by the UAV and maximized by the target given by

$$g(\zeta) = g_1(\vartheta) + g_2(\varphi) + g_3(\rho), \quad (11)$$

where ϑ is the azimuth angle in Figure 1, φ is the elevation angle in Figure 2, $\rho = \sqrt{\zeta_1^2 + \zeta_2^2}$, and

$$\begin{aligned} g_1(\vartheta) &= (\lambda_1 \max\{\vartheta - \bar{\vartheta}, 0, \vartheta - \bar{\vartheta}\})^2 \\ g_2(\varphi) &= \lambda_2^2 (\varphi - \pi/4)^2 \\ g_3(\rho) &= (\lambda_3 \rho)^2 \end{aligned} \quad (12)$$

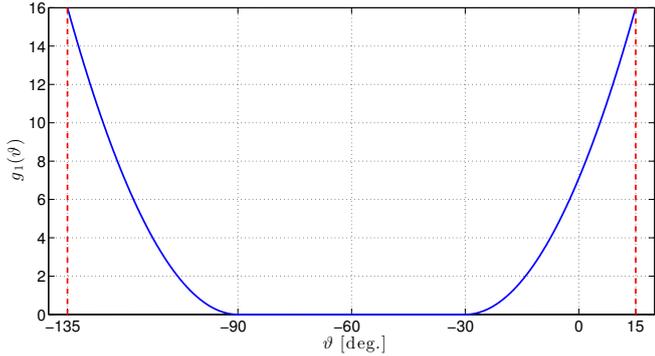


Figure 3: The azimuth cost function $g_1(\vartheta)$. In this particular instance, the pan angle limitations, θ_ℓ and θ_u , are indicated by the horizontal dashed lines at $\vartheta = -135^\circ$ and $\vartheta = 15^\circ$, respectively. The azimuth angle ϑ is penalized near these extremities with $\underline{\vartheta} = -90^\circ$, $\bar{\vartheta} = -30^\circ$, and $\lambda_1 = 16\pi$.

with $\lambda_i > 0$ and $\bar{\vartheta} \geq \underline{\vartheta}$. Note that in general, an elevation angle of $\varphi = \pi/4$ is achieved at some nonzero planar distance ρ , and hence there is a tradeoff between minimizing $g_2(\varphi)$ and minimizing $g_3(\rho)$. Also, since the elevation angle is measured with respect to the airframe, φ is typically not equal to $\pi/4$ when ρ equals the UAV's altitude, since the UAV often has a nonzero roll angle. For example, if the target is directly to the left of the UAV as depicted in Figure 1, then the planar distance that achieves $\varphi = \pi/4$ is $\rho = h_a / \tan(\pi/4 - \phi)$, where h_a denotes the UAV's altitude. Thus, if $\phi = -\pi/6$ and $h_a = 170$ [m], then $\rho = 45.55$ [m] achieves $\varphi = \pi/4$, and consequently the UAV can get quite close to the target while avoiding extreme gimbal angles. Moreover, there is no unique planar distance that achieves $\varphi = \pi/4$, and hence a dynamic optimization is advantageous in determining the optimal trajectories that minimize the UAV's planar distance to target as much as possible while simultaneously avoiding extreme gimbal angles.

An example of $g_1(\vartheta)$ is given in Figure 3. The plot depicts the zero-cost interval of azimuth angles as well as the quadratic penalization that occurs as the azimuth angles leave this range and approach the extremities. Since the actions available to each player and cost objective have been described, the gameplay setup and associated dynamic programming solution are presented next.

2.3. Game Setup and Objective

Although the dynamics of each agent were introduced in continuous-time, each agent chooses its control action at discrete time steps and applies the control action over the T -second ZOH period. Thus, the two-player, zero-sum dynamic game is played on a finite time interval of length K according to the discrete-time dynamics described in Section 2.1. At stage (time-step) $k \in \{0, \dots, K-1\}$, the UAV's action r_k belongs to the action space $U(\phi_k)$ defined in (3), and the target's control action pair \mathbf{d}_k belongs to the action space $D(v_k)$ defined in (7).

The game is played with a perfect state-feedback information structure, i.e., the players have access to the entire state, uncorrupted by noise, in order to decide on their actions. Furthermore, each player decides control actions according to a *behavioral policy*, which is a decision rule that associates to each state $\zeta \in \mathcal{Z}$ at stage k a probability distribution over the possible actions available to that player (see [27], Chapter 8). Therefore, when a player finds itself in a particular state $\zeta \in \mathcal{Z}$, it selects an action *randomly* according to the probability distribution given by the behavioral policy for ζ . The probability distributions over the UAV actions and target actions will belong to the following state-dependent probability simplexes:

$$\mathcal{A}(\phi) = \left\{ \boldsymbol{\alpha} \in \mathbb{R}^{m(\phi)} : \sum_i \alpha_i = 1, \alpha_i \geq 0 \quad \forall i \right\} \quad (13)$$

$$\mathcal{B}(v) = \left\{ \boldsymbol{\beta} \in \mathbb{R}^{n(v)} : \sum_j \beta_j = 1, \beta_j \geq 0 \quad \forall j \right\}, \quad (14)$$

respectively, where $m(\phi) = |U(\phi)|$ is the number of actions available to the UAV and $n(v) = |D(v)|$ is the number of action pairs available to the target. Accordingly, the behavioral policies for the UAV and target comprise time-dependent mappings $\gamma_k : \mathcal{Z} \rightarrow \mathcal{A}(\phi)$ and $\kappa_k : \mathcal{Z} \rightarrow \mathcal{B}(v)$, respectively, and the control actions are realizations of state-dependent probability distributions defined by the behavioral policies:

$$r_k \sim \gamma_k(\zeta[k]), \quad \mathbf{d}_k \sim \kappa_k(\zeta[k]), \quad \forall k \in \{0, \dots, K-1\},$$

where $\zeta[k]$ denotes the state ζ at stage k of the game.

A particular behavioral policy of the UAV is denoted by γ and belongs to the action space Γ_1 , which is the set of admissible behavioral policies for the UAV, i.e., the set of all length K sequences of mappings from \mathcal{Z} to $\mathcal{A}(\phi)$. Similarly, a particular behavioral policy for the target is denoted by κ and belongs to the action space Γ_2 , which is the set of admissible behavioral policies for the target, i.e., the set of all length K sequences of mappings from \mathcal{Z} to $\mathcal{B}(v)$.

Because the state of the game evolves stochastically, the function for the UAV to minimize and the target to maximize is

$$J(\gamma, \kappa) = E \left[\sum_{k=0}^{K-1} g(\zeta[k]) \middle| \zeta[0] \right],$$

where $\zeta[0] \in \mathcal{Z}$ and $E[\cdot]$ denotes expectation. Finally, the goal is to determine a saddle-point pair of behavioral policies (γ^*, κ^*) such that

$$J(\gamma^*, \kappa) \leq J(\gamma^*, \kappa^*) \leq J(\gamma, \kappa^*), \quad \forall \gamma \in \Gamma_1, \quad \forall \kappa \in \Gamma_2$$

for all $\zeta[0] \in \mathcal{Z}$. From the first condition, if the UAV is playing optimally with γ^* , the target (maximizer) can do no better (and may do worse) if it plays with any other admissible policy besides κ^* . And from the second condition,

if the target is playing optimally with κ^* , the UAV (minimizer) can do no better (and may do worse) if it plays with any other admissible policy besides γ^* . Thus, (γ^*, κ^*) constitutes a *saddle-point* pair of *equilibrium* policies from which no player will deviate, lest it do strictly worse.

2.4. Dynamic Programming Solution

Dynamic programming can be used to determine the optimal policies γ^*, κ^* and the corresponding value of the game $J(\gamma^*, \kappa^*)$ through *value iteration* as presented in Chapter 16 of [27]. This method centers on the *value function* $V_k(\zeta)$, which is also referred to as the *cost-to-go* from $\zeta \in \mathcal{Z}$ at time $k \in \{0, 1, \dots, K\}$. The value function is initialized at the final time as $V_K(\zeta) = g(\zeta)$. To determine the value function at the remaining times, an intermediate time and state-dependent matrix $\mathbf{G}(k, \zeta) \in \mathbb{R}^{m(\phi) \times n(v)}$ is first introduced and is given by

$$[\mathbf{G}_{ij}(k, \zeta)] := g(\zeta) + V_{k+1}(\mathbf{f}(\zeta, r^{(i)}, \mathbf{d}^{(j)})), \quad (15)$$

where $\mathbf{f}(\cdot)$ refers to the dynamics of Section 2.1, and $r^{(i)}$ and $\mathbf{d}^{(j)}$ are the i th and j th elements of the action spaces $U(\phi)$ and $D(v)$, respectively. For $k \in \{0, 1, \dots, K-1\}$, the cost-to-go is computed (offline) in reverse chronological order according to the following recursion

$$\begin{aligned} V_k(\zeta) &= \min_{\alpha \in \mathcal{A}(\phi)} \max_{\beta \in \mathcal{B}(v)} \alpha^\top \mathbf{G}(k, \zeta) \beta \\ &= \max_{\beta \in \mathcal{B}(v)} \min_{\alpha \in \mathcal{A}(\phi)} \alpha^\top \mathbf{G}(k, \zeta) \beta, \end{aligned} \quad (16)$$

where the second equality holds due to the Minimax Theorem (see [27], Chapter 5). As the optimizations in (16) are performed, the behavioral control policies are formed as

$$\gamma_k^*(\zeta) = \arg \min_{\alpha \in \mathcal{A}(\phi)} \max_{\beta \in \mathcal{B}(v)} \alpha^\top \mathbf{G}(k, \zeta) \beta \quad (17)$$

$$\kappa_k^*(\zeta) = \arg \max_{\beta \in \mathcal{B}(v)} \min_{\alpha \in \mathcal{A}(\phi)} \alpha^\top \mathbf{G}(k, \zeta) \beta, \quad (18)$$

and the procedure yields $J(\gamma^*, \kappa^*) = V_0(\zeta)$, $\forall \zeta \in \mathcal{Z}$.

For computational feasibility, the computation of the value function is limited to a finite number of points through the introduction of $Z \subset \mathcal{Z}$, which is a finite subset of the state space having N_s distinct elements, along with the quantization function $\mathbf{q} : \mathcal{Z} \rightarrow Z$, defined by

$$\mathbf{q}(\zeta) := \arg \min_{\mathbf{w} \in Z} \|\zeta - \mathbf{w}\|_1. \quad (19)$$

Computation of the value function is approximated by first redefining $\mathbf{G}(k, \zeta)$ as

$$[G_{ij}(k, \zeta)] := g(\zeta) + V_{k+1} \left(\mathbf{q}[\mathbf{f}(\zeta, r^{(i)}, \mathbf{d}^{(j)})] \right) \quad (20)$$

and then computing the value function according to (16) only for $\zeta \in Z$. Similarly, the behavioral policies are computed for $\zeta \in Z$ according to (17) and (18), but with $\mathbf{G}(k, \zeta)$ as defined in (20), instead of (15). In practice,

the control policies for $\zeta \in \mathcal{Z} \setminus Z$ at time k are evaluated using $\gamma_k^*(\mathbf{q}(\zeta))$ and $\kappa_k^*(\mathbf{q}(\zeta))$.

In general, determining (16) requires solving a linear program, which can be computationally burdensome for large N_s . However, when

$$\min_i \max_j G_{ij}(k, \zeta) = \max_j \min_i G_{ij}(k, \zeta), \quad (21)$$

where $i \in \{1, 2, \dots, m(\phi)\}$ and $j \in \{1, 2, \dots, n(v)\}$, then one does not have to solve a linear program, as a saddle-point equilibrium exists in pure policies, and one can set

$$\gamma_k^*(\zeta) = \mathbf{e}_i, \quad i := \arg \min_i \max_j G_{ij}(k, \zeta) \quad (22)$$

$$\kappa_k^*(\zeta) = \mathbf{e}_j, \quad j := \arg \max_j \min_i G_{ij}(k, \zeta), \quad (23)$$

where \mathbf{e}_i and \mathbf{e}_j are the i th and j th vectors of the canonical basis of $\mathbb{R}^{m(\phi)}$ and $\mathbb{R}^{n(v)}$, respectively.

The optimization for the hardware experiments was performed with the parameter data included in Tables 1, 2, and 3. Note that each state is uniformly sampled over a bounded interval (Table 3), and the discrete state space Z comprises the Cartesian product of the resulting discrete sets. This entails that no intuition regarding the nature of the optimal solution is exploited, e.g., discarding regions of the state space wherein the target is on the right-hand side of the UAV. Thus, the dynamic programming solution remains quite general, and, even for target tracking problems with no gimbal constraints, i.e., those in which $\lambda_1 = \lambda_2 = 0$ in (12), the computational effort remains the same.

To gain insight into the nature of the corresponding control policy, consider a projection of the state space onto the (ζ_1, ζ_2) -plane with $(\zeta_3, \zeta_4, \zeta_5) = (0, 0, 8)$ in Figure 4. Thus, this picture depicts the control actions for any position of the UAV relative to that of the target when the relative heading between the UAV and target is zero, the UAV's roll angle is zero, and the target is traveling at 8 m/s, which is just under half of the UAV's speed. The line of states in the first quadrant that corresponds to mixed control policies and $r = 0^\circ$ likely arises from the angular argument of $g_1(\vartheta)$ in (12) being wrapped to the interval $[-\pi, \pi)$. Otherwise, the first quadrant primarily depicts the UAV beginning a left turn ($r = -15^\circ$) in order to perform a counterclockwise loop around target. In the upper-leftmost parts of the first, second, and third quadrants, the UAV turns right ($r = +15^\circ$), to get behind the target and to its right. There is also a significant portion of the state space in the third quadrant for which the UAV simply continues its course ($r = 0^\circ$). Supposing the target holds its velocity and the UAV holds its course, the UAV will eventually cross over into the fourth quadrant due to the ground speed differential, wherein it begins turning left sooner if it is either too close or too far from the target. In any case, when $\zeta_1 \geq 56$ [m], the UAV needs to turn left in order to loop around the target in the counterclockwise direction. Lastly, Figure 4 illustrates the fact that a vast

Table 1: Optimization Parameters

Parameter:	ϑ	ϑ	λ_1	λ_2	λ_3	K
Value:	$-\pi/2$	$-\pi/6$	$16/\pi$	$20/\pi$	$3/200$	15

Table 2: Parameters in UAV and Target Dynamics

Parameter	Description	Value	Units
s	airspeed	18	m/s
α_g	gravitational acceleration	9.81	m/s^2
α_ϕ	inverse roll time constant	2	s^{-1}
-	altitude	170	m
Δ	max roll change	15	deg.
$\bar{\omega}$	max turn rate	$2\pi/24$	rad./s
\bar{v}	max target speed	14	m/s
T	zero-order hold period	2	s

Table 3: Sets for State Space Discretization

Set	Description	Value	Units
X	relative positions	$\{-192, -188, \dots, 192\}$	m
Ψ	relative headings	$\{0, 15, \dots, 345\}$	deg.
C	roll commands	$\{0, \pm 15, \pm 30\}$	deg.
W	target speeds	$\{0, 2, \dots, 14\}$	m/s
Z	discrete state space	$X^2 \times \Psi \times C \times W$	-

majority of the UAV control actions in the discrete state space Z are deterministic, as the white area corresponding to mixed policies is small. In fact, (21) holds for more than 95% of the states in Z , which greatly reduces the number of linear programs that need to be solved.

3. Stochastic Optimal Control Design

This section presents the stochastic optimal control approach to vision-based target tracking with a single small UAV, which is an attractive alternative to the computationally intensive game theoretic control design in Section 2. This approach utilizes more detailed models of the UAV and target to better characterize the actual dynamics encountered in a real-world implementation.

In Section 2, the UAV was modeled as a kinematic unicycle with first-order rotational dynamics. In practice, the UAV's roll dynamics are far more complex, and the UAV's airspeed typically experiences significant stochasticity. To capture the effects of both of these observed characteristics, stochasticity is introduced into the UAV model without adding additional states beyond those encountered in the game theoretic control design. This leads to a stochastic optimal control problem that is solved over a finite planning horizon using dynamic programming.

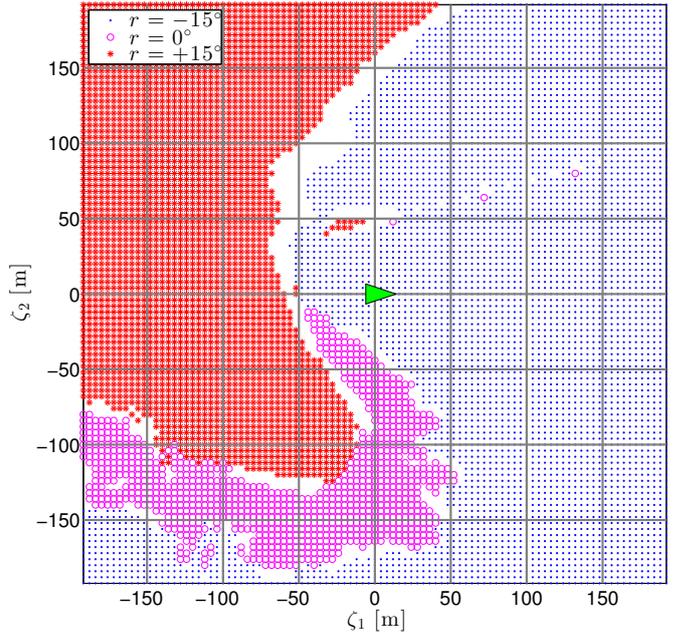


Figure 4: Control surface for game theoretic control policy projected onto the (ζ_1, ζ_2) plane with $(\zeta_3, \zeta_4, \zeta_5) = (0, 0, 8)$. Any points that appear to be missing are those for which the control action is not deterministic, i.e., the control policies are mixed and not pure.

3.1. Overview of Stochastic Dynamics

The UAV and target states are assumed to evolve stochastically according to discrete-time Markov Decision Processes, whose states are described in Section 2.1. Accordingly, the probability of transitioning from the current UAV state ξ to the next UAV state ξ' under the roll command r is given by the controlled state transition probability function $p_a(\xi' | \xi, r)$. Likewise, the probability of transitioning from the current target state η to the next target state η' is given by the state transition probability function $p_g(\eta' | \eta)$.

Rather than deriving explicit formulas for these state transition probabilities, empirical characterizations of the individual agent kinematics are developed that allow one to draw Monte Carlo samples ξ'_i and η'_i , $i \in \{1, 2, \dots, N_p\}$, from appropriate conditional probability density functions. To generate sample trajectories of the combined state ζ , one simply combines ξ'_i and η'_i according to equations (9) and (10). This provides an empirical characterization of the overall stochastic dynamics that evolve according to a controlled state transition probability function $p(\zeta' | \zeta, r)$. The ability to sample this state transition probability will suffice to effectively approximate the dynamic programming solution.

3.2. Stochastic UAV Kinematics

In practice, UAVs are subject to environmental disturbances, such as wind gusts, that introduce stochasticity into the dynamics. Although a real UAV's kinematics are

Table 4: Stochastic target motion parameters

Parameter:	σ_a	\bar{v}	$\bar{\sigma}$	σ_f	ν_1	ν_2	ϱ
Value:	0.5	18	0.5	0.1	1/8	1/2	7
Units:	m/s ²	m/s	rad./s	rad./s	-	-	m

$\phi_i(T, r)$. The samples thus have two sources of randomness: stochasticity in the roll-angle dynamics captured by the collection of roll-angle trajectories $\Phi(r)$ and stochasticity in the airspeed. Figure 7 illustrates the stochastic UAV model.

3.3. Stochastic Target Kinematics

To model the behavior of an operator driving the ground vehicle safely and casually, yet unpredictably, the same nonholonomic ground vehicle model given in Section 2.1 is used, where the continuous time dynamics are those given by (4) and each target control input is still held constant for T seconds. However, since evasive action is no longer assumed, control actions for the target are now drawn from continuous probability density functions that are independent of the UAV state.

While the target's acceleration is assumed to be a zero-mean Gaussian with fixed variance σ_a^2 , its turn rate is generated from a more complex model. Specifically, the target's turn rate begins as a zero-mean Gaussian random variable whose variance depends on the current target speed v according to $\sigma_\omega^2(v) = h^2(v/\bar{v})$, where \bar{v} is the maximum target speed and

$$h(v) = \begin{cases} (\bar{\sigma}/2) [1 + \cos(\pi(\nu - \nu_1)/\nu_1)] & 0 \leq \nu < \nu_1 \\ \bar{\sigma} & \nu_1 \leq \nu < \nu_2 \\ \frac{\bar{\sigma} - \sigma_f}{2} \left[1 + \cos\left(\pi \frac{\nu - \nu_2}{1 - \nu_2}\right) \right] + \sigma_f & \nu_2 \leq \nu < 1, \end{cases}$$

which is an asymmetric cosine-tapered window with maximum value $\bar{\sigma}$, final value σ_f , and parameters ν_1 and ν_2 describing the beginning and end points of the window with $0 < \nu_1 < \nu_2 < 1$. This is illustrated in Figure 8 with the parameters included in Table 4. The fact that $\sigma_\omega^2(0) = 0$ captures the assumption that the target is a nonholonomic vehicle that does not turn in place. Furthermore, as the target's speed increases from zero, it makes sharper turns more frequently, and there is even a range over which this frequency of sharp turns remains constant. Beyond this range, as its speed increases, the target makes softer turns in order to avoid losing traction/control. Thus, the normally distributed stochastic turn rate with speed-dependent variance models safe casual driving of the ground vehicle.

To add another degree of realism, the zero-mean Gaussian random variable with variance $\sigma_\omega^2(v) = h^2(v/\bar{v})$ used to generate the turn rate ω is "saturated" to enforce a minimum turn radius ϱ . Specifically, suppose that an acceleration sample a_i and turn rate sample $\tilde{\omega}_i$ are drawn

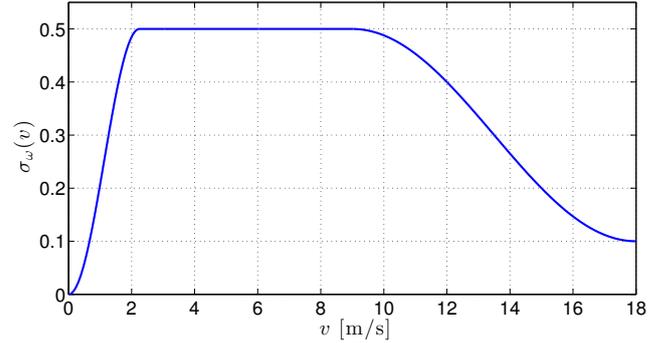


Figure 8: The speed dependent standard deviation of the normally distributed turn rate.

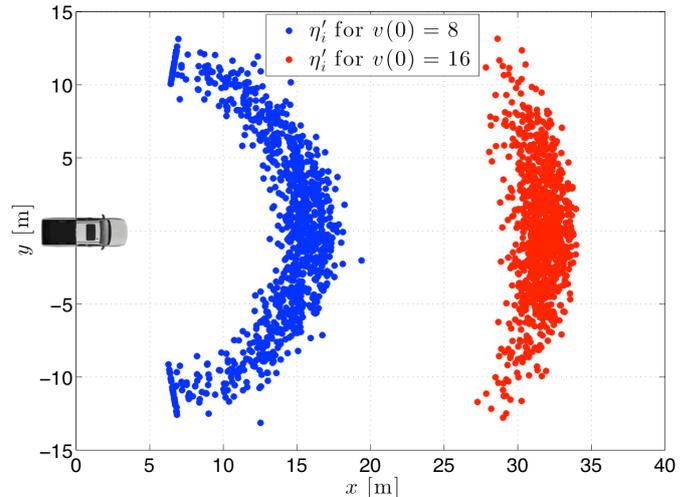


Figure 9: Sample trajectories generated from the stochastic target motion model. The two initial target states depicted with different colors correspond to identical initial positions at the origin, but two distinct initial speeds of 8 and 16 [m/s]. For each initial condition, $N_p = 1,000$ samples are generated.

from $\mathcal{N}(0, \sigma_a^2)$ and $\mathcal{N}(0, \sigma_\omega^2(v))$, respectively. The turn rate ω_i that is actually used is given by

$$\omega_i = \begin{cases} \bar{\omega}_i, & \tilde{\omega}_i > \bar{\omega}_i \\ \tilde{\omega}_i, & -\bar{\omega}_i \leq \tilde{\omega}_i \leq \bar{\omega}_i \\ -\bar{\omega}_i, & \tilde{\omega}_i < -\bar{\omega}_i \end{cases}$$

where $\bar{\omega}_i := \min\{v'_i/\varrho, v/\varrho\}$ is the maximum turn rate consistent with a minimum turn radius ϱ , and $v'_i = a_i T + v$ is the target speed at the end of the ZOH period. Figure 9 illustrates the stochastic target model with the parameters of Table 4 for two different initial conditions. In this figure, the scenario for the slower initial target speed has a final position distribution with a higher radius of curvature, which entails that the target makes sharper turns. The saturation on the tails of the turn-rate distribution becomes evident as well.

3.4. Control Objective and Dynamic Programming Solution

The stochastic optimal control problem is to determine the optimal control feedback control policy $\mu_k^* : \mathcal{Z} \rightarrow \mathcal{C}$, $k \in \{0, 1, \dots, K-1\}$, that minimizes

$$J(\zeta[0]) = E \left[\sum_{k=0}^K g(\zeta[k]) \middle| \zeta[0] \right], \quad \forall \zeta[0] \in \mathcal{Z}, \quad (24)$$

where $K \in \mathbb{N}$, $E[\cdot]$ denotes expectation, $g(\zeta)$ is given by (11), and the state ζ is a Markov Decision Process that evolves according to the transition probability $p(\zeta' | \zeta, r)$ determined by the models in Section 3.2 and Section 3.3, under the feedback law $r_k = \mu_k^*(\zeta)$.

Similar to the first control design, dynamic programming is used to minimize the criterion (24) using *value iteration* as presented in [16]. The method hinges on the *value function*, or *cost-to-go* from state $\zeta \in \mathcal{Z}$ at time $k \in \{0, 1, \dots, K-1\}$, which is defined as

$$V_k(\zeta) := g(\zeta) + \min_{r_k, r_{k+1}, \dots, r_{K-1}} E \left[\sum_{n=k+1}^K g(\zeta[n]) \right].$$

For $k = K$, one takes $V_K(\zeta) = g(\zeta)$, and the cost-to-go for $k \in \{0, 1, \dots, K-1\}$ is computed (offline) in reverse chronological order according to the following recursion

$$\begin{aligned} V_k(\zeta) &= g(\zeta) + \min_{r \in U(\phi)} E [V_{k+1}(\zeta') | \zeta, r] \\ &= g(\zeta) + \min_{r \in U(\phi)} \int V_{k+1}(\zeta') p(\zeta' | \zeta, r) d\zeta', \end{aligned} \quad (25)$$

which holds due to Bellman's principle of optimality (see [29], Chapter 6). As the minimization is performed, the optimal control policy is also formed as

$$\mu_k^*(\zeta) = \arg \min_{r \in U(\phi)} \left(g(\zeta) + E [V_{k+1}(\zeta') | \zeta, r] \right). \quad (26)$$

Performing the sequence of computations in (25) for $k \in \{0, 1, \dots, K-1\}$ ultimately yields $J^*(\zeta) = V_0(\zeta)$, $\forall \zeta \in \mathcal{Z}$, where $J^*(\zeta)$ is the minimum value of (24) under the feedback law (26). The main hurdle in computing (25) is the expectation, which is overcome in this work through a replacement with an empirical average computed using samples drawn according to the stochastic UAV and target models developed in Sections 3.2 and 3.3:

$$V_k(\zeta) \approx g(\zeta) + \min_{r \in U(\phi)} \frac{1}{N_p} \sum_{i=1}^{N_p} V_{k+1}(\zeta'_i). \quad (27)$$

As in Section 2.4, the computation of the value function is limited to a finite number of points in a set $Z \subset \mathcal{Z}$ that has N_s elements. Accordingly, with the quantization function $\mathbf{q}(\cdot)$ from (19), the computation of the value function and

optimal policy is approximated by

$$\begin{aligned} V_k(\zeta) &\approx g(\zeta) + \min_{r \in U(\phi)} \frac{1}{N_p} \sum_{i=1}^{N_p} V_{k+1}(\mathbf{q}(\zeta'_i)) \\ \mu_k^*(\zeta) &= \arg \min_{r \in U(\phi)} \left(g(\zeta) + \frac{1}{N_p} \sum_{i=1}^{N_p} V_{k+1}(\mathbf{q}(\zeta'_i)) \right), \end{aligned}$$

which only requires the evaluation of the value function over the finite set Z . In practice, the optimal roll command r_k for a state $\zeta \in Z$ is determined using $r_k = \mu_k^*(\mathbf{q}(\zeta))$.

Dynamic programming was performed for the hardware experiments with the parameters given in Tables 1, 4, 5, and 6, where $N_p = 1,000$ sample points were used to approximate the state transition probability. Note that, as in the solution for the game theoretic approach, each state is uniformly sampled over a bounded interval, and the overall discrete state space Z comprises the Cartesian product of the resulting discrete sets, as indicated in Table 6. Thus, no intuition into the nature of the optimal solution is exploited, and the dynamic programming solution to the stochastic optimal control problem remains quite general, requiring no additional computational effort even when there are no gimbal constraints, i.e., when $\lambda_1 = \lambda_2 = 0$ in (12).

To understand the nature of the resultant UAV control policy, consider a projection of the state space onto the (ζ_1, ζ_2) -plane with $(\zeta_3, \zeta_4, \zeta_5) = (0, 0, 8)$, which is shown in Figure 10. In the first quadrant, roughly below the line $\zeta_2 = (13/27)\zeta_1 + 117$ (dashed), the UAV begins turning left ($r = -15^\circ$) to perform a counterclockwise loop about the target, which is a costly maneuver as the distance will increase significantly during this act. However, if the UAV is above the said line, it rather begins turning right ($r = +15^\circ$) to move behind the target and to its right. The UAV does this same maneuver if it is in the left half of the second quadrant or the upper left corner of the third quadrant. Furthermore, the large region in the third quadrant wherein the UAV maintains a straight course ($r = 0^\circ$) extends neither below $\zeta_2 = -154$ nor to the right of $\zeta_1 = 0$, which is not true for the game theoretic control policy γ^* . Since the UAV turns left below $\zeta_2 = -154$, it is prevented from straying too far from the target. Supposing the UAV holds its course in the third quadrant and the target maintains its velocity, the UAV will cross into the fourth quadrant, where it begins turning left immediately to perform a counterclockwise turn about the target. This characteristic results in the UAV turning sooner than in the first control design and results in noticeably different behavior, which will become evident in the flight test results.

4. Hardware Setup and Flight Test Results

This section presents field test results of a single camera-equipped UAV performing vision-based target tracking with

Table 5: Parameters in Stochastic UAV dynamics

Parameter	Description	Value	Units
\bar{s}	nominal airspeed	18	m/s
σ_s^2	airspeed variance	16/25	m ² /s ²
α_g	gravitational acceleration	9.81	m/s ²
-	altitude	170	m
Δ	max roll change	15	deg.
T	zero-order hold period	2	s

Table 6: Sets for State Space Discretization

Set	Description	Value	Units
X	relative positions	$\{-252, -248, \dots, 252\}$	m
Ψ	relative headings	$\{0, 15, \dots, 345\}$	deg.
C	roll commands	$\{0, \pm 15, \pm 30\}$	deg.
W	target speeds	$\{0, 1, \dots, 17\}$	m/s
Z	discrete state space	$X^2 \times \Psi \times C \times W$	-

the control policies described in Sections 2 and 3. It begins with a description of the hardware setup for the field experiments whereby both the game theoretic and stochastic optimal control policies were tested. Next, it highlights representative flight test results to illustrate the behavior and performance of each control policy. Finally, the section concludes with a quantitative comparison of both control policies to objectively determine the preferred method for a real-world implementation.

4.1. Experimental Setup

A flight experiment was conducted in February, 2012 with a single UAV at the McMillan Airfield, Camp Roberts, CA. Toyon Research Corporation was responsible for launching, monitoring, and landing the UAV as well as driving the target ground vehicle.

A Unicorn/Zagi flying wing was used to test the vision-based target tracking control policies. This particular UAV platform is depicted in Figure 11 with its main payload, which is a gimbaled video camera. The particular pan and tilt limitations are those described in Figures 1 and 2, respectively, wherein the particular pan limits of the gimbal mechanism are $\theta_\ell = -135^\circ$ and $\theta_u = 15^\circ$. The corresponding FOR limits are $\vartheta_\ell = -152^\circ$ and $\vartheta_u = 32^\circ$.

The UAV's autopilot adjusts the throttle and two elevons to maintain pitch, airspeed, altitude, and the commanded roll angle. The particular ZOH time of $T = 2$ seconds was chosen to be roughly the settling time for the Unicorn's roll dynamics. The number of discrete time steps in each dynamic optimization was chosen to make the planning horizon ($KT = 30$ seconds) longer than the time the UAV would spend in an orbit around the target (≈ 20 seconds) at max bank. Furthermore, since both types of control policies were tested on the same sortie for a minimum of

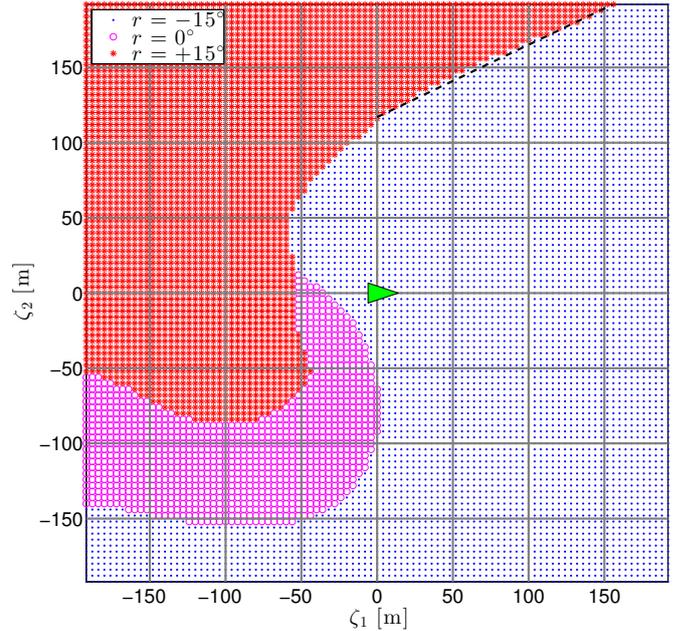


Figure 10: Control surface for stochastic optimal control policy projected onto the (ζ_1, ζ_2) plane with $(\zeta_3, \zeta_4, \zeta_5) = (0, 0, 8)$.



Figure 11: The Unicorn comprises a 60" expanded polypropylene foam wing that houses four lithium polymer batteries driving an electric motor attached to a push propeller.

15 minutes each, the limited planning horizon of 30 seconds was addressed by applying the control policies in a *receding horizon* fashion. This entails that, in the game theoretic case, $r_k \sim \gamma_0^*(\zeta)$, $\forall k \in \mathbb{Z}_{\geq 0}$, and in the stochastic optimal control case, $r_k = \mu_0^*(\zeta)$, $\forall k \in \mathbb{Z}_{\geq 0}$. In this manner, the UAV considered the impact its control decisions had on the viewing geometry (11) up to 30 seconds into the future, which proved beneficial for good long term performance.

During the tracking experiment, a human operator drove the ground vehicle casually, yet unpredictably, along the dirt roads of Camp Roberts, while the UAV was controlled by a ground control station (GCS). This is illustrated in Figure 12. Although the target remained on the roads for the duration of the tracking experiments, this did not benefit the tracking algorithm since the UAV's control policies were computed with no information regarding road networks. The UAV's control actions were computed using MATLAB[®], which communicated with the GCS and the GPS receiver onboard the truck to acquire the

UAV telemetry and target data and to determine the roll command according to the particular control policy being tested on the UAV. The roll command was then sent back to the GCS and relayed to the UAV.

In real-world conditions, steady winds are often encountered having speeds that constitute a significant portion of a small UAV's airspeed. While the policies presented do not address heavy winds that alter the UAV's kinematics (projected onto the ground) significantly, light to moderate winds can be addressed in an approximate manner by instead altering the target's *apparent* ground velocity. In particular, Saunders notes in [12] that a constant target velocity and steady wind can be generalized to just a steady wind. We take a similar approach and combine the wind and target velocity to form the target's apparent ground velocity, since the coordinate frame defined by (9) and (10) is centered on the target. More specifically, denoting the wind by $\mathbf{w} \in \mathbb{R}^2$, one can take the target's apparent heading $\hat{\psi}_g$ and speed \hat{v} to be the following:

$$\hat{\psi}_g = \text{atan2}(v \sin \psi_g - w_2, v \cos \psi_g - w_1) \quad (28)$$

$$\hat{v} = \sqrt{(v \cos \psi_g - w_1)^2 + (v \sin \psi_g - w_2)^2}, \quad (29)$$

where atan2 is the four-quadrant inverse tangent function. One then uses $\hat{\psi}_g$ and \hat{v} in place of ψ_g and v in (9) and (10) to determine ζ . In reality, wind truly alters the UAV's ground velocity $\mathbf{v}_g \in \mathbb{R}^2$ and hence its groundspeed. More specifically, $\mathbf{v}_g = (s \cos(\psi) + w_1, s \sin(\psi) + w_2)$ per the wind triangle [30]. However, since we have planned for a range of target speeds and only a single UAV groundspeed, it is convenient to use the preceding transformation, which we shall see is quite effective for even moderate wind speeds. Also, note that the heading of the wind does not correspond to the meteorological definition of wind, i.e., $\psi_w = \text{atan2}(w_2, w_1)$ is the direction *in* which the wind is blowing rather than the direction *from* which it is blowing. Furthermore, an estimate of the wind field \mathbf{w} is typically reported by the autopilot of a small UAV and can be accurate to within 0.5 [m/s] of the actual wind speed in practical settings [31]. Moreover, due to the robustness of the policies to either uncertain or worst-case relative dynamics between the UAV and target, estimation error in the wind speed is unlikely to cause any significant degradation in performance.

While this approximation was employed during the flight tests, its effects were quite negligible, as the average wind speed (as measured by the UAV) was less than 0.5 [m/s] over each 15-minute experiment. Since wind speeds are often greater than those that were experienced during this particular experiment, we address the issue of the control policies' robustness to steady winds at the end of the section and keep our focus on robustness to unpredictable target motion.

4.2. Game Theory Results

Three representative minutes of flight are now presented to highlight the key features of vision-based target



Figure 12: A single UAV tracks the unpredictable ground target at Camp Roberts, CA with one of the optimized-based control policies. This image was captured by an independent UAV flying at a higher altitude and illustrates the typical imagery sent to the ground control station for video processing.

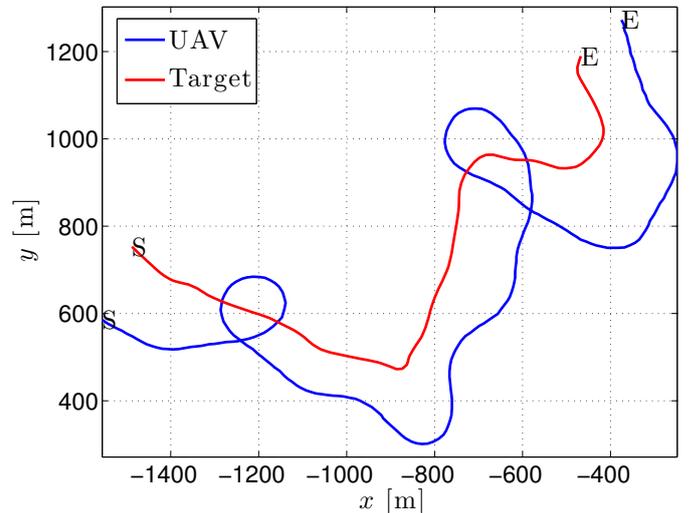


Figure 13: Trajectory of the UAV over 3 minutes as it tracks the target with the game theoretic control policy. An “S” indicates a starting position and “E” an ending position.

tracking with the game theoretic control policy γ^* . The target and UAV trajectories during this window are provided in Figure 13. In the figure, the UAV keeps to the right of the target and occasionally makes counterclockwise turns about the target. Such behavior enables the UAV to stay close to the target in spite of the speed differential and simultaneously keep the target in its FOR. Another characteristic of the UAV's trajectory is that it never passes over the target, which would minimize the distance to the target but lead to an elevation angle close to 90° . Such behavior illustrates the tradeoff between minimizing the distance to the target and avoiding a large elevation angle. Overall, the UAV exhibits standoffish behavior with the game theoretic control policy, meaning that it appears

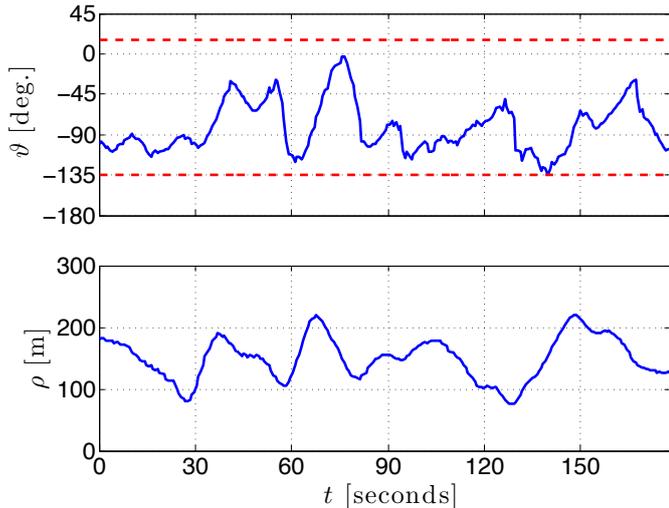


Figure 14: Critical components of the viewing geometry performance with the game theoretic control policy: azimuth ϑ and 2-D distance to target ρ . The mechanical limits of the gimbal, θ_ℓ and θ_u , are indicated by dashed lines in the plot of azimuth angle.

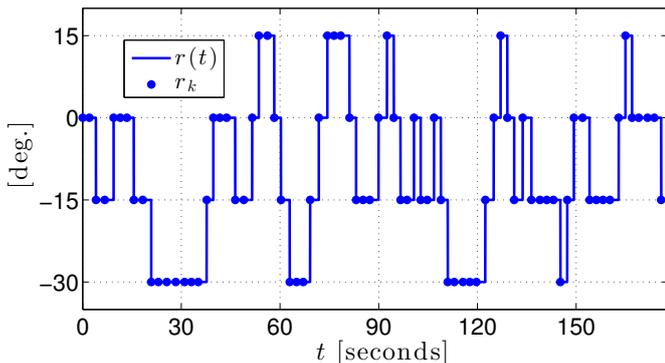


Figure 15: Roll command sequence under the game theoretic control policy γ^* .

reluctant to perform loops that would likely make it vulnerable to an evasive target.

Figure 14 highlights the critical components of the viewing geometry. The azimuth angle was kept within the mechanical limits of the gimbal mechanism, and hence the target was kept within the FOV during this entire 3 minute window. However, the azimuth varies significantly in this window, as it takes on values near both extremes. This is quite visible in the time interval from $t = 60$ to $t = 75$ seconds when the gimbal slews from the back left of the UAV to the front. The gimbal also nearly hit its lower mechanical limit near $t = 140$ seconds. Another feature of the viewing geometry is that the planar distance to target oscillates a fair amount, as it varies between roughly 100 and 200 meters during this 3-minute window.

Lastly, Figure 15 illustrates the roll command sequence under the game theoretic policy γ^* . One immediately apparent characteristic of this plot is that the maximum roll

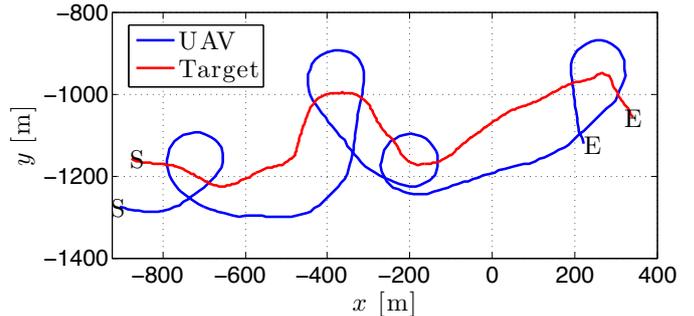


Figure 16: Trajectory of the UAV over 3 minutes as it tracks the target with the stochastic optimal control policy. An “S” indicates a starting position and “E” an ending position.

command of $+30^\circ$ is never used; however, the positive roll command of 15° is employed to achieve the standoffish behavior shown in Figure 13. Note how the roll-angle action space of (3) limits changes in roll to $\Delta = 15^\circ$ degrees, and hence chattering between roll commands of opposite polarity is avoided.

4.3. Stochastic Optimal Control Results

Three representative minutes of flight during which roll commands were governed by μ^* are now presented to highlight the salient features of vision-based target tracking with the stochastic optimal control policy μ^* . The UAV and target trajectories are depicted in Figure 16. As with the game theoretic controller, the UAV keeps to the right of the target, and performs loops as necessary to compensate for the speed differential between itself and the target vehicle.

One immediate difference between the two control policies is that during the same time window the UAV performed twice as many loops under the stochastic optimal control policy. One might attribute this noticeable difference to the target speed profile over the 3 minute window for both cases; however, the target traveled nearly 1,739 [m] in the game theoretic case and 1,562 [m] in the stochastic optimal control case at average speeds of about 9.77 [m/s] and 8.83 [m/s], respectively. The average speed difference of just under 1 [m/s] is unlikely the primary source of this feature. Rather, this behavioral differences is best explained by the differences in the control policies themselves. In particular, as noted earlier in the discussion of the stochastic optimal control policy, the UAV makes loops earlier when it is traveling to the right of the target with zero roll and zero relative heading, as shown in the fourth quadrant of the control surface in Figure 10. With the game theoretic control policy, the UAV takes a more precautionary stance against the target (assumed to be evasive), and accordingly, exhibits a reluctance to perform loops around the target.

Figure 17 illustrates the critical components of the viewing geometry, namely the azimuth angle and planar distance to target. One immediately noticeable feature of the

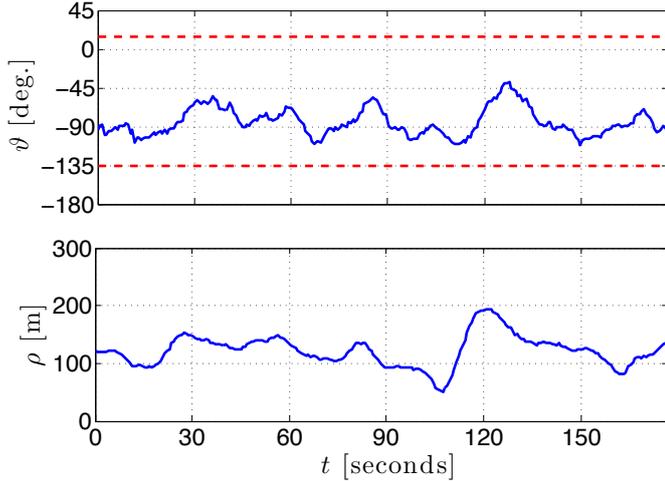


Figure 17: Critical components of the viewing geometry performance with the stochastic optimal control policy: azimuth ϑ and 2-D distance to target ρ . The mechanical limits of the gimbal, θ_ℓ and θ_u , are indicated by dashed lines in the plot of azimuth angle.

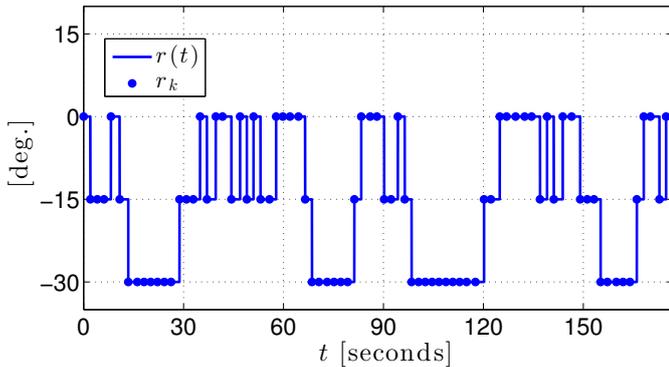


Figure 18: Roll command sequence under the stochastic optimal control policy μ^* .

azimuth angle ϑ is that it oscillates very little in comparison to that of the game theoretic case. Furthermore, the gimbal never slews from one extreme to the next, nor does it come close to either extreme.

Concerning the planar distance to target ρ , there is one instance where it reaches nearly 200 [m], yet for the majority of the time window, the distance oscillates between 100 and 150 [m]. Moreover, the amplitude of the oscillations is nearly half that of the game theoretic case, and both the average value and variance of ρ are also significantly smaller.

Finally, Figure 18 depicts the roll command sequence under the stochastic optimal control policy μ^* . The most prominent feature of this figure is that no positive roll commands are ever employed. Rather, one will notice a slight chattering behavior between $r = 0^\circ$ and $r = -15^\circ$. Also, during this particular 3-minute time window, anytime the roll command $r = -30^\circ$ degrees is used the UAV is performing a left turn, which cannot be said of the game

Table 7: Statistics over 15 minute window

Statistic	γ^*	μ^*	Units
min ϑ	-166.7	-140.3	deg.
max ϑ	16.63	20.19	deg.
var ϑ	1026.8	663.3	deg. ²
avg ρ	150.9	131.8	m
var ρ	2,104	1,372	m ²
avg $g(\zeta)$	10.76	7.753	N/A

theoretic control policy γ^* .

4.4. Quantitative Comparison

Quantitative statistics are now presented to solidify the advantages of one technique over the other. Since each approach was tested in flight for a minimum of 15 minutes, statistics are presented for select quantities over a whole 15 minute window for the purpose of an objective comparison. The statistics for critical parameters resulting from both control approaches are presented in Table 7, where the game theoretic approach is represented by its corresponding control policy γ^* and the stochastic optimal control approach is represented by its corresponding control policy μ^* .

The first parameter to consider is the azimuth angle ϑ . Both approaches caused the gimbal to saturate and hit each of its mechanical limits of -135° and 15° at least once, as indicated by the min and max statistics for ϑ . However, the azimuth angle never exceeded the FOR limits of -152° and 32° with the optimal control policy μ^* , though it exceeded the lower bound once with the game theoretic control policy γ^* . Moreover, the UAV never lost sight of the target with μ^* and lost sight of it just once with γ^* . As noted in the previous section, the azimuth angle varied considerably more with the game theoretic approach than with the stochastic approach, which is indicated by the variance statistic of the said parameter.

Now consider the planar distance to target ρ . Its average value was nearly 18 [m] less with μ^* than with γ^* , and its variance was also significantly less with μ^* . Coincidentally, the stochastic optimal control approach also offered a near 35% reduction in the variances of both the azimuth and distance parameters, and finally, it also achieved the lower average viewing geometry cost.

From the quantitative results of Table 7, the stochastic optimal control approach proves to be the superior method for the application of vision-based target tracking, as it performed better on all fronts and yielded significantly lower variances for both the azimuth and distance parameters. The most noteworthy feature of this approach is that it never lost sight of the target even once in its 15 minutes of flight time while maintaining proximity to the target.

Posing the problem of vision-based target tracking as a two-player zero-sum game is a natural way to address the

problem when one wishes to ensure robustness to worst-case / evasive target motion. However, for a more casual target vehicle, the precautionary behavior from the behavioral control policy may reduce viewing geometry performance. Another plausible contributor to reduced performance is the fact that the construction of the game theoretic policy assumed first-order roll dynamics, whereas the stochastic approach incorporated a more realistic model of the roll dynamics (Figure 6) by treating them as a source of stochasticity. Thus, to improve performance with the game theoretic approach, one could incorporate a second-order model for the roll dynamics of the UAV. Nevertheless, having lost sight of the target only once in 15 minutes, the current game theoretic control policy is still an effective means of vision-based target tracking and potentially much more robust for an evasive target.

4.5. Wind Considerations

While there was relatively little wind during flight testing, steady wind generally constitutes a significant portion of a small UAV’s airspeed, thereby altering the aircraft’s ground speed significantly. Hence, we performed high fidelity simulations of the existing policies in steady winds to determine their robustness margins to such wind. In particular, the UAV was simulated using the high fidelity flight simulator Aviones while the target trajectory was that encountered during the flight test for the stochastic optimal control policy. The wind field estimate of the simulated autopilot was utilized in (28) and (29) to accommodate nonzero winds with the original control policies. Moreover, we conducted multiple 15-minute experiments to test each policy’s performance for various constant-speed winds. Since the target’s primary direction of travel was west, we simulated steady winds to the east in the flight simulator for the duration of each experiment. Moreover, the UAV primarily experienced a headwind while tracking the target in each experiment. Wind was incorporated into each controller using the approximation described at the end of Section 4.1.

The results for the game theoretic control policy and the stochastic optimal control policy are provided in Tables 8 and 9, respectively, where the middle four columns indicate the absolute wind speed of each scenario. From Table 8, the game theoretic policy γ^* exceeded both upper and lower FOR limits in all of the cases that were tested. In addition, the last four quantities in the fourth column (6-m/s scenario) are significantly greater than their counterparts in the third column (4.5-m/s scenario). Hence, even though the increase in wind speed was only 1.5 [m/s], the upper limit of the game theoretic control policy’s ability to track in steady wind was crossed as the wind was increased from 4.5 to 6 [m/s]. Thus, γ^* was not tested in 7.5-m/s wind. From Table 9, the stochastic optimal control policy μ^* fared quite well in the first 3 cases, never losing sight in the first two cases and losing sight of the target only once in the third case, though the target exited the FOR by less than 1 degree, as seen in the first

Table 8: Statistics for game theoretic policy in steady wind over 15 minute window

Statistic	3 m/s	4.5 m/s	6 m/s	7.5 m/s	Units
min ϑ	-165.5	-164.6	-172.3	-	deg.
max ϑ	61.09	55.16	86.15	-	deg.
var ϑ	1,378	1,273	2,318	-	deg. ²
avg ρ	158.5	182.4	217.5	-	m
var ρ	2,967	5,890	9,415	-	m ²
avg $g(\zeta)$	12.28	14.81	23.74	-	N/A

Table 9: Statistics for stochastic optimal control policy in steady wind over 15 minute window

Statistic	3 m/s	4.5 m/s	6 m/s	7.5 m/s	Units
min ϑ	-132.0	-136.1	-152.6	-131.5	deg.
max ϑ	8.688	20.22	30.5	45.52	deg.
var ϑ	634.7	954.6	927.3	1,529	deg. ²
avg ρ	138.8	140.0	146.3	190.8	m
var ρ	1,692	2,385	2,542	8,181	m ²
avg $g(\zeta)$	7.813	8.846	9.244	15.82	N/A

row of the fourth column. In the fourth scenario of 7.5-m/s wind, the last four quantities in the fifth column made significant jumps, which indicated that the boundaries of this policy’s capabilities were being crossed, even though the policy lost sight of the target only twice.

Based on these studies, the conservative upper limits for wind speeds that can be tolerated by the policies are 4.5 [m/s] (1/4 of the UAV’s airspeed) for the game theoretic case and 6 [m/s] (1/3 of the UAV’s airspeed) for the stochastic optimal control case. One would surely improve the UAV’s performance in the wind speeds that were tested and perhaps even greater wind speeds by incorporating wind velocity into the models; however, the dimensionality of the problem would grow, thereby presenting considerable computational challenges for basic dynamic programming techniques. Another issue that arises by addressing heavy winds is that the range of target speeds that can be tracked is reduced, as we must have that the combined wind / target speed given by (29) must be strictly less than the UAV’s airspeed for the tracking problem to be well posed. Moreover, there is a tradeoff between the speed of the vehicle that can be tracked and the amount of wind that can be tolerated.

5. Overall Conclusion and Future Work

While the existing literature offers methods for target tracking using continuous-time feedback control laws [6–12] or optimization based methods [13–15], these individual works make assumptions that simplify the UAV dynamics, target motion, and/or sensor visibility constraints, thereby hindering the feasibility of a real world implementation with actual hardware. This paper has detailed

the design of two optimization-based control policies for vision-based target tracking, where strict trajectories must be flown by an underactuated UAV to maintain visibility and proximity to an unpredictable ground target. Both control policies have been flight tested, thereby verifying their robustness to environmental disturbances, unpredictable target motion, and unmodeled dynamics.

The approach based on stochastic optimal control proved most effective through an appropriate design choice of a cost function to facilitate good viewing geometry and also by directly addressing the nontrivial roll dynamics shown in Figure 6 and the unpredictable changes in target velocity illustrated in Figure 10. This paper has shown that, with an appropriate choice of cost function, one can use relatively simple kinematic models for motion planning under uncertainty so long as one incorporates sufficient stochasticity into the model.

Future work involves testing the stochastic optimal control policy with estimates from the target state constructed from the vision sensor onboard the UAV. The vision-based measurements of the target position not only suffer from errors proportional to the UAV's distance from the target as detailed in [22], but they are also corrupted by outliers. Accordingly, the second author has been developing a robust filter that utilizes ℓ_1 regularization to reject outliers and simultaneously provide a filtered estimate of the target state. The coupled problem of control and state estima-

tion must be tested, both in simulation and in flight tests, to verify the robustness of the control policy to noise in the estimate of the target state. Incorporating stochastic wind into the dynamical models is also another avenue for future work, as this would enable the stochastic optimal control policy to operate robustly in heavier winds.

The final venue for future work is to extend our work to two UAVs as in [22], but with the gimbal constraints of Section 2.2 and the stochastic kinematics of Sections 3.2 and 3.3. The goal would be for the two UAVs to coordinate their distances to the target to gather the best joint vision-based measurements, which entails that at least one UAV is always relatively close to the target with the target in its field of regard. By achieving distance coordination with two UAVs, one should see a significant reduction in the estimate of the target state produced by the robust filter.

Acknowledgments

Research supported by the Institute for Collaborative Biotechnologies through grant W911NF-09-0001 from the U.S. Army Research Office.

The authors would like to thank Gaemus Collins, Mike Wiatt, Chris Stankevitz, and Paul Filitchkin of Toyon Research Corporation for their help in conducting the flight experiments.

References

- [1] J. Saunders, Obstacle Avoidance, Visual Automatic Target Tracking, And Task Allocation For Small Unmanned Air Vehicles, Ph.D. thesis, Brigham Young University, 2009.
- [2] G. E. Collins, C. R. Stankevitz, J. Liese, Implementation of a Sensor Guided Flight Algorithm for Target Tracking by Small UAS, in: Ground/Air Multi-Sensor Interoperability, Integration, and Networking for Persistent ISR II, vol. 8047, SPIE, 2011.
- [3] J. Lee, R. Huang, A. Vaughn, X. Xiao, J. K. Hedrick, M. Zenaro, R. Sengupta, Strategies of path-planning for a UAV to track a ground vehicle, in: The Second Annual Symposium on Autonomous Intelligent Networks and Systems, Menlo Park, CA, 2003.
- [4] C. R. Husby, Path generation tactics for a UAV following a moving target, Master's thesis, University of Washington, 2005.
- [5] R. W. Beard, A Class of Flight Trajectories for Tracking Ground Targets with Micro Air Vehicles, in: Mediterranean Conf. on Control and Automation, 2007.
- [6] E. Lalish, K. Morgansen, T. Tsukamaki, Oscillatory Control for Constant-Speed Unicycle-Type Vehicles, in: IEEE Conf. on Decision and Control, 2007.
- [7] N. Regina, M. Zanzi, UAV guidance law for ground-based target trajectory tracking and loitering, in: Aerospace Conference, 2011 IEEE, ISSN 1095-323X, 1 –9, doi: 10.1109/AERO.2011.5747522, 2011.
- [8] E. W. Frew, Lyapunov Guidance Vector Fields for Unmanned Aircraft Applications, in: American Control Conf., 2007.
- [9] T. H. Summers, Cooperative Shape and Orientation Control of Autonomous Vehicle Formations, Ph.D. thesis, University of Texas at Austin, 2010.
- [10] V. N. Dobrokhodov, I. I. Kaminer, K. D. Jones, R. Ghabche- loo, Vision-Based Tracking and Motion Estimation for Moving targets using Small UAVs, in: American Control Conf., 2006.
- [11] Z. Li, N. Hovakimyan, V. Dobrokhodov, I. Kaminer, Vision-based Target Tracking and Motion Estimation Using a Small UAV, in: IEEE Conf. on Decision and Control, Atlanta, GA, 2010.
- [12] J. Saunders, R. W. Beard, Tracking a Target in Wind Using a Micro Air Vehicle with a Fixed Angle Camera, in: American Control Conference, Seattle, WA, 3863–3868, 2008.
- [13] R. Anderson, D. Milutinović, A Stochastic Approach to Dubins Feedback Control for Target Tracking, in: IEEE / RSJ Conf. on Intelligent Robots and Systems, 3917–3922, 2011.
- [14] S. S. Ponda, R. M. Kolacinski, E. Frazzoli, Trajectory Optimization for Target Localization Using Small Unmanned Aerial Vehicles, in: AIAA Guidance, Navigation, and Control Conference, 2009.
- [15] S. A. Miller, Z. A. Harris, E. K. P. Chong, A POMDP Framework for Coordinated Guidance of Autonomous UAVs for Multi-target Tracking, EURASIP J. Adv. Signal Process (2009) 1–17.
- [16] S. Thrun, W. Burgard, D. Fox, Probabilistic Robotics, The MIT Press, 2005.
- [17] H. Oh, S. Kim, A. Tsourdos, B. A. White, Decentralised Stand-off Tracking of Moving Targets Using Adaptive Sliding Mode Control for UAVs, Journal of Intelligent & Robotic Systems (2013) 1–15.
- [18] E. W. Frew, Sensitivity of Cooperative Geolocation to Orbit Coordination, in: AIAA Guid., Nav., and Control Conf., Hilton Head, SC, 3869–3892, 2007.
- [19] M. Wheeler, B. Schrick, W. Whitacre, M. Campbell, R. T. Rysdyk, R. A. Wise, Cooperative Tracking of Moving Targets by a Team of Autonomous UAVs, in: IEEE / AIAA 25th Digital Avionics Systems Conference, 2006.
- [20] D. B. Kingston, Decentralized Control of Multiple UAVs for perimeter and target surveillance, Ph.D. thesis, Brigham Young University, 2007.
- [21] M. Stachura, A. Carfang, E. W. Frew, Cooperative target tracking with a communication limited active sensor network, in: International Workshop on Robotic Wireless Sensor Networks, Marina Del Ray, CA, 2009.
- [22] S. A. P. Quintero, F. Papi, D. J. Klein, L. Chisci, J. P. Hespanha, Optimal UAV Coordination for Target Tracking using Dynamic Programming, in: IEEE Conf. on Decision and Control, Atlanta, GA, 2010.
- [23] H. Oh, S. Kim, H.-S. Shin, A. Tsourdos, B. White, Coordinated standoff tracking of groups of moving targets using multiple UAVs, in: Control & Automation (MED), 2013 21st Mediterranean Conference on, IEEE, 969–977, 2013.
- [24] R. Isaacs, Differential Games, John Wiley and Sons, 1965.
- [25] J. Dobbie, Solution of Some Surveillance-Evasion Problems by the Methods of Differential Games, in: Proceedings of the 4th International Conference on Operational Research, 1966.
- [26] J. Lewin, J. V. Breakwell, The Surveillance-Evasion Game of Degree, Journal of Optimization Theory and Applications 16.
- [27] J. P. Hespanha, An Introductory Course in Noncooperative Game Theory, available at <http://www.ece.ucsb.edu/~hespanha/published>, 2011.
- [28] URL <http://aviones.sourceforge.net>, 2007.
- [29] F. L. Lewis, D. Vrabie, V. L. Syrmos, Optimal Control, John Wiley and Sons, Hoboken, New Jersey, 3 edn., 2012.
- [30] R. W. Beard, T. W. McLain, Small unmanned aircraft: Theory and practice, Princeton University Press, 2012.
- [31] J. W. Langelaan, N. Alley, J. Neidhoefer, Wind field estimation for small unmanned aerial vehicles, Journal of Guidance, Control, and Dynamics 34 (4) (2011) 1016–1030.