

Low-Altitude Road Following Using Strap-down Cameras on Miniature Air Vehicles

Joseph Egbert and Randal W. Beard
Electrical and Computer Engineering
Brigham Young University

October 19, 2010

Abstract

This paper addresses the problem of autonomously maneuvering a miniature air vehicle (MAV) to follow a road using computer vision as the primary guidance sensor. We focus on low altitude flight with the objective of maximizing the pixel density of the road in the image. The airframe is assumed to be a bank-to-turn fixed-wing MAV with a downward-looking strap-down camera. The road is identified in the image using HSV classification, flood-fill operations, and connected component analysis. The main contribution of the paper is the derivation of explicit roll-angle and altitude-above-ground-level (AGL) constraints that guarantee that the road will remain in the footprint of the camera, assuming a flat earth. The effectiveness of our approach is demonstrated through high fidelity simulation and through flight results.

1 Introduction

This paper addresses the problem of tracking a visually distinct connected structure on the ground using a miniature air vehicle equipped with a strap down camera. Potential structures include roads, pipelines, and perimeters of forest fires. To be specific, we will focus our attention in this paper on road following.

A potential road-following application of interest to the military is convoy support where the objective is to detect roadside bombs and ambushes along the anticipated route of the convoy. Surveillance and reconnaissance of these routes can be accomplished using low-altitude miniature air vehicles (MAVs) which relay video data to human operators associated with the convoy. Passive sensing devices like video cameras are covert and are therefore essential for this type of military operation. The GPS position of a road may not be available, and even if it is, the MAV's GPS receiver may be jammed by an adversary. Vision-based control using automated road tracking can meet the reconnaissance needs for convoy support while accommodating for potential GPS jamming without revealing the MAV's presence and position.

There are many applications where borders are not well defined or adequately mapped. Fire fighters confront situations where the extent of a fire may be unknown and changing in time. High-altitude aircraft and satellites are often requested for overhead imagery but these resources are not always available. Surveillance of fire perimeters using manned helicopters and aerial vehicles are limited because updrafts resulting from the fire create dangerous flying conditions. For safety reasons, manned aircraft are not used at night to survey forest fires. MAVs are an attractive alternative because they are relatively inexpensive to manufacture and replace, and because they can be used to obtain high-resolution imagery of the fire and its perimeter.

Studies that specifically address border and road following using cameras on-board unmanned ground vehicles (UGVs) have appeared in the literature for more than a decade [1], [2], [3]. Each of these studies presents a method to effectively extract the road from the video stream, project the estimated road position onto the inertial frame, and maneuver the vehicle to follow the road. A vanishing point in the image is used to simplify computer-vision algorithms and provide improved road-directional bearings. A vanishing point is the point in the image where the road merges with the horizon. In [1], an edge-detection algorithm is used to extract the road edges. By analyzing a smaller image window, the algorithm processes more complex calculations in real-time including estimating the road location. In [2], the road is registered onto a virtual map. Using this virtual map, sophisticated path corrections and intersection analysis improve the UGVs estimated path as it travels. The ability for a UGV to adjust to non-ideal road conditions is demonstrated in [3], where two assumptions are used to ensure that the road is effectively classified by processing the on-board video. The first assumption is that the vehicle begins operation on the road and generally remains there while the guidance algorithm is active. Secondly, ground-vehicle movement and general road structure are estimated using the vanishing point of the road. These assumptions significantly ease computation and decision complexity in the guidance loop.

Overhead views obtained by MAVs provide an alternate perspective to strategically important areas that cannot be obtained by ground vehicles. Researchers have only recently begun to address the specific guidance issues associated with road following using video cameras on fixed-wing MAVs. Earlier studies have addressed some problems associated with vision-based guidance of UAVs. Vision-based landing of an unmanned helicopter has been demonstrated in [4] and [5]. Low-altitude path planning has been proposed where a virtual map is updated based on the data captured from an onboard camera in [6] and a path is planned to avoid terrain obstacles. In [7], a ground-based object is localized using a fixed-wing MAV, and errors and difficulties associated with vision analysis are discussed.

Studies addressing the guidance issues associated with fixed-wing MAV road following have been reported in [8], [9], [10]. Flight navigation based solely on captured images from an onboard strap-down camera has been demonstrated in [9], where the camera is pointed toward the center of the earth. In [10], it has been shown that MAVs can follow various locally linear structures using

visual feedback from a strap-down camera pointed downward with a specified look-ahead distance. The look-ahead distance is selected heuristically using simulation studies. By looking ahead, the UAV can curve onto the road segment in front of the vehicle. References [9] and [10] assume that the air vehicle can perform skid-to-turn which significantly simplifies the problem of maintaining the road in the footprint of the camera.

While [9] and [10] demonstrate vision-based path following using a fixed-wing MAV, they do not address the problem of maintaining the road in the camera footprint using bank-to-turn maneuvers. The primary contribution of this paper is to develop altitude-above-ground-level (AGL) and roll-angle constraints that ensure that the road remains in the camera footprint for bank-to-turn MAVs. Forcing a bank-to-turn MAV to follow a road is difficult because the image footprint and road are not physically linked. Consider the scenario where a MAV is flying parallel to a road with a ground track that is offset a fixed distance to the left of the road. In order to maneuver so that its ground track is directly over the road, the MAV must bank to the right. However, when it banks right, the camera footprint moves to the left and if the bank angle is steep enough, the road may leave the footprint of the camera. Because the ground track is initially offset to the left of the road, it will be imaged on the right side of the camera footprint and the distance from the center of the image is a monotonic function of the AGL. Therefore, the allowable bank angles that maintain the road in the camera footprint is also a function of AGL. Similar to [10], these constraints apply to any locally-linear structure including roads, fire perimeters, and pipelines.

While a gimballed camera would resolve some of the issues described above, size and weight constraints limit the use of gimbals on MAVs. A gimballed camera is able to adjust the camera angle during flight to obtain the preferred camera view. Alternatively, a strap-down camera limits MAV maneuverability during road tracking. Real-time computer vision and control strategies must be implemented to maintain the proper flight trajectory for the MAV so that the road remains in the camera field-of-view. The maneuverability limits are a function of the camera's horizontal field of view $\bar{\eta}$, road width w_{road} , the expected turn angle of the road μ , and the MAV's AGL h .

Computer vision algorithms are used to extract the road from the video stream and estimate its location and heading in world coordinates. To be successful, computer vision classification algorithms rely upon certain assumptions which are often derived from training sets or object characteristics. A training set is a collection of images used to teach the model the object's characteristics. Computer vision techniques for road following largely derive from ground vehicle applications. Reference [11] uses several previous images to adaptively classify the road using reverse optical flow to estimate the evolution of its structure. In [12], two approaches are introduced to deal with spatial and temporal changes in the road. Two cameras with different iris settings are used to filter out illumination differences. In [13], dominant orientations are extracted from the image. Using the vanishing point and dominant structural orientations allows the algorithm to extract an estimate of the road. The algorithm in [13] can

determine if it is off-road and estimate a new direction toward the lost road. In [14], a method to correctly adjust to lighting conditions for effective road classification is presented. A training image is used to establish the road’s color bounds and these color bounds are updated as the vehicle travels to adjust to gradual road changes. In [15], a computer-vision approach is presented for a MAV that removes video noise as well as classifying objects like roads, trees, and sky.

Several papers not associated with unmanned vehicles proved useful in developing an effective road-classification algorithm for MAVs. In [16], the authors produce a region-based query system that allows objects in the image to be classified more correctly. An adaptive-classification approach using a statistical threshold is presented in [17]. Methods to increase processing speed of connected-component and histogram analysis are presented in [18]. Effective color spaces for road extraction are discussed in [19], [3], and [2]. Hue-based color models (hue-saturation-intensity HSI and hue-saturation-value HSV) are more effective color spaces for extracting asphalt roads. In this paper, the strategies presented in [14] are implemented using HSV as the color space. The image-analysis strategy will be discussed in further detail in Section 3.

The paper is organized as follows. In Section 2 we give an overview of the system architecture. The computer vision algorithms used to classify and extract the road from the video stream are discussed in Section 3. In Section 4 we derive roll-angle and AGL constraints that guarantee that the road will remain in the camera footprint. Results for tracking a road whose world coordinates are unknown a priori, but whose geometry is known, are discussed in Section 4.3. Tracking a road whose world coordinates are known a priori from digital-elevation map (DEM) data, is discussed in Section 4.5. In Section 5 we present simulation and experimental results to demonstrate the validity of the derived constraints. Conclusions and future directions are discussed in Section 6. A preliminary version of the results contained in this paper, but without flight validations, were presented in [20]

2 Architecture

The high level tasks for road following are separated into two categories: computer vision, and guidance and control. The block diagram in Figure 1 illustrates the information flow of the road-following algorithm. The Road Classifier uses the method developed in [17] to segment road and non-road pixels and creates a binary-image representation. Noise is removed during this process using flood-fill and erosion techniques. The Road Direction Finder uses component orientation and geometry to determine a road direction, and outputs a road position and heading angle in camera-frame coordinates. The Localizer transforms the road position and heading angle into world coordinates. Using the localized coordinates and telemetry information of the MAV, the Constraints Governor outputs acceptable world heading and tracking-error data to the Path Following algorithm which computes a desired heading for the MAV using the technique

described in [21]. The autopilot converts the desired heading to control-surface commands.

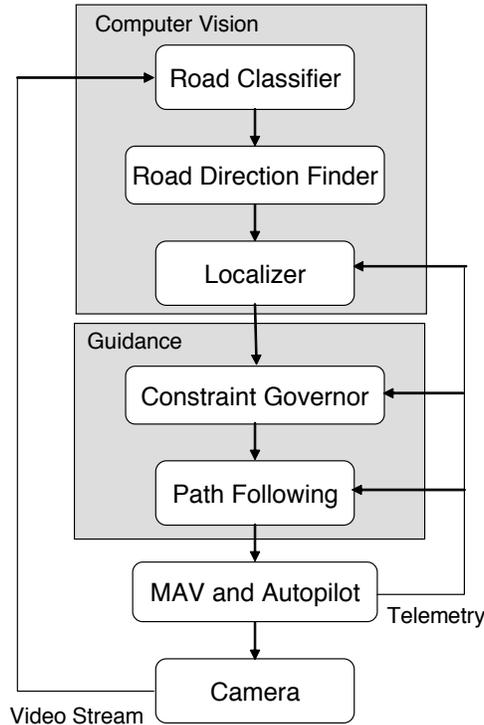


Figure 1: This figure illustrates the guidance loop architecture for road following. The video stream is processed in the Computer Vision block to find the location and direction of the road with respect to the MAV. Geometric analysis is used to derive constraints on roll angle and AGL. These constraints help maintain the road in the camera footprint. A path following guidance algorithm maneuvers the MAV to follow the road.

Several failsafes are implemented by the Constraints Governor to minimize road-estimation error and smooth the captured video. If the road leaves the camera footprint, the computer-vision algorithm is bypassed and the MAV is commanded to follow the predicted path derived from the last successful pass through the guidance loop. When the estimated path reenters the camera footprint, the computer-vision algorithm resumes operation. To smooth the video, the gain on the roll rate is reduced to limit chatter. As explained in Section 4, the roll angle is also limited according to the calculations of the expected geometry of the road. The previous road estimate is also used if the Road Classifier is unable to obtain a connected component of a certain size.

The architecture shown in Figure 1 allows the processor on the MAV to perform the essential operations for stable flight. The road-following guidance

loop requires knowledge of the MAV position and orientation to localize the road. The GPS provides the MAV's inertial position but is not necessary for the guidance loop because the relative tracking error between the MAV and the road is obtained from pitch, roll, and yaw estimates which are independent of absolute world coordinates.

3 Road Classification and Directional Analysis

In this section we discuss the Computer Vision block of the architecture shown in Figure 1. Image analysis requires the majority of the computational resources and must be carefully considered in the design of the guidance-loop design. The first step is to segment the image into road and non-road components. Once the road location in the image has been determined, the Road Direction Finder and Localizer calculate the road's relative inertial position and the corresponding desired heading for the MAV.

Real-time requirements for road following limit the potential computer-vision algorithms available to classify the road. The Road Classifier and Direction Finder may produce a bottleneck in the guidance loop causing delays in the commands sent to the MAV. Because of this, using a computationally efficient Road Classifier and Road Direction Finder will improve the stability of the guidance loop. Computing data using several threads for computer vision and MAV control provide even greater efficiency. The studies cited in Section 1 concerning computer-vision techniques were used to create an effective road-classification system and accelerate the algorithm.

Our road-classification algorithm uses the hue-saturation-value (HSV) color space to extract road pixels. The red-green-blue (RGB) color space is inadequate because the green value contributes little to asphalt detection as explained in [2] and the red and blue values provide weaker discriminating power. In [19], HSV is shown to provide excellent discrimination power for asphalt roads. The H and V values suffice for effective road-pixel classification. We are encouraged to disregard the S value because it is the noisiest. By disregarding the S value, the classification can be simplified to two dimensions, the H and V values. The Road Classifier classifies each pixel and maps it from the H-V domain to a statistical value λ using the mean and covariance of the road pixels derived from a training set. A threshold determined from the training set assigns the pixel as road or non-road to create a binary image. Unfortunately, both HSV and RGB lack reliable methods to classify road pixels in shadowed areas. Other color spaces can be useful depending on the characteristics of the road. Color spaces that are better for extracting road pixels in shadowed areas or changing cloud cover are suggested in [3] and [1]. In the case that the MAV is tracking a fire perimeter, an infrared or near-infrared camera can be used to extract pixels representing hot regions.

Pixel-wise analysis and classification of the image cause holes and noise in the resulting binary image that are removed using a flood-fill operation. Flood filling allows the Road Classifier to include larger sections of the road normally outside

the classification threshold. Vehicles or other objects that are on the road are filled and included in the directional analysis. The image is also eroded to reduce erroneous road-pixel classification at the road edges and speckle noise. Finally, the largest component is assumed to be the road while all other components are removed from the image. The H-V histogram information of the resulting road component is stored and used to classify future frames. The resulting binary image is composed of pixels classified as road and non-road.

Figure 2 illustrates the process used by the Road Classifier and Road Direction Finder in extracting the road and determining its direction. Note that the order of operation can have a significant impact on the classification results. Figure 2 shows that by filling the components first, a solid road component can be extracted from the image for analysis. Noise is removed after the filled road component is extracted.

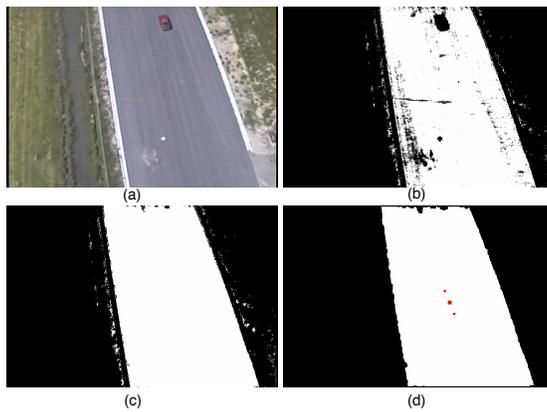


Figure 2: This figure illustrates the steps used to extract the road from the image. Figure (a) shows the image captured by the camera. Figure (b) shows the results of classifying using the H-V color space to find the road. Figure (c) shows the results of the flood-fill operation. In Figure (d) the noise components are removed by eliminating smaller connected components and eroding the large road component.

The binary image along with its component histogram (i.e. histogram of the black and white pixels along horizontal and vertical directions of the image) is passed to the Road Direction Finder. The position of the road is found by calculating the mean of the component histogram as

$$m_x = \frac{\sum xf(x)}{\sum f(x)}, \quad (1)$$

$$m_y = \frac{\sum yf(y)}{\sum f(x)}, \quad (2)$$

where $f(x)$ is the horizontal histogram of the road component and $f(y)$ is the vertical histogram of the road component. The orientation Θ of the road is

found by using the variance

$$\sigma_{ij} = \frac{\sum_x \sum_y (x - m_x)^i (y - m_y)^j f(x, y)}{\sum f(x, y)}, \quad (3)$$

where $f(x, y)$ is the histogram associated with the particular image directions represented by i and j . These directions are horizontal $(i, j) = (2, 0)$, vertical $(0, 2)$, or diagonal $(1, 1)$. The variances are calculated to derive the orientation Θ of the road in the image plane given as

$$\Theta = \frac{1}{2} \tan^{-1} \frac{2\sigma_{11}}{\sigma_{20}\sigma_{02}}. \quad (4)$$

In urban terrain, roads possess frequent intersections. When two possible road directions appear, the algorithms described above does not always yield a correct solution. Therefore, an intersection-analysis algorithm was developed to account for these situations. Intersections are detected by searching the image edges for branches in the road. Directions are estimated using the edges of the image until the intersection leaves the camera footprint. This method allows the MAV to select which branch of the intersection it will follow depending on a human supervisor's preference.

4 Roll-Angle and AGL Constraints Imposed by Road Following

If the road stays within the camera field-of-view, the computer vision algorithm described in Section 3 will extract the road and determine its direction in the inertial frame. The objective of this section is to derive roll angle and AGL constraints that guarantee that the road will remain in the camera footprint, given certain assumptions on the flight path and on the geometry of the road.

We consider three general classes of roads. The first class of roads will those that are composed of straight line segments where the angle between subsequent line segments is bounded by $0 < \mu \leq \pi/2$. The second class of roads will be curved roads with bounded curvature. For both of these classes of roads, we assume that the the bound on the angle μ and the bound on the curvature of the road is known to the guidance algorithm. However, we do not assume that the specific geometry and inertial position of the road is known. The third class of roads are those where the geometry and position of the road is known a priori from, for example, a DEM map. Since the world coordinates of the road are known, a feedforward term can be added to the guidance loop and thereby significantly relax the roll angle and AGL constraints necessary to keep the road in the camera field-of-view.

To simplify the analysis, we assume that the optical axis of the camera is pointed out of the belly of the MAV and that the MAV maneuvers at constant altitude. In other words, the flight path angle is $\gamma = 0$, and the gimbal azimuth angle measured from the body-frame x -axis, which points out the nose of the

aircraft, is $\alpha_{az} = 0$, and the gimbal elevation angle measured from the body $x-y$ plane, where the y -axis points out the right wing of the MAV, is $\alpha_{el} = -\pi/2$. We will assume a flat-earth model and that the MAV is flying in zero-wind conditions. The limit on the camera's horizontal field-of-view is denoted by $\bar{\eta}$, and we assume that the autopilot maintains a constant AGL h and MAV speed v .

4.1 Camera-Frame to Inertial-Frame Projection

The relevant reference frames include the inertial frame, the body frame, the gimbal frame, and the camera frame. The inertial frame whose axes are denoted by (X^i, Y^i, Z^i) , is a fixed frame with X^i directed North, Y^i directed East, and Z^i directed toward the center of the earth. The orientation of the body frame whose axes are denoted by (X^b, Y^b, Z^b) is obtained from the inertial frame by first rotating about Z^i by the yaw angle ψ , then rotating about the transformed y -axis by the pitch angle θ , and then rotating about the transformed x -axis by the roll angle ϕ . The body frame is centered at the MAV center of mass. The gimbal frame is obtained by rotating the body frame about Z^b by the azimuth angle α_{az} , and then rotating the resulting frame about Y^g by the elevation angle α_{el} . The orientation of the camera frame, denoted (X^c, Y^c, Z^c) , originates at the optical center with X^c pointing right in the image, Y^c pointing down in the image, and Z^c directed along the optical axis.

The camera geometry is shown in Figure 3. During road following, the road will be primarily oriented from top to bottom in the image. We will therefore focus on the position of the road relative to the camera axis X^c . If ϵ_x is the pixel associated with the point where the road crosses the camera axis X^c , and if f is the focal length of the camera measured in pixels, then define the associated angle as $\eta_x = \tan^{-1}(\epsilon_x/f)$, as shown in Figure 3. Define $\bar{\eta}$ to be the angle associated with the right-most pixel along the x -axis in the camera frame. The objective in this section is to find the world coordinates (q_N, q_E) associated with the pixel ϵ_x .

Let

$$\hat{w}^c = [\sin \eta_x, 0, \cos \eta_x]^T \quad (5)$$

be the unit vector from the center of mass to the inertial point $(q_N, q_E, 0)^T$ on the ground, expressed in camera coordinates. Note that \hat{w}^c is measured from camera coordinates. The vector \hat{w}^c is transformed into inertial coordinates using the transformation

$$\hat{w}^i = R_b^i R_g^b R_c^g \hat{w}^c, \quad (6)$$

where

$$R_b^i = \begin{pmatrix} c_\psi & -s_\psi & 0 \\ s_\psi & c_\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_\theta & 0 & s_\theta \\ 0 & 1 & 0 \\ -s_\theta & 0 & c_\theta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_\phi & -s_\phi \\ 0 & s_\phi & c_\phi \end{pmatrix} \quad (7)$$

is the transformation from body to inertial coordinates and $s_\varphi \triangleq \sin \varphi$ and

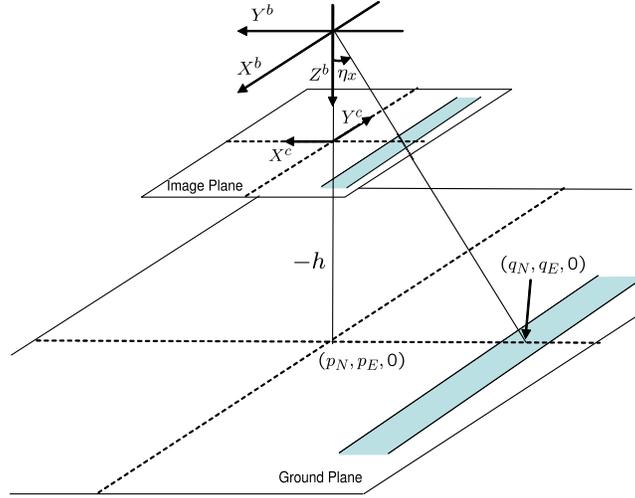


Figure 3: This figure depicts the camera geometry. The MAV is located at position $(p_N, p_E, -h)$. The world coordinates of the road are denoted by $(q_N, q_E, 0)$.

$c_\varphi \triangleq \cos \varphi$, and where

$$R_g^b = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad (8)$$

is the transformation from the gimbal to the body frame, and where

$$R_c^g = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad (9)$$

is the transformation from the camera to the gimbal frame. We include here a gimbal frame because the nominal orientation for the camera is out the nose of the aircraft, and in this paper the camera is assumed to point along the body frame z -axis. The fixed rotation R_c^g reorders the axes accordingly. Therefore assuming that $\theta \approx \gamma = 0$ we get

$$\hat{w}^i = \begin{pmatrix} -s_\psi c_\phi s_{\eta_x} + s_\psi s_\phi c_{\eta_x} \\ c_\psi c_\phi s_{\eta_x} - c_\psi s_\phi c_{\eta_x} \\ s_\phi s_{\eta_x} + c_\phi c_{\eta_x} \end{pmatrix}.$$

Normalizing the z -component results in the vector

$$\hat{d}^i = \begin{pmatrix} \frac{-s_\psi c_\phi s_{\eta_x} + s_\psi s_\phi c_{\eta_x}}{s_\phi s_{\eta_x} + c_\phi c_{\eta_x}} \\ \frac{c_\psi c_\phi s_{\eta_x} - c_\psi s_\phi c_{\eta_x}}{s_\phi s_{\eta_x} + c_\phi c_{\eta_x}} \\ 1 \end{pmatrix}.$$

Therefore, the inertial position of the road associated with pixel location $(\epsilon_x, 0)$ is given by

$$\begin{pmatrix} q_N \\ q_E \\ 0 \end{pmatrix} = \begin{pmatrix} p_N \\ p_E \\ -h \end{pmatrix} + h\hat{d}^i. \quad (10)$$

4.2 Roads composed of straight lines

In this section we assume that the road is composed of straight line segments and that the angle between subsequent segments is $\mu \in [-\pi/2, \pi/2]$, as shown in Figure 4. We will not address acute angles where $|\mu| > \frac{\pi}{2}$, because for acute angles, the future path of the road will be seen in the camera footprint before the turn, and will therefore enable predictive methods, which will be described in Section 4.5. Our objective is to derive roll-angle and AGL constraints so that, when the MAV encounters the turn in the road and maneuvers to track it, the road will be maintained in the camera footprint throughout the bank-to-turn maneuver. The location of the bend in the road is not known to the MAV a priori, and is estimated by the computer vision algorithm from the video stream.

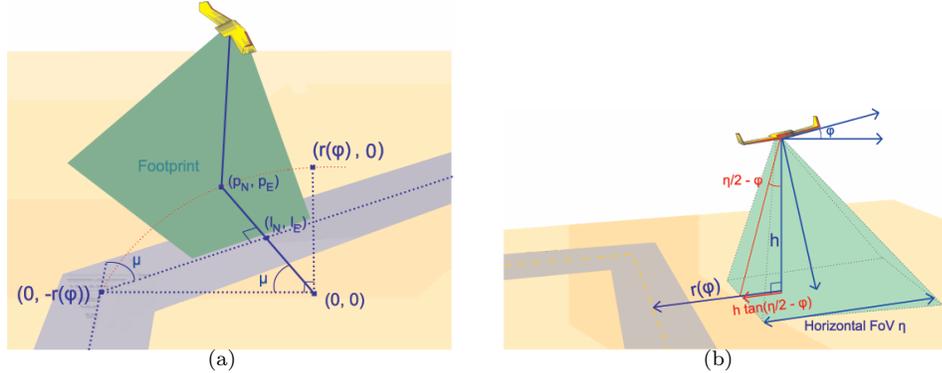


Figure 4: A depiction of the MAV following a road with a turn of angle $\mu \in [-\pi/2, \pi/2]$. Figure (a) depicts the MAV positioned where the tracking error is maximum. Figure (b) characterizes the geometric relationship between roll angle ϕ , AGL h , and the horizontal field of view $\bar{\eta}$.

Figure 4 depicts the MAV following a turn of magnitude μ . It is assumed that the flight path of the MAV is aligned with the road before it encounters the turn, and that the MAV is in trimmed flight. We also assume that the MAV maintains the same commanded roll angle ϕ throughout the turn, and that the change from a zero degree roll angle to the commanded roll angle is achieved instantaneously at the corner of the road. Assuming coordinated turn conditions, the radius of the flight path when the MAV maintains a constant

angle ϕ is given by

$$r(\phi) \triangleq \frac{v^2}{g \tan \phi}. \quad (11)$$

We ignore several factors that may impact the roll and AGL constraints including wind, initial tracking error, and roll rate constraints.

In Figure 4, the MAV is at the point of maximum deviation from the road while maintaining a constant commanded roll angle ϕ . Assuming without loss of generality, that the initial road segment is headed North, and choosing the center of the inertial coordinate system as the center of the turning circle, the position where the MAV's tracking error is maximum is given by

$$\begin{aligned} p_N &= r(\phi) \sin \mu, \\ p_E &= -r(\phi) \cos \mu. \end{aligned} \quad (12)$$

Denote the road coordinates closest to the MAV position as (l_N, l_E) , as shown in Figure 4(a). The equation for the road after the turn is

$$y = \frac{x}{\tan \mu} + \frac{r(\phi)}{\tan \mu}, \quad (13)$$

where y is the East coordinate and x is the North coordinate of the road. The distance between the road and the coordinates of the MAV (p_N, p_E) are given by

$$d^2 = (x - p_E)^2 + (y - p_N)^2 = (x - p_E)^2 + \left(\frac{x}{\tan \mu} + \frac{r(\phi)}{\tan \mu} - p_N\right)^2. \quad (14)$$

Maximizing (14) produces the closest road coordinates

$$\begin{aligned} l_N &= r(\phi) \sin \mu \cos \mu, \\ l_E &= -r(\phi) \cos^2 \mu. \end{aligned} \quad (15)$$

With reference to Figure 4, the distance between the MAV and (l_N, l_E) is $r(\phi)(1 - \cos \mu)$. Also, the distance between the MAV and the edge of the camera field of view is given by $h \tan(\frac{\bar{\eta}}{2} - \phi)$ where $\bar{\eta}$ is the field-of-view of the camera. Therefore, the road will remain in the camera field of view if

$$r(\phi)(1 - \cos \mu) \leq h \tan\left(\frac{\bar{\eta}}{2} - \phi\right). \quad (16)$$

Solving (16) for the AGL constraint gives

$$h(\phi, \mu) \geq \frac{v^2(1 - \cos \mu)}{g \tan \phi \tan\left(\frac{\bar{\eta}}{2} - \phi\right)}, \quad (17)$$

where g is the gravity constant. To derive a minimum AGL constraint, we first minimize the right-hand-side of Equation (17) with respect to ϕ , resulting in

the optimal roll angle $\phi^* = \frac{\bar{\eta}}{4}$. The ground AGL constraint is therefore given by

$$h^*(\mu) = \frac{v^2(1 - \cos \mu)}{g \tan^2(\frac{\bar{\eta}}{2})}. \quad (18)$$

The AGL constraint increases monotonically from $h^*(\mu = 0) = 0$ for a straight road, to the maximum AGL constraint $h^*(\mu = \frac{\pi}{2}) = v^2/g \tan^2(\bar{\eta}/2)$.

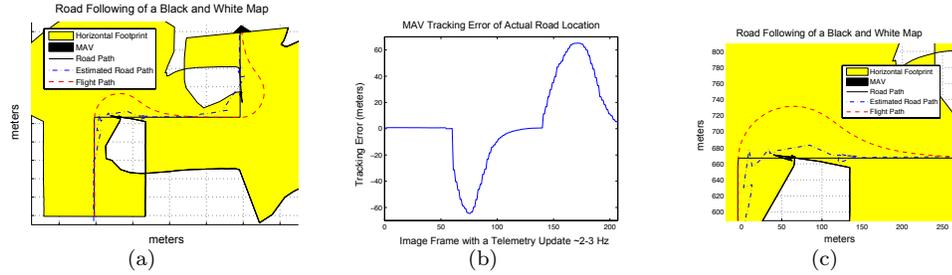


Figure 5: Simulation results for two subsequent turns of $\mu = \pi/2$. Figure (a) shows the MAV flight path, road pathway, and the region captured by the horizontal camera footprint through two turns of $\mu = \pi/2$. Figure (b) plots the MAV tracking error in relation to the actual road throughout the maneuver. Figure (c) shows a closeup of the footprint edge during the first turn.

To illustrate, Figure 5 depicts the flight path and camera footprint of a MAV flying at the the AGL and roll angle constraints when $\mu = \pi/2$. The groundspeed of the MAV is $v = 13$ m/s, and the field-of-view of the camera is $\bar{\eta} = \frac{\pi}{3}$, resulting in a roll angle constraint of $\phi^* = \frac{\pi}{12}$ and an AGL constraint of $h^* = 240$ m. As seen in Figure 5, the MAV successfully maintains the road in the camera throughout the two turns of $\mu = \pi/2$. In Figure 5 the shaded area is the locus of points swept out by the projection of the horizontal axis or x -axis of the camera frame onto the inertial frame.

To maintain the entire road in the camera footprint we must account for the road width. In this case, Equation (16) becomes

$$r(\phi)(1 - \cos \mu) + \frac{w_{road}}{2} \leq h \tan(\frac{\bar{\eta}}{2} - \phi), \quad (19)$$

where w_{road} is the (known) width of the road. The lower bound on AGL is therefore

$$h(\phi, \mu) \geq \frac{v^2(1 - \cos \mu) \pm g \tan \phi (w_{road}/2)}{g \tan \phi \tan(\frac{\bar{\eta}}{2} - \phi)}. \quad (20)$$

The minimum turn angle, and AGL constraints are therefore given by

$$\begin{aligned} \phi^* &= \arg \min_{0 \leq |\phi| \leq \bar{\phi}} h(\phi, \mu). \\ h^*(\mu) &= h(\phi^*, \mu). \end{aligned} \quad (21)$$

For this case, a closed form solution does not exist, and must be determined numerically. To solve this problem we have used the Matlab Optimization Toolbox.

4.3 Flying below the AGL constraint

Suppose that we select the commanded altitude based on an expected turn angles of μ , and then unexpectedly encounter sharper turns. Alternatively, if our primary objective is to image the road, the AGL constraint derived in the previous section may be too large to allow the desired level of camera resolution. Therefore, we may choose to fly below the AGL constraint resulting in the road leaving the footprint of the camera, requiring the road to be visually recaptured.

An obvious method for recapturing the road is to bank at the maximum bank angle $\pm\bar{\phi}$ until the road reenters the field of view. While this scheme results in a minimum time recapture, the resulting video will likely be erratic. In addition, the camera footprint may significantly deviate from the road, which may result in the vision system identifying a nearby road or structure as the path that needs to be followed.

As an alternative, we propose the introduction of a roll angle constraint that minimizes the distance between the camera footprint and the road while the road is not in the image. By doing so overhead imagery closest to the road, that may still contain useful information, will be obtained. The roll-angle constraint also provides a smoother viewing experience for the operator.

The minimum distance between the camera footprint and the road is determined from the closest road coordinates and the closest projected-pixel coordinates to the road. The projected-pixel location is obtained from Equation (10) as

$$\begin{aligned} q_N &= p_N + h \left(\frac{-s_\mu c_\phi s_{\frac{\eta}{2}} + s_\mu s_\phi c_{\frac{\eta}{2}}}{s_\phi s_{\frac{\eta}{2}} + c_\phi c_{\frac{\eta}{2}}} \right), \\ q_E &= p_E + h \left(\frac{c_\mu c_\phi s_{\frac{\eta}{2}} - c_\mu s_\phi c_{\frac{\eta}{2}}}{s_\phi s_{\frac{\eta}{2}} + c_\phi c_{\frac{\eta}{2}}} \right), \end{aligned} \quad (22)$$

where the MAV coordinates are given in Equation (12), and the road coordinates corresponding to the maximum deviation from the road are given in Equation (15). The roll angle that minimizes the distance between the camera footprint and the road for a specified AGL is found by finding the minimum distance between the closest road location given by Equation (15) and the closest localized pixel given by Equation (22) with respect to ϕ as

$$\phi^* = \arg \min_{0 \leq |\phi| \leq \bar{\phi}} ((q_N - l_N)^2 + (q_E - l_E)^2), \quad (23)$$

where $\bar{\phi}$ is the maximum roll angle allowed by the autopilot. The solution to Equation (23) is not necessarily $\bar{\phi}$. An aggressive roll angle will reduce the MAV turning radius but the camera footprint will sweep out less distance from the MAV. Equation (23) resolves the conflict between the MAV turn radius and

the footprint limit which are both functions of ϕ . The desired roll angle is solved numerically from Equation (23) and used as the roll-angle constraint in situations where the MAV operates below the AGL constraint.

Assuming a MAV speed of $v = 13$ m/s, a camera field of view $\bar{\eta} = \frac{\pi}{3}$, and a road width $w_{road} = 10$ m, the corresponding AGL and roll-angle constraint are shown in Figure 6. For situations where the MAV operates at the AGL constraint, the roll-angle constraint is set to $\frac{\bar{\eta}}{4}$, which guarantees the capture of the road. As shown in Figure 6(b), for turn angles between 0 and 55 degrees the MAV operates above the AGL constraint allowing it to conduct more aggressive turn maneuvers. For turn angles above 55 degrees the roll-angle constraint minimizes the distance between the road and the camera footprint.

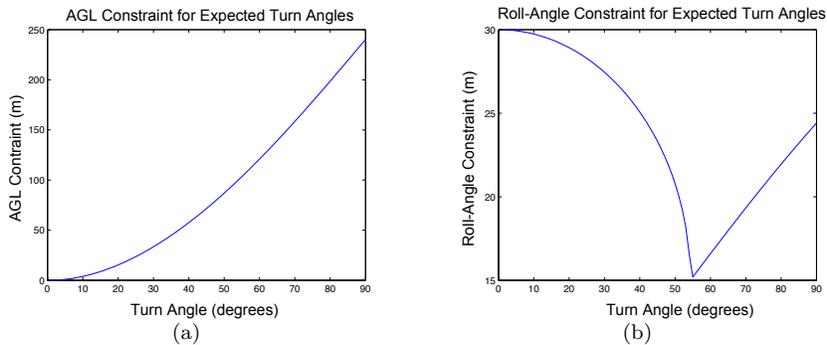


Figure 6: Figure (a) shows the relationship between the turn angle μ and the AGL constraint for a MAV speed of 13 m/s, and a horizontal field of view $\bar{\eta} = \frac{\pi}{3}$. Figure (b) shows the relationship between μ and the roll-angle constraint when the MAV is at an AGL of 100 m. For turn angles less than 55 degrees, the MAV operates above the AGL constraint which allows more aggressive roll-angles. For turn angles greater than 55 degrees, the roll angle constraint minimizes the distance between the footprint and the road.

4.4 Curved Roads

This section considers visual tracking of curved roads with bounded spatial frequency. Without loss of generality we assume that the road is sinusoidally shaped, as shown in Figure 7, with a constant East coordinate and a North coordinate given by $\alpha \cos(\beta\omega)$, where α is the amplitude and β is the spatial frequency.

As depicted in Figure 4(b), when the MAV is in wings-level flight, the farthest distance the camera footprint captures to the right or left of the MAV is $q_E(\phi = 0) = h \tan \frac{\bar{\eta}}{2}$. Therefore, the road will remain in the camera footprint during straight, wings-level flight if the AGL and the amplitude of the deviation

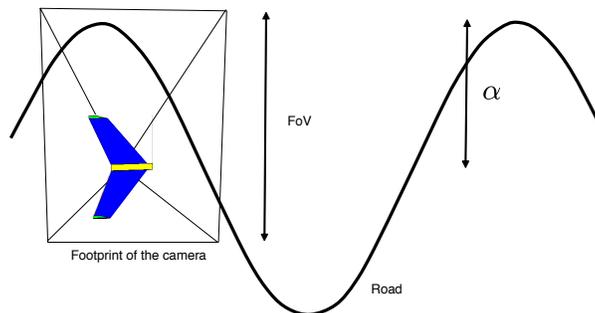


Figure 7: This figure depicts a MAV following a sinusoidal road. At the AGL derived in Equation (24), the camera footprint may cover the road without maneuvering.

satisfy

$$h \geq \frac{\alpha}{\tan \frac{\eta}{2}}. \quad (24)$$

Maintaining the road in the footprint of the camera is most difficult at the points of maximum curvature which occur when $\beta\omega = n\pi$, where n is an integer. By commanding a constant roll angle while tracking one of the periods of the sinusoidal road, a worst-case design can be obtained by approximating the MAV's flight path as a circle. The commanded roll angle will change during observation of the sinusoidal path to follow the road more tightly. Therefore, the roll-angle constraint that we will derive is conservative. Because the commanded roll angle is approximated as a constant, the camera observation limit of the footprint will also trace out a circular path as shown in Figure 8. For a righthanded orbit, the observation limit of the footprint is the inertial position of the rightmost pixel along the x -axis in the image.

For a dead-end, the MAV will need to turn 180 degrees without losing the road. Setting $\mu = \pi/2$ in Equation (18) gives the AGL constraint necessary to maintain the road through a quarter orbit. If the footprint is twice as large in the inertial frame, the MAV will successfully track the road through a half orbit. Therefore, the AGL constraint needed to maintain any inertial coordinate or road in the camera footprint is twice the AGL constraint for $\mu = \pi/2$:

$$h^* = \frac{2v^2}{g \tan(\frac{\eta}{4})^2}. \quad (25)$$

Flying at or above this AGL guarantees that any sinusoidally varying road will be in the camera footprint regardless of the road's curvature. Above this AGL, any continuous road can be maintained in the camera footprint by limiting the roll-angle constraint to $\phi \leq \frac{\eta}{4}$.

When flying below the AGL constraint (25), the roll-angle constraint ϕ^* is calculated using an approach similar to that used in Section 4.3. The road location with the greatest curvature ($\beta\omega = n\pi$) will dictate the maneuverability

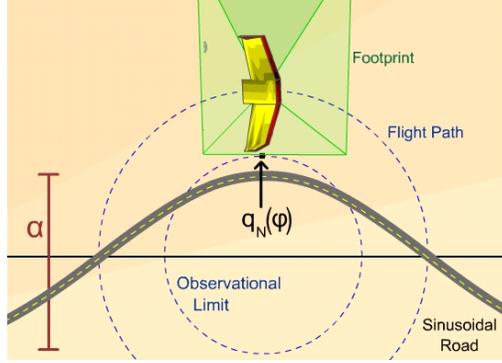


Figure 8: This figure depicts the MAV flight path, road path, and observation limit (the inertial position of the rightmost pixel). The maximum distance the MAV deviates from the road occurs at the locations of greatest curvature.

constraints, as shown in Figure 8. The road will be in the camera footprint if the curvature of the observation limit is greater than the curvature of the road. The observation circle is given by

$$y = (q_N(\phi)^2 - x^2)^{\frac{1}{2}}, \quad (26)$$

where x and y are the North and East coordinates of the observation limit, and where

$$q_N(\phi) = \frac{v^2}{g \tan \phi} + h \left(\frac{-c_\phi s_{\frac{\bar{\eta}}{2}} + s_\phi c_{\frac{\bar{\eta}}{2}}}{s_\phi s_{\frac{\bar{\eta}}{2}} + c_\phi c_{\frac{\bar{\eta}}{2}}} \right), \quad (27)$$

is the North location of the far right coordinate of the footprint.

Comparing the curvature of the road path and the observation limit yields

$$\frac{\partial^2}{\partial \omega^2} \alpha \cos(\beta \omega) = \frac{\partial^2}{\partial x^2} (q_N(\phi)^2 - x^2)^{\frac{1}{2}}. \quad (28)$$

where ω and x are both in units of meters. Evaluating the partial derivatives at $\beta \omega = n\pi$ and $x = 0$ we obtain $\beta^2 \alpha = \frac{1}{q_N(\phi^*)}$. Therefore, the road will remain in the camera footprint if $\beta^2 \alpha \leq \frac{1}{q_N(\phi^*)}$. The road width can be easily accounted for by adjusting the amplitude α to $\alpha_w = \alpha + w_{road}$ which gives

$$q_N(\phi^*) \leq \frac{1}{\beta^2 \alpha_w}. \quad (29)$$

Figure 9 shows the relationship between the roll-angle constraint and the AGL h , the road amplitude α , and the horizontal field-of-view $\bar{\eta}$. The road is a circle with curvature $\beta = \frac{\pi}{2\alpha} \text{ m}^{-1}$. Figure 9(a) plots the minimum roll-angle constraint as a function of AGL when the curvature amplitude is $\alpha = 140 \text{ m}$ and the field-of-view limit is $\bar{\eta} = 60$ degrees. Figure 9(b) plots the minimum roll-angle constraint as a function of curvature amplitude when the AGL is

$h = 100$ m and the field-of-view limit is $\bar{\eta} = 60$ degrees. Figure 9(c) plots the minimum roll-angle constraint as a function of field-of-view limit when the AGL is $h = 100$ m and the curvature amplitude is $\alpha = 140$ m.

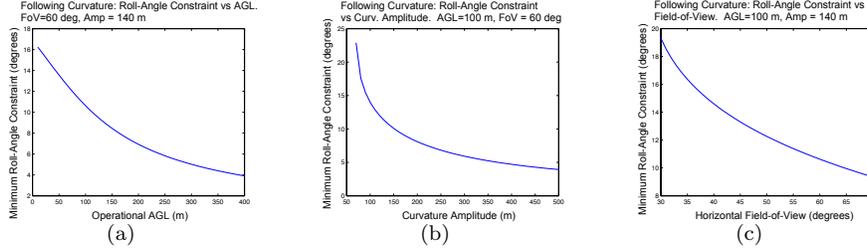


Figure 9: Minimum roll-angle constraint relationships for curved roads. These graphs show how the AGL h , the road curvature α , and the field-of-view constraint $\bar{\eta}$ influence the minimum roll-angle constraint. The MAV airspeed is 13 m/s and a road width $w_{road} = 0$ m.

In Figure 10 the road is constructed with circles of approximately 140 m radius. The MAV follows the road at an AGL of 100 m and a speed of 13 m/s. With these values, $\alpha = 140$ m and $\beta = \frac{\pi}{2\alpha} \text{ m}^{-1}$. Solving for the minimum roll angle constraint we get $\phi^* = 10.6$ degrees. The MAV follows the simulated curved road over 3000 m with an average tracking error of 7.49 m. The roll-angle constraint was set to 15 degrees to allow feedback control to track the road. Figure 10(a) shows the MAV flight path, path of the road, the estimated path of the road, and the horizontal footprint as it tracks the simulated road. The tracking error is shown in Figure 10(b), and the road-estimation error is shown in Figure 10(c).

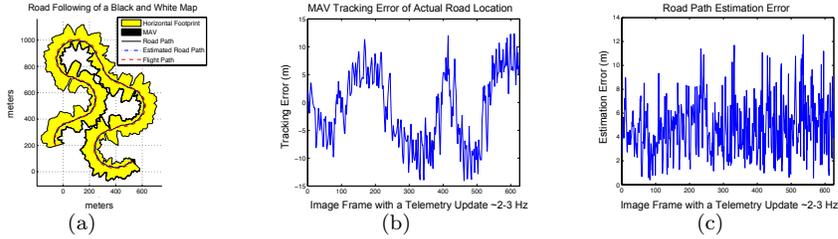


Figure 10: This figure presents the results of simulation of a MAV following a road consisting of ~ 140 m radius circles at an AGL h of 100 m and a roll-angle constraint ϕ^* set at 15 degrees over a distance of 3000 m.

4.5 Tracking roads whose position is known a priori

Knowledge of the road location and geometry may be available before flight using, for example, digital elevation maps obtained from Google Earth. Since the geometry of the path and the location of the turn is known, the MAV is able to anticipate the turn and, as shown in Figure 11, bank so that the road remains in the image. The known road scenario is significantly easier than the unknown road scenario since banking on the inside of the turn pushes the road into the camera footprint, whereas banking on the outside of the turn pushes the road out of the camera footprint.

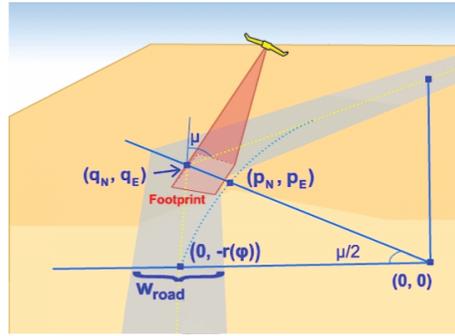


Figure 11: This figure depicts a MAV tracking a road whose turn angle μ and the inertial position of the turn are known a priori. In this scenario, the MAV is able to anticipate the turn and bank on the inside of the angle.

Figure 11 depicts the proposed MAV flight path. If our objective is simply to maintain the road in the camera field-of-view, it is not necessary to command a roll angle beyond the angle where the camera footprint captures the world horizon, which for a right-turn angle is a roll angle $\phi = \frac{\pi}{2} - \frac{\bar{\eta}}{2}$. If we desire to maintain the roll angle within the bound $|\phi| \leq \bar{\phi}$, then the roll angle constraint will be given by

$$\phi^* = \min \left\{ \left(\frac{\pi}{2} - \frac{\bar{\eta}}{2} \right), \bar{\phi} \right\}.$$

Using Equation (17), AGL constraint is given by

$$h^*(\mu) = \frac{v^2(1 - \cos \mu)}{g \tan \phi^* \tan(\frac{\bar{\eta}}{2} + \phi^*)}.$$

Even for an aggressive turn like $\mu = \frac{\pi}{2}$, the minimum AGL condition is trivial to achieve because the MAV maneuvers in a way to point the camera toward the road while it turns. The AGL required to track the road is only a few meters above the ground for a wide range of horizontal fields of view. With a horizontal field-of-view of $\bar{\eta} = 1$ degree, a maximum roll angle $\bar{\phi} = 30$ degrees, and a MAV speed $v = 13$ m/s, the minimum AGL constraint would be 3.9 m. Larger horizontal fields-of-view will result in an AGL constraint even closer to

the ground. Since the MAV will never approach the minimum AGL constraint, only a few meters above the ground, this constraint does not play a major role in the guidance algorithm.

However, for a fixed AGL, a roll-angle constraint can still be used to aid the road-following guidance loop. The roll angle constraint will depend on the desired flight path. The remainder of this section will derive three roll-angle constraints. The first constraint will center the bend of the road in the camera footprint. The second constraint will minimize distance between the MAV flight path and the road. The third constraint will minimize the tracking error between the road and the center of the camera footprint throughout the bank-to-turn maneuver. We will assume that constant control commands are provided to the MAV while it turns to ensure steady video imaging.

4.5.1 Roll Angle Constraint that Centers the Footprint on the Bend in the Road

This section derives a roll-angle constraints that centers the camera footprint on the bend in the road. The AGL is fixed at h . Figure 11 depicts the maximum distance the MAV travels from the road, where the MAV location is

$$\begin{aligned} p_N &= r(\phi) \sin\left(\frac{\mu}{2}\right) \\ p_E &= -r(\phi) \cos\left(\frac{\mu}{2}\right), \end{aligned} \quad (30)$$

and the heading angle is $\psi = \frac{\mu}{2}$, where $r(\phi)$ is given in Equation (11)

Image coordinates are calculated by using the transformation from the camera frame to the inertial frame. The case of centering the camera footprint precisely on the bend is mathematically equivalent to setting the horizontal field-of-view $\bar{\eta}$ equal to zero and minimizing the distance between the camera path and the bend coordinates. For a given AGL h , the inertial coordinates that correspond to the center of the camera footprint are

$$\begin{aligned} q_N &= p_N + h \sin\left(\frac{\mu}{2}\right) \tan \phi, \\ q_E &= p_E - h \cos\left(\frac{\mu}{2}\right) \tan \phi. \end{aligned} \quad (31)$$

The position of the bend in the road is given by

$$\begin{aligned} l_N &= \frac{-r(\phi)}{\tan \mu} + \frac{r(\phi)}{\sin \mu} \\ l_E &= -r(\phi). \end{aligned} \quad (32)$$

The roll angle ϕ^* is found by minimizing the distance between the world location corresponding to the center of the camera and the coordinates in the bend in the road as

$$\phi^* = \arg \min_{0 \leq |\phi| \leq \bar{\phi}} ((q_N - l_N)^2 + (q_E - l_E)^2). \quad (33)$$

The angle ϕ^* is used to obtain the coordinates where the MAV must begin its maneuver to capture the center of the bend. The MAV must begin its maneuver at a distance of $\frac{-r(\phi^*)}{\tan \mu} + \frac{r(\phi^*)}{\sin \mu}$ from the bend in the road. Using this path-planning scheme, the camera footprint will deviate from the road at the beginning and end of the maneuver, and at these points the road may leave the image depending on the field-of-view. The maneuver will center the road bend in the footprint at the midpoint of the maneuver. The road width does not effect the design of this constraint. Figure 12(a) shows the flight path and footprint following a 70 degree turn when the center of the footprint of the camera is constrained to pass through the center of the bend.

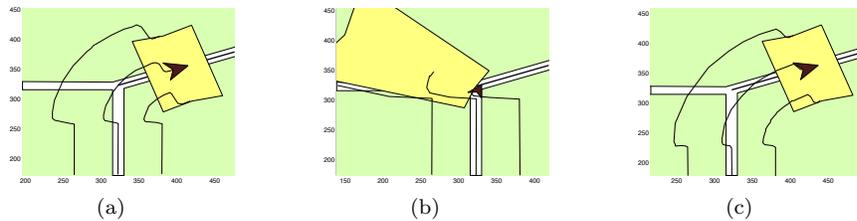


Figure 12: This figure shows simulation results of a MAV tracking a road whose position is known a priori with a 70 degree turn at 100 m AGL, airspeed $v = 13$ m/s, and horizontal field-of-view $\bar{\eta} = \frac{\pi}{3}$, using the three constraints discussed in this section. In Figure (a) the MAV executes a 11.04 degree roll angle 61.88 m before the road turn, centering the footprint on the bend in the road. In Figure (b) the MAV executes a 33.31 degree roll angle 18.37 m before the road turn, minimizing the distance between the road and the MAV. In Figure (c) the MAV executes a 7.4 degree roll angle 92.97 m before the road turn, minimizing the distance between the road and the center of the camera footprint throughout the maneuver.

4.5.2 Roll Angle Constraint that Minimizes the Distance Between MAV Flight Path and Road

As an alternative to centering the camera on the bend in the road, minimizing the distance between the MAV flight path and the road provides several benefits. For example, estimation errors from camera-to-world projection can be reduced by decreasing the MAV proximity to the road. Another benefit is that the pixel density of the road in the image will be maximized, thus helping the Road Classifier more effectively segment the road.

The position of the bend in the road is again given by Equation (32). As shown in Figure 11, for a right hand turn, the rightmost pixel constrains the MAV's ability to follow the road. For a road that bends to the right, the constraining pixel is located at $\eta_x = \frac{\bar{\eta}}{2}$. Using the result of Section 4.1, the

North and East inertial position corresponding to the rightmost pixel $(\frac{\bar{\eta}}{2}, 0)$ is

$$\begin{aligned} q_N &= p_N + h \left(\frac{-s_\mu c_\phi s_{\frac{\bar{\eta}}{2}} + s_\mu s_\phi c_{\frac{\bar{\eta}}{2}}}{s_\phi s_{\frac{\bar{\eta}}{2}} + c_\phi c_{\frac{\bar{\eta}}{2}}} \right), \\ q_E &= p_E + h \left(\frac{c_\mu c_\phi s_{\frac{\bar{\eta}}{2}} - c_\mu s_\phi c_{\frac{\bar{\eta}}{2}}}{s_\phi s_{\frac{\bar{\eta}}{2}} + c_\phi c_{\frac{\bar{\eta}}{2}}} \right). \end{aligned} \quad (34)$$

The coordinates from Equation (34) can be inserted into Equation (33) and the minimum roll angle ϕ^* can be found. A roll-angle constraint of $\phi = \phi^* + \epsilon$, where $|\epsilon|$ is less than the accepted roll-angle deviation, will allow the guidance loop to adjust to the path errors detected by the vision system. Figure 12(b) shows the flight path and sensor footprint when the MAV follows a 70 degree turn and the roll angle is constrained to minimize the distance to the road. Because the road is barely captured in the camera footprint at the center of the turn and it may initially be lost when the MAV begins the maneuver, this roll constraint should be the maximum limit for a given turn angle μ .

4.5.3 Roll Angle Constraint that Minimizes the Footprint Tracking Error over the Entire Road

Using a constraint that minimizes the tracking error between the entire road and the camera footprint helps the MAV capture the most area surrounding the road. This constraint is preferred if surveillance of the entire road is the desired objective. The constraints discussed in the preceding paragraphs focus on imaging the bend in the road, which may result in losing track of the road at the point where the MAV begins its maneuver. The inertial position corresponding with the center pixel of the camera is found using Equation (10) with $\bar{\eta} = 0$. From Figure 11, the area between the footprint center and the road during the maneuver can be used to determine the optimal roll-angle constraint. By considering the first half of the maneuver, we can find the tracking error by calculating the area between the East coordinates of the center of the footprint $q_E = -r(\phi)\cos(\psi) - h \cos \frac{\mu}{2} \tan \phi$, and the road before the bend $l_E = -r(\phi)$ as

$$\phi^* = \arg \min_{0 \leq |\phi| \leq \phi_{\max}} \int_{\psi=0}^{\psi=\frac{\mu}{2}} (-r(\phi) \cos \psi - h \cos \frac{\mu}{2} \tan \phi + r(\phi)) \delta \psi, \quad (35)$$

where ϕ_{\max} is the maximum roll angle. Figure 12(c) shows the flight path and footprint following a 70 degree turn where the roll angle is constrained to minimize the tracking area between the footprint and the road.

5 Simulation and Flight Results

5.1 Simulation Environment

The road-following algorithm was tested extensively in simulation in order to improve the guidance loop's robustness. The simulations tested the constraint

results as well as the entire implementation of the road-following algorithm. Simulation results were obtained using the Aviones simulator that was developed at BYU for small UAV research. The environment allows all flight algorithms to be tested in the lab before use in the field. Aviones incorporates a full physics model with six degrees of freedom. Aviones renders the MAV for display and provides a video stream to the Virtual Cockpit of the simulated environment captured from the onboard camera. BYU’s Virtual Cockpit software interfaces directly with Aviones, mimicking the hardware. The simulation software emulates the Kestrel autopilot by converting Rabbit processor code into standard C.

Figure 13 shows the relationship between the simulation environment and the flight test environment. In simulation, Aviones simulates the flight dynamics and the sensors, and produces synthetic images from the camera. In addition, the emulated Rabbit processor is used to estimate the states and control the MAV. The advantage of our framework is that the autopilot code, the virtual cockpit, and the frame processor are identical for both simulation and during the flight test.

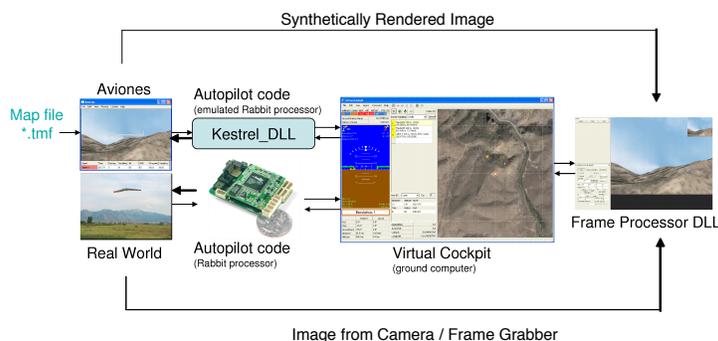


Figure 13: This diagram shows the relationship between software and hardware features in the MAV testbed. The Virtual Cockpit serves as the central hub to the flight system. Captured video is processed in the Frame Processor DLL. Through the Virtual Cockpit, the interpreted image data is sent to the controller located on the Autopilot.

5.2 Testbed Description

Figure 14 shows the key components of the testbed used by the Brigham Young University (BYU) Multiple AGent Intelligent Coordination and Control (MAG-ICC) Lab. Figure 14(a) shows the Procerus [22] Kestrel autopilot equipped with a Rabbit 3400 29 MHz processor, rate gyros, accelerometers, and absolute-pressure and differential-pressure sensors. The autopilot measures 2x1.37x0.47 inches and weighs 16.7 grams. Figure 14(b) shows the airframe used for the flight tests. The airframe is a 5 foot wingspan XS EPP foam flying wing se-

lected for its durability, ease of component installation, and flying characteristics. Embedded in the airframe are the Kestrel autopilot, batteries, a 1000 mW 900 MHz radio modem, a GPS receiver, a video transmitter, and an analog camera. Figure 14(c) shows the ground-station components. A laptop runs the Virtual Cockpit and the Frame Processor and interfaces with the MAV through a communication box. Video is transmitted to ground via a 2.4 GHz analog transmitter and captured on the laptop using an Imperx VCE-PRO PCMCIA frame grabber, which provides 640x480 images at 30 frames per second. An RC transmitter is used as a stand-by fail-safe mechanism to facilitate safe operations.

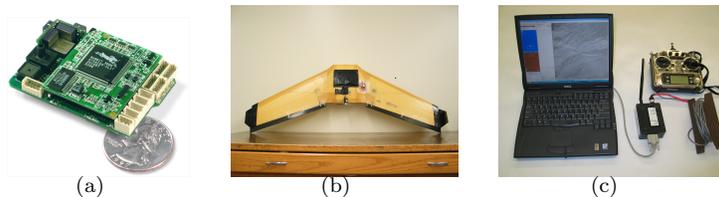


Figure 14: The figure shows the (a) Kestrel autopilot, (b) Kevlar-covered airframe, and (c) Ground-station components.

Telemetry information is recorded on the ground station, and guidance commands are sent to the MAV at 10 Hz. Since video is processed on the ground, the vision algorithm is only required to run at the modem-connection speed because road directions obtained from the image are localized to the inertial frame using the telemetry data. Therefore, the computer vision algorithm is designed to return data at 10 Hz. The inner loops that use accelerometers, rate gyros, and pressure sensors to control attitude, altitude, and airspeed run at approximately 100 Hz.

The road-following algorithm is based on the vector field method described in [23]. The MAV follows GPS waypoints until it reaches the road, switching to a road-following waypoint specified by the computer vision Road Direction Finder and Localizer shown in Figure 1. The road-following algorithm then follows the road continually updating the road waypoint using vision.

5.3 Flight Results

The road-following guidance strategy developed in this paper was used to successfully track over a mile of Goshen Canyon Road in Mona, Utah. The map of the road is shown in Figure 15. The aircraft flew at an altitude of 30 meters and at a velocity of 13 m/s.

The road was segmented from the image using the methods discussed in Section 3. The auto-gain on the camera was disabled so that the Image Classifier received consistent image hues from one frame to the next. A camera frame along with the corresponding classification result is shown in Figure 16.



Figure 15: This figure is a map of Goshen Canyon Road in Mona, UT. The curvature of the road has a maximum amplitude $\alpha = 400$ m so that the roll-angle constraint cannot be set below 5 degrees in order to follow the road at 70 m AGL.

As explained in Section 3, the original image is classified according to its H and V values. The resulting binary image is flood filled and then eroded. Finally, the smallest components are removed to obtain the classification. Using the telemetry information from the autopilot, the road location and its heading are estimated in the inertial frame.



Figure 16: This figure shows one camera frame captured during the road following demonstration and the corresponding result of the classification algorithm.

Figure 17 shows the path taken by the MAV with the estimation of the road position and overlays of the camera horizontal footprint. The estimation of the road position is less accurate when the MAV is banking because of estimation error in the roll angle. The MAV is able to successfully follow the road over long stretches and detailed imagery of the road is obtained throughout the flight. Wind causes the MAV to fly at a slight crab angle. The road possesses a curvature of amplitude $\alpha = 400$ m. With this curvature the MAV's roll-angle constraint cannot be set below 5 degrees.

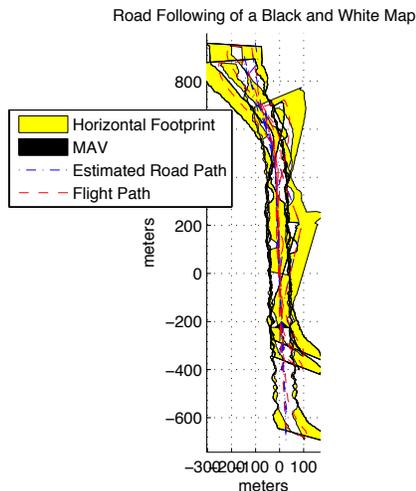


Figure 17: This figure shows the MAV flight path over Goshen Canyon Road, the estimation of the road path, and the overlay of the horizontal camera footprint. The overlay shows that the MAV was able to maintain the camera footprint over the road during the flight.

Figure 18 shows the road tracking error and the estimate of the road orientation during repeated road following experiments. To facilitate experimentation from a fixed ground station, the MAV is aligned with the road using GPS coordinates. When the road is in the camera footprint, the road following guidance loop is enabled. When the operator begins to lose visual line-of-sight to the MAV, a "turn around" signal is sent to the autopilot. Upon receiving this signal, the road following loop is disabled, and a 180 degree turn maneuver is executed using GPS guidance. After the turn, the road following loop is automatically re-enabled and road following resumes. The grey areas in Figure 18 correspond to turn-around maneuvers.

Figure 18(a) shows that the road tracking error consistently converges to a steady state error that is within a few meters of the road despite the presence of wind, road classification errors, camera-pointing errors, and communication latency. Figure 18(b) shows the estimate of the heading of the road produced by the vision system. The estimated heading is roughly 180 degrees when the MAV is tracking the road Southbound and approximately zero degrees when it is tracking the road Northbound.

Wind does cause some road-estimation error during flight. The Kestrel autopilot estimates steady state wind and MAV heading angle over the period of the flight by observing multiple samples of airspeed from the differential pressure sensor, groundspeed from GPS, and course angle from GPS. However, gusts and unsteady wind will cause errors in the heading angle used to localize the road. The MAV guidance loop will then follow the erroneous road direction.

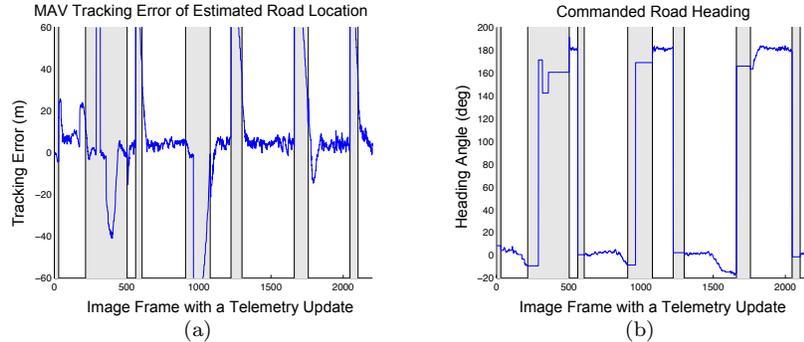


Figure 18: These figure show the tracking error and estimated heading direction of Goshen Canyon Road. The shaded regions correspond to turn-around maneuvers where the road following guidance loop was not active. Figure (a) shows the estimated tracking error which consistently converges to a small steady error. Figure (b) shows the estimated heading direction of Goshen Canyon road.

Additional errors are caused by estimation errors in the roll and pitch angles which are obtained by correcting rate gyro information with accelerometer measurements. However, despite these issues, the system operates reasonably well in windy conditions. Direct image based servoing may result in reduced tracking error and is the subject of another paper [24].

6 Conclusions and Future Work

This paper provides roll-angle and AGL constraints for a bank-to-turn MAV using a strap-down camera. Constraint derivations are derived for roads with sharp turns and roads with bounded curvature. Roll angle constraints are necessary for bank-to-turn MAVs because the roll-angle directly influences the shape and position of the camera footprint. The AGL determines the size of the footprint. The roll-angle and AGL constraints are derived to prevent the MAV from maneuvering in such a way that the road leaves the footprint of the camera.

The road-following guidance loop was shown to work in both simulation and flight tests. The road is successfully extracted from the image using statistical classification techniques in the HSV color space. The guidance loop can be altered to follow other visually distinct boundaries. For example, with forest fires, the Road Classifier can be adjusted to follow hot/cold perimeters rather than roads. In these other cases, the constraints derived in this paper are directly applicable.

The AGL and roll-angle constraints derived in this paper can be improved by considering the entire footprint and turn angles greater than 90 degrees. The pixels at the top of the image will relax the imposed constraints. These pixels provide a prediction element to the constraint problem which adjusts

the heading direction and increases the AGL limit. Turn angles greater than 90 degrees will also allow the MAV to predict future turns, thus relaxing the constraints. The current AGL and roll-angle solutions can also be relaxed by considering variable flight paths. Processing delay and tracking error between the MAV and the road will tighten the constraints.

The road-following guidance loop can be used in the development of other useful algorithms and mission planning. For example, autonomous vision based landing on a road is an obvious application. The guidance loop can also work in conjunction with tracking algorithms, where the MAV follows a road until it locates a target on the road and then switches to tracking the target.

Acknowledgements

This research was supported by NASA under STTR contract No. NNA04AA19C to Scientific Systems Company, Inc and Brigham Young University and by NSF award no. CCF-0428004. We would like to thank Josh Bodily for producing Figures 4, 7, 8 and 11. We also acknowledge the assistance of P.B. Sujit and Evan Anderson in obtaining the flight results. Portions of this work were performed while the second author was a National Research Council Fellow at the Air Force Research Laboratory, Munitions Directorate, Eglin AFB.

References

- [1] R. Wallace, K. Matsuzaki, Y. Goto, J. Crisman, J. Webb, and T. Kanade, "Progress in robot road-following," *IEEE International Conference on Robotics and Automation*, vol. 3, pp. 1615–1621, 1986.
- [2] D. Kuan, G. Phipps, and A. C. Hsueh, "Autonomous robotic vehicle road following," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 5, pp. 648 – 658, September 1988.
- [3] D. Song, H. N. Lee, J. Yi, and A. Levandowski, "Vision-based motion planning for an autonomous motorcycle on ill-structured road," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, October 2006.
- [4] S. Saripalli, J. F. Montgomery, and G. S. Sukhatme, "Vision-based autonomous landing of an unmanned aerial vehicle," *IEEE International Conference on Robotics and Automation*, May 2002.
- [5] O. Shakernia, R. Vidal, C. S. Sharp, Y. Ma, and S. Sastry, "Multiple view motion estimation and control for landing an unmanned aerial vehicle," *IEEE International Conference on Robotics and Automation*, May 2002.
- [6] B. Sinopoli, M. Micheli, G. Donato, and T. J. Koo, "Vision based navigation for an unmanned aerial vehicle," *IEEE International Conference on Robotics and Automation*, 2001.

- [7] J. Redding, T. McLain, R. Beard, and C. Taylor, "Vision-based target localization from a fixed-wing miniature air vehicle," *American Control Conference*, June 2006.
- [8] [Http://www.calccit.org/c3uv/](http://www.calccit.org/c3uv/).
- [9] E. Frew, T. McGee, Z. Kim, X. Xiao, S. Jackson, M. Morimoto, S. Rathinam, J. Padiyal, and R. Sengupta, "Vision-based road-following using a small autonomous aircraft," *IEEE Aerospace Conference*, March 2004.
- [10] S. Rathinam, Z. Kim, A. Soghikian, and R. Sengupta, "Vision based following of locally linear structures using an unmanned aerial vehicle," *IEEE Conference on Decision and Control, and the European Control Conference*, December 2005.
- [11] D. Lieb, A. Lookingbill, and S. Thrun, "Adaptive road following using self-supervised learning and reverse optical flow," in *Proceedings of Robotics: Science and Systems*, Cambridge, USA, June 2005.
- [12] J. Crisman and C. Thorpe, "Color vision for road following," *SPIE Conference on Mobile Robots*, November 1988.
- [13] C. Rasmussen and T. Korah, "On-vehicle and aerial texture analysis for vision-based desert road following," *IEEE Computer Vision and Pattern Recognition*, vol. 3, pp. 66–66, June 2005.
- [14] R. Wallace, "Robot road following by adaptive color classification and shape tracking," *IEEE International Conference on Robotics and Automation*, vol. 4, pp. 258–263, March 1987.
- [15] S. Todorovic and M. Nechyba, "A vision system for intelligent mission profiles of micro air vehicles," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 6, pp. 1713–1725, 2004.
- [16] J. Fauqueur and N. Boujemaa, "Progress in robot road-following," *16th Conference on Pattern Recognition*, vol. 3, pp. 1027–1030, 2002.
- [17] E. Saber, A. M. Tekalp, R. Eschbach, and K. Knox, "Automatic image annotation using adaptive color classification," *Graphical Models and Image Processing*, vol. 58, no. 2, pp. 115 – 126, March 1996.
- [18] D. Bader and J. Jaja, "Parallel algorithms for image histogramming and connected components with an experimental study," *Parallel and Distributed Computing*, vol. 35, pp. 173–190, 1996.
- [19] X. Lin and S. Chen, "Color image segmentation using modified HSI system for road following," *IEEE International Conference on Robotics and Automation*, April 1991.

- [20] J. Egbert and R. W. Beard, “Low altitude road following constraints using strap-down eo cameras on miniature air vehicles,” in *Proceedings of the American Control Conference*, New York City, July 2007, pp. 353–358.
- [21] D. R. Nelson, D. B. Barber, T. W. McLain, and R. W. Beard, “Vector field path following for small unmanned air vehicles,” *American Control Conference*, June 2006.
- [22] [Http://www.procerusuav.com/](http://www.procerusuav.com/).
- [23] D. R. Nelson, D. B. Barber, T. W. McLain, and R. W. Beard, “Vector field path following for miniature air vehicles,” *IEEE Transactions on Robotics*, vol. 37, no. 3, pp. 519–529, June 2007.
- [24] R. S. Holt and R. W. Beard, “Vision-based road-following using proportional navigation,” *IEEE Transactions on Control Systems Technology*, (in review).