Geolocation using video sensor measurements

Conference Paper · August 2007

DOI: 10.1109/ICIF.2007.4408156 · Source: IEEE Xplore

1 author:

Mahendra Mallick
Independent Consultant
119 PUBLICATIONS **1,545** CITATIONS

Geolocation using Video Sensor Measurements

Mahendra Mallick
Toyon Research Corporation
6800 Cortona Drive
Goleta, CA93117
mmallick@toyon.com

Abstract – Ground target tracking using electro-optical and infrared video sensors onboard unmanned aerial vehicles has drawn a great deal of interest in recent years due to the evolution of inexpensive video sensors and platforms. We present algorithms for geolocation using pixel location measurements which are based on the perspective transformation and includes radial and tangential lens distortions. The covariance of geolocation error takes into account the errors in pixel location, intrinsic and extrinsic camera parameters, and terrain height. Pixel coordinates of the optical center, focal distances along the X and Y axes, lens distortion parameters, and the skew parameter constitute the intrinsic camera parameters. The extrinsic camera parameters include the sensor position and sensor attitude relative a local east-north-up coordinate frame. Numerical results are presented using simulated data. Our results show that the errors in the Euler angles used to represent the sensor attitude is the dominant source of geolocation error.

Keywords: Aerial surveillance, ground target tracking, unmanned aerial vehicle, geolocation, video sensor, perspective transformation, lens distortion parameters, intrinsic and extrinsic camera parameters.

1 Introduction

Aerial surveillance [10], ground target tracking [2], and geolocation using video (electro-optical (EO) and infrared (IR)) sensors onboard unmanned aerial vehicles (UAVs) have drawn a great deal of interest in recent years. The video cameras and UAV platforms are relatively inexpensive compared with radar sensors and platforms. Therefore, multiple UAVs can be deployed in a given area of interest (AOI). In a video tracking problem, centroid pixel locations of moving targets in an image frame are extracted using video processing algorithms [6] which represent the kinematic measurements for the video tracking problem. Two types of video tracking can be performed; tracking in the image plane and tracking in a ground coordinate frame (GCF) using one or more video sensors. In the former case a single video sensor is used and position and velocity of a target are estimated in the image plane in pixels and pixels/s or pixels/frame. In the later case, position and velocity of a target are estimated relative to a GCF and camera intrinsic parameters (focal distances, pixel coordinates of the optical center, skew parameter, and lens distortion parameters) and extrinsic parameters (position and attitude) of the camera are used. The extrinsic camera parameters are obtained from the onboard global positioning system (GPS) and inertial navigation system (INS). The intrinsic camera parameters are estimated using camera calibration algorithms [17, 18].

Track initiation is an essential component of a tracking algorithm [3, 4]. Two types of track initiation algorithms, namely, the single-point [12] and two-point difference [3] track initiation algorithms are known. Given the pixel coordinates corresponding to an object point, camera parameters, and terrain data we can estimate the 3D location of the object point on the surface of the Earth and calculate the associated error covariance. This process is known as the geolocation for video sensor measurements. Once we have geolocation estimates, we can perform track initiation using the single-point [12] or two-point difference [3] algorithms.

The organization of the paper is as follows. Section 2 describes various coordinate frames used for the problem and Section 3 presents the video sensor measurement model which includes the perspective transformation, lens distortion, and video measurement error model. The video measurement error model addresses errors in pixel coordinates, sensor position, sensor attitude, and terrain height. Section 4 presents the geolocation algorithm which calculates the geolocation estimate and corresponding covariance using all error sources. Finally, Sections 6 and 7 present numerical results and conclusions.

2 Coordinate Frames and Convention

The geolocation algorithm uses a number of coordinate frames and considers transformations of 3-vectors among coordinate frames. We assume that all coordinate frames are right-handed and orthogonal unless specified otherwise. An uppercase roman letter (e.g. A) is used to denote the coordinate frame A. We use $\mathbf{b} \in \Re^3$ and $\mathbf{b}^{\mathrm{A}} \in \Re^3 := [b_x^{\mathrm{A}} \quad b_y^{\mathrm{A}} \quad b_y^{\mathrm{A}}]'$ to represent a 3-vector in a coordinate-free representation and in the A frame,

respectively, where (b_x^A, b_y^A, b_z^A) represent the Cartesian components of $\mathbf{b}^A$ along the *X*, *Y*, and Z axes of the A frame. Next, we describe the coordinate frames.

Topographic coordinate frame (T frame) - A topographic coordinate frame (T frame) is commonly used in ground target tracking where the origin of the T frame is chosen at a given geodetic longitude (λ_0), geodetic latitude (ϕ_0), and zero geodetic height $(h_0 = 0)$. The *X*, *Y*, and *Z* axes of the T frame are along the local East, North, and upward direction, respectively.

Sensor coordinate frame (S frame) - The origin of the S frame is at the *optical center* of the camera with geodetic coordinates (λ_s, ϕ_s, h_s). The *Z* axis is along the general downward direction along the optical axis. The *X* axis is to the right hand side of the image (Figure 1). The *Y* axis completes the right-handed coordinate frame. The image points and normalized image points are expressed in the S frame. We assume that an East-North-Up frame at the sensor position has the same orientation as the T frame. This is a valid assumption if the size of the tracking AOI is small compared with the radius of the earth.

Local North-East-Down coordinate frame (N frame) – The *X*, *Y*, and *Z* axes of the N frame are along the local North, East, and downward directions, respectively. If we are considering the N frame for the video sensor, then the origin of the N frame is on the reference ellipsoid with geodetic coordinates $(\lambda_s, \phi_s, 0)$.

Digital image coordinate frame (I frame) – The axes of the I frame and S frame are parallel but the origins are different. The origin of the I frame is at the upper left corner of the image and lies in the image plane (Figure 1). The *X* axis is to the right hand side of the image and the *Y* axis makes an angle θ, with the *X* axis, measured in the clockwise direction. If $\theta \neq \pi/2$, then a non-zero skew exists. The *Z* axis is parallel to the optical axis and, therefore, parallel to the Z axis of the S frame.

3 Video Sensor Measurement Model

The position vector of an object point from the T frame origin is given by

$$\mathbf{o}^T := \begin{bmatrix} o_x^T & o_y^T & o_z^T \end{bmatrix}' = \begin{bmatrix} x & y & z \end{bmatrix}', \tag{3-1}$$

The position of the sensor can be specified using the geodetic longitude (λ_s), geodetic latitude (ϕ_s), and geodetic height (h_s). Given the geodetic coordinates of the T frame origin and the geodetic coordinates of the S frame, we can calculate the position vector of the sensor $\mathbf{s}^T$ from the T frame origin, expressed in the T frame,

$$\mathbf{s}^T := \begin{bmatrix} s_x^T & s_y^T & s_z^T \end{bmatrix}' = \begin{bmatrix} x_s & y_s & z_s \end{bmatrix}'. \tag{3-2}$$

3.1 Perspective Transformation

We first consider the error-free case. Let $\mathbf{q}^S \in \Re^3$ denote the error-free image point corresponding to the object point $\mathbf{o}^T \in \Re^3$ and the sensor position vector $\mathbf{s}^T \in \Re^3$, as shown in Figure 1. Non-negligible terrain height variations can occur in certain cases for ground target tracking problems. In order to handle the terrain height variation, we use the *perspective transformation* [1, 7] for image formation. The simpler model, homography [7] is not suitable for handling terrain height variations.

Let (q_x^S, q_y^S) denote the *X* and *Y* coordinates of the image point in the image plane relative to the S frame corresponding to the object point $\mathbf{o}$. Then from Figure 1

$$\mathbf{q}^S := \begin{bmatrix} q_x^S & q_y^S & f \end{bmatrix}' \tag{3-3}$$

where *f* is the effective focal length of the camera. Figure 1 shows the relationship between the object and image points. From Figure 1

$$\mathbf{r}^S = \mathbf{o}^S - \mathbf{s}^S = \mathrm{T}_T^S(\mathbf{o}^T - \mathbf{s}^T), \tag{3-4}$$

where T_T^S is the 3×3 *sensor attitude matrix* relative to the T frame and is an orthogonal matrix.

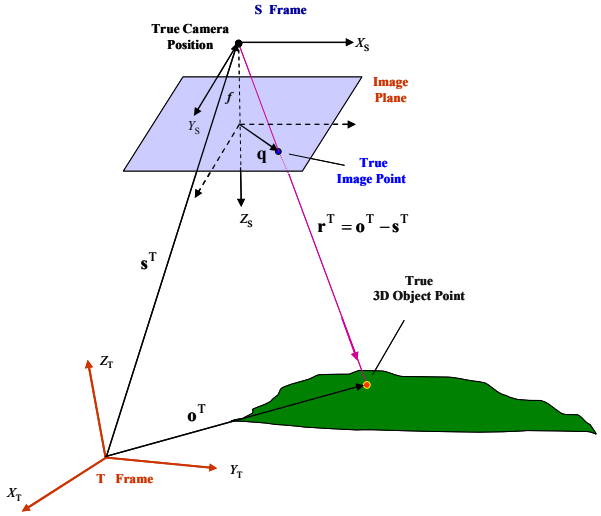

Figure 1. The orientation of the sensor frame (S frame) relative to the topographic coordinate frame (T frame). The true image point is derived from the true object point by the perspective transformation.

For a video sensor onboard an UAV, T_T^S is given by

$$\mathrm{T}_T^S = \mathrm{T}_N^S \mathrm{T}_T^N, \tag{3-5}$$

where

$$\mathrm{T}_T^N = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{3-6}$$

The attitude of the sensor $\mathrm{T}_{\mathrm{N}}^{\mathrm{S}}$ relative to the local N frame can be expressed using three Euler angles and an Euler sequence [16]. Using the (3, 2, 1) Euler sequence of Euler angles $\boldsymbol{\theta} \coloneqq (\theta_1, \theta_2, \theta_3)$ to represent $\mathrm{T}_{\mathrm{N}}^{\mathrm{S}}$ [16], we get

$$\mathrm{T}_{\mathrm{N}}^{\mathrm{S}}(\boldsymbol{\theta}) = \begin{bmatrix} \mathrm{c}\theta_2\mathrm{c}\theta_1 & \mathrm{c}\theta_2\mathrm{s}\theta_1 & -\mathrm{s}\theta_2 \\ -\mathrm{c}\theta_3\mathrm{s}\theta_1 + \mathrm{s}\theta_3\mathrm{s}\theta_2\mathrm{c}\theta_1 & \mathrm{c}\theta_3\mathrm{c}\theta_1 + \mathrm{s}\theta_3\mathrm{s}\theta_2\mathrm{s}\theta_1 & \mathrm{s}\theta_3\mathrm{c}\theta_2 \\ \mathrm{s}\theta_3\mathrm{s}\theta_1 + \mathrm{c}\theta_3\mathrm{s}\theta_2\mathrm{c}\theta_1 & -\mathrm{s}\theta_3\mathrm{c}\theta_1 + \mathrm{c}\theta_3\mathrm{s}\theta_2\mathrm{s}\theta_1 & \mathrm{c}\theta_3\mathrm{c}\theta_2 \end{bmatrix}, \tag{3-7}$$

where, for simplicity, we use the convention

$$\mathrm{c}\theta \coloneqq \cos\theta, \quad \mathrm{s}\theta \coloneqq \sin\theta. \tag{3-8}$$

The perspective transformation [1, 7] for the sensor geometry in Figure 1 is given by

$$\begin{bmatrix} q_x^{\mathrm{S}} \\ q_y^{\mathrm{S}} \end{bmatrix} = \frac{f}{r_z^{\mathrm{S}}} \begin{bmatrix} r_x^{\mathrm{S}} \\ r_y^{\mathrm{S}} \end{bmatrix}. \tag{3-9}$$

The normalized image point $\mathbf{q}_{\mathrm{n}}^{\mathrm{S}}$ is defined by [17]

$$\mathbf{q}_{\mathrm{n}}^{\mathrm{S}} \coloneqq \frac{\mathbf{q}^{\mathrm{S}}}{f} = [q_x^{\mathrm{S}}/f \quad q_y^{\mathrm{S}}/f \quad 1]'. \tag{3-10}$$

For simplicity in notation, define

$$\mathbf{x}_{\mathrm{n}} = \begin{bmatrix} x_{\mathrm{n}} \\ \mathrm{y}_{\mathrm{n}} \end{bmatrix} \coloneqq \begin{bmatrix} q_{\mathrm{n},x}^{\mathrm{S}} \\ q_{\mathrm{n},y}^{\mathrm{S}} \end{bmatrix} = \frac{1}{r_z^{\mathrm{S}}} \begin{bmatrix} r_x^{\mathrm{S}} \\ r_y^{\mathrm{S}} \end{bmatrix}. \tag{3-11}$$

Using (3-4), we get

$$x_{\mathrm{n}}(\mathbf{o}^{\mathrm{T}}, \mathbf{s}^{\mathrm{T}}, \mathrm{T}_{\mathrm{T}}^{\mathrm{S}}) = \frac{\mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(1,1)(o_x^{\mathrm{T}} - s_x^{\mathrm{T}}) + \mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(1,2)(o_y^{\mathrm{T}} - s_y^{\mathrm{T}}) + \mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(1,3)(o_z^{\mathrm{T}} - s_z^{\mathrm{T}})}{\mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(3,1)(o_x^{\mathrm{T}} - s_x^{\mathrm{T}}) + \mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(3,2)(o_y^{\mathrm{T}} - s_y^{\mathrm{T}}) + \mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(3,3)(o_z^{\mathrm{T}} - s_z^{\mathrm{T}})}, \tag{3-12}$$

$$y_{\mathrm{n}}(\mathbf{o}^{\mathrm{T}}, \mathbf{s}^{\mathrm{T}}, \mathrm{T}_{\mathrm{T}}^{\mathrm{S}}) = \frac{\mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(2,1)(o_x^{\mathrm{T}} - s_x^{\mathrm{T}}) + \mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(2,2)(o_y^{\mathrm{T}} - s_y^{\mathrm{T}}) + \mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(3,3)(o_z^{\mathrm{T}} - s_z^{\mathrm{T}})}{\mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(3,1)(o_x^{\mathrm{T}} - s_x^{\mathrm{T}}) + \mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(3,2)(o_y^{\mathrm{T}} - s_y^{\mathrm{T}}) + \mathrm{T}_{\mathrm{T}}^{\mathrm{S}}(3,3)(o_z^{\mathrm{T}} - s_z^{\mathrm{T}})}.$$

Examining (3-12) and (3-5)-(3-7), we note that the normalized image point depends on the object point $\mathbf{o}^{\mathrm{T}}$, sensor position $\mathbf{s}^{\mathrm{T}}$, and the Euler angles $(\theta_1, \theta_2, \theta_3)$. The sensor position $\mathbf{s}^{\mathrm{T}}$, and the Euler angles $(\theta_1, \theta_2, \theta_3)$ constitute the video camera *exterior parameter vector* $\boldsymbol{\eta}_{\mathrm{e}}$

$$\boldsymbol{\eta}_{\mathrm{e}} \coloneqq [(\mathbf{s}^{\mathrm{T}})' \quad \boldsymbol{\theta}']'. \tag{3-13}$$

3.2 Lens Distortion

Let $\mathbf{x}_{\mathrm{d}}$ denote the normalized image point including distortion

$$\mathbf{x}_{\mathrm{d}} \coloneqq \begin{bmatrix} x_{\mathrm{d}} \\ y_{\mathrm{d}} \end{bmatrix}. \tag{3-14}$$

Then

$$\mathbf{x}_{\mathrm{d}} = \beta_{\mathrm{radial}} \mathbf{x}_{\mathrm{n}} + \mathbf{d}, \tag{3-15}$$

where the radial lens distortion parameter β_{radial} and tangential lens distortion vector $\mathbf{d}$ are given by [5, 8, 17]

$$\beta_{\mathrm{radial}} = 1 + k_1\rho^2 + k_2\rho^4 + k_5\rho^6, \tag{3-16}$$

$$\rho^2 = x_{\mathrm{n}}^2 + y_{\mathrm{n}}^2, \tag{3-17}$$

$$\mathbf{d} = \begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} 2k_3x_{\mathrm{n}}y_{\mathrm{n}} + k_4(\rho^2 + 2x_{\mathrm{n}}^2) \\ k_3(\rho^2 + 2y_{\mathrm{n}}^2) + 2k_4x_{\mathrm{n}}y_{\mathrm{n}} \end{bmatrix}. \tag{3-18}$$

The X and Y components of $\mathbf{x}_{\mathrm{d}}$ in the S frame are

$$\begin{aligned} x_{\mathrm{d}} &= \beta_{\mathrm{r}}x_{\mathrm{n}} + d_x, \\ y_{\mathrm{d}} &= \beta_{\mathrm{r}}y_{\mathrm{n}} + d_y. \end{aligned} \tag{3-19}$$

Define the image distortion coefficient vector

$$\mathbf{k} \coloneqq [k_1 \quad k_2 \quad k_3 \quad k_4 \quad k_5]'. \tag{3-20}$$

The pixel coordinates (u, v) which include the effects of lens distortion and non-orthogonality of the U (X) and V (Y) axes of the digitized image frame (I frame) are given by [11, 17]

$$\begin{aligned} u &= f_u(x_{\mathrm{d}} + \alpha_{\mathrm{c}}y_{\mathrm{d}}) + u_0, \\ v &= f_v y_{\mathrm{d}} + v_0, \end{aligned} \tag{3-21}$$

where

f_u, f_v : focal distances along the U and V axes,

u_0, v_0 : U and V pixel coordinates of the optical center,

α_{c} : *skew coefficient* which equals the cotangent of the angle θ between the U and V axes of the digitized image, and

$$\alpha_{\mathrm{c}} = \cot\theta. \tag{3-22}$$

The focal distances are functions of the effective focal length f, scale factors (s_u, s_v) along the X and Y axes, and the angle θ,

$$f_u = f\,s_u, \qquad f_v = \frac{fs_v}{\sin\theta}. \tag{3-23}$$

We note that the model in (3-21) is purely deterministic. If the U and V axes are orthogonal, then the skew coefficient α_{c} is zero and (3-21) and (3-23), reduce, respectively to

$$\begin{aligned} u &= f_u x_{\mathrm{d}} + u_0, \\ v &= f_v y_{\mathrm{d}} + v_0, \end{aligned} \tag{3-24}$$

$$f_u = f\,s_u, \qquad f_v = fs_v. \tag{3-25}$$

Let $\boldsymbol{\eta}_{\mathrm{i}}$ denote the video camera *intrinsic parameter vector* defined by

$$\boldsymbol{\eta}_{\mathrm{i}} := [\mathbf{k}' \quad f_u \quad f_v \quad u_0 \quad v_0 \quad \alpha_{\mathrm{c}}]'. \tag{3-26}$$

Summarizing, the pixel coordinates which include the effects of perspective transformation, scaling, offset, skew, and lens distortion depend on a total number of sixteen camera parameters which define the *camera parameter vector* $\boldsymbol{\eta}$

$$\boldsymbol{\eta} := [\boldsymbol{\eta}_{\mathrm{e}}' \quad \boldsymbol{\eta}_{\mathrm{i}}']'. \tag{3-27}$$

The number of camera parameters for some cameras may be less than sixteen. For example, if the skew and tangential lens distortion are absent and the radial distortion effect is of fourth order, then there are twelve parameters.

3.3 Video Measurement Error Model

Now we consider the noisy pixel location represented by $\mathbf{z} := [z_u \quad z_v]'$ corresponding to the object point $\mathbf{o}^{\mathrm{T}}$. The noisy video camera measurement model based on the perspective transformation, lens distortion, scaling, offset, and skew effects is given by

$$\mathbf{z} = \begin{bmatrix} u \\ v \end{bmatrix} + \mathbf{n} = \mathbf{h}(\mathbf{o}^{\mathrm{T}}; \mathbf{s}^{\mathrm{T}}, \mathrm{T}_{\mathrm{T}}^{\mathrm{S}}, \boldsymbol{\eta}) + \mathbf{n}, \tag{3-28}$$

where $\mathbf{n}$ is the zero-mean Gaussian measurement noise with covariance $\mathbf{R}$

$$\mathbf{n} \sim N(0, \mathbf{R}) \tag{3-29}$$

and $\mathbf{h}$ is the nonlinear measurement function of the object point, sensor position, sensor attitude matrix, and intrinsic camera parameter vector.

In practice, the true sensor position vector $\mathbf{s}^{\mathrm{T}}$ and Euler angles $\boldsymbol{\theta}$ are not known. Rather, the measured sensor position $\hat{\mathbf{s}}^{\mathrm{T}}$ and Euler angles $\hat{\boldsymbol{\theta}} := (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$ are available from the GPS and INS which contain errors $\tilde{\mathbf{s}}^{\mathrm{T}}$ and $\tilde{\boldsymbol{\theta}}$, respectively. We assume that $\tilde{\mathbf{s}}^{\mathrm{T}}$ and $\tilde{\boldsymbol{\theta}}$ are zero-mean Gaussian with covariances $\mathbf{R}_{\tilde{\mathbf{s}}^{\mathrm{T}}}$ and $\mathbf{R}_{\tilde{\boldsymbol{\theta}}}$, respectively. Thus, the error models for the sensor position and Euler angles are described by

$$\hat{\mathbf{s}}^{\mathrm{T}} = \mathbf{s}^{\mathrm{T}} + \tilde{\mathbf{s}}^{\mathrm{T}}, \quad \tilde{\mathbf{s}}^{\mathrm{T}} \sim N(0, \mathbf{R}_{\tilde{\mathbf{s}}^{\mathrm{T}}}), \tag{3-30}$$

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \tilde{\boldsymbol{\theta}}, \quad \tilde{\boldsymbol{\theta}} \sim N(0, \mathbf{R}_{\tilde{\boldsymbol{\theta}}}). \tag{3-31}$$

The intrinsic parameters can be estimated quite accurately [17, 18] using the least squares estimator (LSE). Therefore, it is a reasonable approximation to ignore the errors in the estimated intrinsic parameter vector $\hat{\boldsymbol{\eta}}_{\mathrm{i}}$.

4 Geolocation

Given a measured pixel location $\mathbf{z}$, sensor position $\hat{\mathbf{s}}^{\mathrm{T}}$ and Euler angles $\hat{\boldsymbol{\theta}}$, and estimated intrinsic camera parameter vector $\hat{\boldsymbol{\eta}}_{\mathrm{i}}$, our objective is to estimate the geolocation $\hat{\mathbf{o}}^{\mathrm{T}}$ and the covariance $\mathbf{P}_{\tilde{\mathbf{o}}^{\mathrm{T}}}$ of geolocation error $\tilde{\mathbf{o}}^{\mathrm{T}}$. The covariance $\mathbf{P}_{\tilde{\mathbf{o}}^{\mathrm{T}}}$ must depend on $\mathbf{R}_{\tilde{\mathbf{s}}^{\mathrm{T}}}$, $\mathbf{R}_{\tilde{\boldsymbol{\theta}}}$, $\mathbf{R}$, and the variance $\sigma^2_{\mathrm{terrain}}$ of terrain height estimation error.

Given the measured pixel location $\mathbf{z}$, we first obtain an estimate $\hat{\mathbf{x}}_{\mathrm{n}}$ of the normalized image point $\mathbf{x}_{\mathrm{n}}$. Once we know $\hat{\mathbf{x}}_{\mathrm{n}}$, then we can obtain an estimate $\hat{\mathbf{u}}^{\mathrm{S}}$ of the unit vector along the line-of-sight (LOS) vector $\mathbf{u}^{\mathrm{S}}$. We can estimate the geolocation using the sensor position $\hat{\mathbf{s}}^{\mathrm{T}}$, the unit vector $\hat{\mathbf{u}}^{\mathrm{S}}$, and the terrain data. A flat earth model for terrain can also be used as an approximation, if the terrain is relatively flat in the tracking AOI.

4.1 Normalized Image Point Estimate $\hat{\mathbf{x}}_{\mathrm{n}}$

Given the measured pixel location **z,** estimates of $(x_{\mathrm{d}}, y_{\mathrm{d}})$ using (4-22) are given by [17]

$$\hat{y}_{\mathrm{d}} = (z_v - v_0)/f_v, \tag{4-1}$$

$$\hat{x}_{\mathrm{d}} = (z_u - u_0)/f_u - \alpha_{\mathrm{c}}\hat{y}_{\mathrm{d}}. \tag{4-2}$$

Then we use steps (1-5) to obtain the estimate $\hat{\mathbf{x}}_{\mathrm{n}}$.

Step 1. Initialize: $\hat{x}_{\mathrm{n}} = \hat{x}_{\mathrm{d}}, \quad \hat{y}_{\mathrm{n}} = \hat{y}_{\mathrm{d}}.$ (4-3)

Step 2. Calculate $\hat{\rho}^2 = \hat{x}_{\mathrm{n}}^2 + \hat{y}_{\mathrm{n}}^2.$ (4-5)

Step 3. Calculate $\hat{\beta}_{\mathrm{radial}}$ and $\hat{\mathbf{d}}$ by (4-17) and (4-19) using $(\hat{x}_{\mathrm{n}}, \hat{y}_{\mathrm{n}})$ and $\hat{\rho}^2$.

Step 4. Calculate

$$\begin{aligned} \hat{x}'_{\mathrm{n}} &= (\hat{x}_{\mathrm{d}} - d_x)/\beta_{\mathrm{radial}}, \\ \hat{y}'_{\mathrm{n}} &= (\hat{y}_{\mathrm{d}} - d_y)/\beta_{\mathrm{radial}}. \end{aligned} \tag{4-6}$$

Step 5. If $|\hat{x}'_{\mathrm{n}} - \hat{x}_{\mathrm{n}}| < \delta_{\mathrm{pix}}$ and $|\hat{y}'_{\mathrm{n}} - \hat{y}_{\mathrm{n}}| < \delta_{\mathrm{pix}}$, then
$(\hat{x}_{\mathrm{n}}, \hat{y}_{\mathrm{n}}) = (\hat{x}'_{\mathrm{n}}, \hat{y}'_{\mathrm{n}})$ represent the desired solution and exit,
Else set
$(\hat{x}_{\mathrm{n}}, \hat{y}_{\mathrm{n}}) = (\hat{x}'_{\mathrm{n}}, \hat{y}'_{\mathrm{n}})$ and go to Step 2.
End

4.2 Geolocation Estimate $\hat{\mathbf{o}}^{\mathrm{T}}$

Given $(\hat{x}_{\mathrm{n}}, \hat{y}_{\mathrm{n}})$, we calculate an estimate of $\mathbf{q}_{\mathrm{n}}^{\mathrm{S}}$ by

$$\hat{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}} = [\hat{x}_{\mathrm{n}} \quad \hat{y}_{\mathrm{n}} \quad 1]'. \tag{4-7}$$

An estimate of the unit vector along the LOS vector $\mathbf{u}^{\mathrm{S}}$ is given by

$$\hat{\mathbf{u}}^{\mathrm{S}} := \frac{\hat{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}}{[(\hat{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}})'\hat{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}]^{1/2}}. \tag{4-8}$$

Therefore, the geolocation estimate is

$$\hat{\mathbf{o}}^{\mathrm{T}} = \hat{\mathbf{s}}^{\mathrm{T}} + \hat{r}\hat{\mathbf{u}}^{\mathrm{T}}, \tag{4-9}$$

where $\hat{r}$ is the estimated range from the S frame origin to the estimated target position on the surface of the Earth and

$$\hat{\mathbf{u}}^{\mathrm{T}} = \hat{\mathrm{T}}_{\mathrm{S}}^{\mathrm{T}}\hat{\mathbf{u}}^{\mathrm{S}}, \qquad \hat{\mathrm{T}}_{\mathrm{S}}^{\mathrm{T}} = \mathrm{T}_{\mathrm{S}}^{\mathrm{T}}(\hat{\boldsymbol{\theta}}). \tag{4-10}$$

We can determine $\hat{r}$ using the terrain data and bisection algorithm [9, 13]. Alternately, we can determine $\hat{r}$ using the flat Earth approximation.

4.3 Covariance of Geolocation Error $\mathbf{P}_{\tilde{\mathbf{o}}^{\mathrm{T}}}$

The error-free target location is

$$\mathbf{o}^{\mathrm{T}}(r) = \mathbf{s}^{\mathrm{T}} + r\mathbf{u}^{\mathrm{T}}. \tag{4-11}$$

where

$$\mathbf{u}^{\mathrm{T}} = \mathrm{T}_{\mathrm{S}}^{\mathrm{T}}\mathbf{u}^{\mathrm{S}}. \tag{4-12}$$

Given the estimated geolocation in (4-9), the corresponding geolocation error is given by

$$\tilde{\mathbf{o}}^{\mathrm{T}}(r) = \tilde{\mathbf{s}}^{\mathrm{T}} + \hat{\mathbf{u}}^{\mathrm{T}}\tilde{r} + \hat{r}\tilde{\mathbf{u}}^{\mathrm{T}}. \tag{4-13}$$

From (4-12) we obtain

$$\tilde{\mathbf{u}}^{\mathrm{T}} = \tilde{\mathrm{T}}_{\hat{\mathrm{S}}}^{\mathrm{T}}(\boldsymbol{\theta})\mathbf{u}^{\mathrm{S}} + \mathrm{T}_{\hat{\mathrm{S}}}^{\mathrm{T}}\tilde{\mathbf{u}}^{\mathrm{S}} = \mathbf{C}'\tilde{\boldsymbol{\theta}} + \mathrm{T}_{\hat{\mathrm{S}}}^{\mathrm{T}}\tilde{\mathbf{u}}^{\mathrm{S}}, \tag{4-14}$$

where $\tilde{\mathrm{T}}_{\mathrm{S}}^{\mathrm{T}}(\boldsymbol{\theta})$ is the error in $\mathrm{T}_{\mathrm{S}}^{\mathrm{T}}(\boldsymbol{\theta})$ due to the error $\tilde{\boldsymbol{\theta}}$. We can show that

$$\tilde{\mathrm{T}}_{\hat{\mathrm{S}}}^{\mathrm{T}}(\boldsymbol{\theta})\mathbf{u}^{\mathrm{S}} = \mathbf{C}'\tilde{\boldsymbol{\theta}}, \tag{4-15}$$

where

$$\mathbf{C} := [\mathbf{c}_1 \quad \mathbf{c}_2 \quad \mathbf{c}_3], \tag{4-16}$$

$$\mathbf{c}_1 := \begin{bmatrix} -\mathrm{c}\theta_2\mathrm{s}\theta_1 u_1^{\mathrm{S}} - (\mathrm{c}\theta_3\mathrm{c}\theta_1 + \mathrm{s}\theta_3\mathrm{s}\theta_2\mathrm{s}\theta_1)u_2^{\mathrm{S}} + (\mathrm{s}\theta_3\mathrm{c}\theta_1 - \mathrm{c}\theta_3\mathrm{s}\theta_2\mathrm{s}\theta_1)u_3^{\mathrm{S}} \\ -\mathrm{s}\theta_2\mathrm{c}\theta_1 u_1^{\mathrm{S}} + \mathrm{s}\theta_3\mathrm{c}\theta_2\mathrm{c}\theta_1 u_2^{\mathrm{S}} + \mathrm{c}\theta_3\mathrm{c}\theta_2\mathrm{c}\theta_1 u_3^{\mathrm{S}} \\ (\mathrm{s}\theta_3\mathrm{s}\theta_1 + \mathrm{c}\theta_3\mathrm{s}\theta_2\mathrm{c}\theta_1)u_2^{\mathrm{S}} + (\mathrm{c}\theta_3\mathrm{s}\theta_1 - \mathrm{s}\theta_3\mathrm{s}\theta_2\mathrm{c}\theta_1)u_3^{\mathrm{S}} \end{bmatrix}, \tag{4-17}$$

$$\mathbf{c}_2 := \begin{bmatrix} \mathrm{c}\theta_2\mathrm{c}\theta_1 u_1^{\mathrm{S}} + (-\mathrm{c}\theta_3\mathrm{s}\theta_1 + \mathrm{s}\theta_3\mathrm{s}\theta_2\mathrm{c}\theta_1)u_2^{\mathrm{S}} + (\mathrm{s}\theta_3\mathrm{s}\theta_1 + \mathrm{c}\theta_3\mathrm{s}\theta_2\mathrm{c}\theta_1)u_3^{\mathrm{S}} \\ -\mathrm{s}\theta_2\mathrm{c}\theta_1 u_1^{\mathrm{S}} + \mathrm{s}\theta_3\mathrm{c}\theta_2\mathrm{s}\theta_1 u_2^{\mathrm{S}} + \mathrm{c}\theta_3\mathrm{s}\theta_2\mathrm{s}\theta_1 u_3^{\mathrm{S}} \\ (-\mathrm{s}\theta_3\mathrm{c}\theta_1 u_1^{\mathrm{S}} + \mathrm{c}\theta_3\mathrm{s}\theta_2\mathrm{s}\theta_1)u_2^{\mathrm{S}} - (\mathrm{c}\theta_3\mathrm{c}\theta_1 + \mathrm{s}\theta_3\mathrm{s}\theta_2\mathrm{s}\theta_1)u_3^{\mathrm{S}} \end{bmatrix}, \tag{4-18}$$

$$\mathbf{c}_3 := \begin{bmatrix} 0 \\ -\mathrm{c}\theta_2 u_1^{\mathrm{S}} + s\theta_3 s\theta_2 u_2^{\mathrm{S}} - \mathrm{c}\theta_3 s\theta_2 u_3^{\mathrm{S}} \\ \mathrm{c}\theta_3\mathrm{c}\theta_2 u_2^{\mathrm{S}} - s\theta_3 c\theta_2 u_3^{\mathrm{S}} \end{bmatrix}. \tag{4-19}$$

Using the small error approximation, we can show that

$$\tilde{\mathbf{u}}^{\mathrm{S}} = \frac{1}{\left\|\hat{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}\right\|}[\mathbf{I} - \hat{\mathbf{u}}^{\mathrm{S}}(\hat{\mathbf{u}}^{\mathrm{S}})']\tilde{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}, \tag{4-20}$$

where we use the approximations $\mathbf{u}^{\mathrm{S}} \approx \hat{\mathbf{u}}^{\mathrm{S}}$, $\mathbf{q}_{\mathrm{n}}^{\mathrm{S}} \approx \hat{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}$. Use of (4-20) in (4-14) gives

$$\tilde{\mathbf{u}}^{\mathrm{T}} = \mathbf{C}'\tilde{\boldsymbol{\theta}} + \mathbf{D}\tilde{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}, \tag{4-21}$$

where

$$\mathbf{D} := \frac{1}{\left\|\hat{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}\right\|}\mathrm{T}_{\hat{\mathrm{S}}}^{\mathrm{T}}[\mathbf{I} - \hat{\mathbf{u}}^{\mathrm{S}}(\hat{\mathbf{u}}^{\mathrm{S}})']. \tag{4-22}$$

Substitution of (4-21) in (4-13) gives

$$\tilde{\mathbf{o}}^{\mathrm{T}}(r) = \tilde{\mathbf{s}}^{\mathrm{T}} + \hat{\mathbf{u}}^{\mathrm{T}}\tilde{r} + \hat{r}(\mathbf{C}'\tilde{\boldsymbol{\theta}} + \mathbf{D}\tilde{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}). \tag{4-23}$$

The error $\tilde{r}$ is dependent on the error $\tilde{h}_{\mathrm{terrain}}$ in terrain height. We assume that the terrain height measurement error $\tilde{h}_{\mathrm{terrain}}$ is zero-mean Gaussian with variance $\sigma_{\mathrm{terrain}}^2$. Let α be the angle between the LOS vector $\hat{\mathbf{u}}^{\mathrm{T}}$ and the local vertical at the geolocation. Then

$$\tilde{r} \approx \cos\alpha\, \tilde{h}_{\mathrm{terrain}}. \tag{4-24}$$

Substitution of (4-24) in (4-23) gives

$$\tilde{\mathbf{o}}^{\mathrm{T}} = \tilde{\mathbf{s}}^{\mathrm{T}} + \cos\alpha\, \hat{\mathbf{u}}^{\mathrm{T}}\tilde{h}_{\mathrm{terrain}} + \hat{r}(\mathbf{C}'\tilde{\boldsymbol{\theta}} + \mathbf{D}\tilde{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}). \tag{4-25}$$

Since the errors $\tilde{\mathbf{s}}^{\mathrm{T}}$, $\tilde{h}_{\mathrm{terrain}}$, $\tilde{\boldsymbol{\theta}}$, and $\tilde{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}$ are assumed as zero-mean Gaussian, the geolocation error $\tilde{\mathbf{o}}^{\mathrm{T}}$ is zero-mean and Gaussian. Since the errors on the right hand side of (4-25) are assumed to be uncorrelated, the geolocation error covariance $\mathbf{P}_{\tilde{\mathbf{o}}^{\mathrm{T}}}$ is given by

$$\begin{aligned} \mathbf{P}_{\tilde{\mathbf{o}}^{\mathrm{T}}} &= \mathbf{R}_{\tilde{\mathbf{s}}^{\mathrm{T}}} + \cos^2\alpha\, \hat{\mathbf{u}}^{\mathrm{T}}(\hat{\mathbf{u}}^{\mathrm{T}})'\sigma_{\mathrm{terrain}}^2 \\ &\quad + \hat{r}^2(\mathbf{C}'\mathbf{R}_{\tilde{\boldsymbol{\theta}}}\mathbf{C} + \mathbf{D}\mathbf{R}_{\tilde{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}}\mathbf{D}'). \end{aligned} \tag{4-26}$$

4.4 Geolocation using the Flat Earth Approximation

We assume that the surface of the Earth in the tracking AOI is approximated by a plane whose height is h_0 above the *XY*-plane of the T frame. We have a measured value $\hat{h}_0$ of the height and we assume that the height measurement error $\tilde{h}_0$ is zero-mean Gaussian with

variance $\sigma^2_{\text{terrain}}$. Then from (2-1), for each object point $z = h_0$ and

$$\mathbf{o}^{\mathrm{T}} := \begin{bmatrix} o_x^{\mathrm{T}} & o_y^{\mathrm{T}} & o_z^{\mathrm{T}} \end{bmatrix}' = \begin{bmatrix} x & y & h_0 \end{bmatrix}'. \tag{4-26}$$

The estimated object position must satisfy (5-9) and taking the *Z* component of (4-9), we get

$$\hat{h}_0 = \hat{s}_z^{\mathrm{T}} + \hat{r}\hat{u}_z^{\mathrm{T}}. \tag{4-27}$$

Using the flat Earth approximation, we get from (4-27)

$$\hat{r} = (\hat{h}_0 - \hat{s}_z^{\mathrm{T}}) / \hat{u}_z^{\mathrm{T}}. \tag{4-28}$$

The geolocation error is still given by (4-13), where

$$\begin{aligned} \tilde{r} &= (\tilde{h}_0 - \tilde{s}_z^{\mathrm{T}}) / \hat{u}_z^{\mathrm{T}} - [(\hat{h}_0 - \hat{s}_z^{\mathrm{T}}) / (\hat{u}_z^{\mathrm{T}})^2]\tilde{u}_z^{\mathrm{T}} \\ &= a_1\tilde{h}_0 - a_1\tilde{s}_z^{\mathrm{T}} + a_2\tilde{u}_z^{\mathrm{T}}, \end{aligned} \tag{4-29}$$

and

$$a_1 = 1/\hat{u}_z^{\mathrm{T}}, \quad a_2 = (\hat{s}_z^{\mathrm{T}} - \hat{h}_0)/(\hat{u}_z^{\mathrm{T}})^2. \tag{4-30}$$

Substitution of (4-29) in (4-13) gives

$$\tilde{\mathbf{o}}^{\mathrm{T}} = \tilde{\mathbf{s}}^{\mathrm{T}} + \hat{\mathbf{u}}^{\mathrm{T}}(a_1\tilde{h}_0 - a_1\tilde{s}_z^{\mathrm{T}} + a_2\tilde{u}_z^{\mathrm{T}}) + \hat{r}\tilde{\mathbf{u}}^{\mathrm{T}}. \tag{4-31}$$

We have

$$\tilde{s}_z^{\mathrm{T}} = \mathbf{e}_3'\tilde{\mathbf{s}}^{\mathrm{T}}, \qquad \tilde{u}_z^{\mathrm{T}} = \mathbf{e}_3'\tilde{\mathbf{u}}^{\mathrm{T}}, \tag{4-32}$$

where $\mathbf{e}_3$ is a unit vector along the *Z* axis of the T frame. Use of (4-32) in (4-31) gives

$$\begin{aligned} \tilde{\mathbf{o}}^{\mathrm{T}} &= (\mathbf{I} - a_1\hat{\mathbf{u}}^{\mathrm{T}}\mathbf{e}_3')\tilde{\mathbf{s}}^{\mathrm{T}} + a_1\hat{\mathbf{u}}^{\mathrm{T}}\tilde{h}_0 + (\hat{r}\mathbf{I} + a_2\,\hat{\mathbf{u}}^{\mathrm{T}}\mathbf{e}_3')\tilde{\mathbf{u}}^{\mathrm{T}} \\ &= \mathbf{B}_{\tilde{\mathbf{s}}^{\mathrm{T}}}\tilde{\mathbf{s}}^{\mathrm{T}} + a_1\hat{\mathbf{u}}^{\mathrm{T}}\tilde{h}_0 + \mathbf{B}_{\tilde{\mathbf{u}}^{\mathrm{T}}}\tilde{\mathbf{u}}^{\mathrm{T}}, \end{aligned} \tag{4-33}$$

where

$$\mathbf{B}_{\tilde{\mathbf{s}}^{\mathrm{T}}} := (\mathbf{I} - a_1\hat{\mathbf{u}}^{\mathrm{T}}\mathbf{e}_3'), \quad \mathbf{B}_{\tilde{\mathbf{u}}^{\mathrm{T}}} := (\hat{r}\mathbf{I} + a_2\,\hat{\mathbf{u}}^{\mathrm{T}}\mathbf{e}_3'). \tag{4-34}$$

Substituting $\tilde{\mathbf{u}}^{\mathrm{T}}$ from (4-21) in (4-33), we get

$$\tilde{\mathbf{o}}^{\mathrm{T}} = \mathbf{B}_{\tilde{\mathbf{s}}^{\mathrm{T}}}\tilde{\mathbf{s}}^{\mathrm{T}} + a_1\hat{\mathbf{u}}^{\mathrm{T}}\tilde{h}_0 + \mathbf{B}_{\tilde{\mathbf{u}}^{\mathrm{T}}}[\mathbf{C}'\tilde{\boldsymbol{\theta}} + \mathbf{D}\tilde{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}]. \tag{4-35}$$

Then the geolocation error covariance $\mathbf{P}_{\tilde{\mathbf{o}}^{\mathrm{T}}}$ is given by

$$\begin{aligned} \mathbf{P}_{\tilde{\mathbf{o}}^{\mathrm{T}}} &= \mathbf{B}_{\tilde{\mathbf{s}}^{\mathrm{T}}}\mathbf{R}_{\tilde{\mathbf{s}}^{\mathrm{T}}}\mathbf{B}'_{\tilde{\mathbf{s}}^{\mathrm{T}}} + a_1^2\hat{\mathbf{u}}^{\mathrm{T}}(\hat{\mathbf{u}}^{\mathrm{T}})'\sigma^2_{\text{terrain}} \\ &\quad + \mathbf{B}_{\tilde{\mathbf{u}}^{\mathrm{T}}}(\mathbf{C}'\mathbf{R}_{\tilde{\boldsymbol{\theta}}}\mathbf{C} + \mathbf{D}\mathbf{R}_{\tilde{\mathbf{q}}_{\mathrm{n}}^{\mathrm{S}}}\mathbf{D}')\mathbf{B}'_{\tilde{\mathbf{u}}^{\mathrm{T}}}. \end{aligned} \tag{4-36}$$

5 Numerical Simulation and Results

We used intrinsic camera parameters similar to the estimated intrinsic camera parameters of the video camera onboard a Procerus™ UAV [15]. Table 1 presents the intrinsic and extrinsic camera parameters used in the simulation. We assume that the intrinsic camera parameters are error-free. Table 1 also presents the standard deviations of the sensor position and Euler angles that specify the sensor attitude.

We used the flat Earth approximation and chose object points on a 11×11 square grid with 10 meters spacing. We assume that the true height of the surface of the Earth relative to the *XY*-plane of the T frame is 100 meters. We used 500 Monte Carlo simulations to calculate various statistics of the geolocation estimate. For each Monte Carlo simulation, we used a measured value of the terrain height which has a standard deviation of three meters. The standard deviation of the pixel location measurement error for each component of (*u*, *v*) is three pixels. The standard deviations of all three Euler angles are equal in our simulation. A sample of the undistorted noise-free and distorted noisy pixels from the first Monte Carlo run is shown in Figure 2.

Table 1. Intrinsic and Extrinsic Camera Parameters

Parameter	Value
Intrinsic Camera Parameters	
Focal distances (pixels)	(548, 556)
Coordinates of the optical center (pixels)	(316.4, 223.0)
Skew coefficient	0.0
Image distortion coefficient vector $(k_1, k_2, ..., k_5)$	(-0.4500, 0.2200, 0.0005, 0.0005,0)
Extrinsic Camera Parameters	
Sensor position (m)	(100, 200, 350)
Standard deviation in sensor position (m)	(2.0, 2.0, 4.0)
Euler angles with (3,2,1) sequence (deg)	(150.0, 5.0, 3.0)
Standard deviation of Euler angles (deg)	(1, 2, 3, 4, 5)

In each Monte Carlo run, we calculated the geolocation estimate for each object point using (5-9) and (5-27), the actual estimation error using the truth data, and the geolocation covariance using (5-35). Then we calculated the mean square error matrix (MSEM) [3, 14] and average covariance using results from 500 Monte Carlo runs. Since the measurement model is nonlinear, the MSEM and average covariance matrix for an object point are not equal, in general. This is due to the bias inherent in all nonlinear estimators. We compare the root mean square (RMS) error (square root of the trace of the MSEM) and square root of the trace of the average covariance matrix (denoted by Sigma) in Figure 3 for the case where the standard deviation of each Euler angle is three degrees.

Results in Figure 3 show that the RMS error and Sigma are close to each other for all object points with the Sigma being slightly lower. This implies that the covariance calculated by (4-36) is slightly underestimated relative to the MSEM. This difference is due to the small bias present in the geolocation estimate as shown in Figure 4.

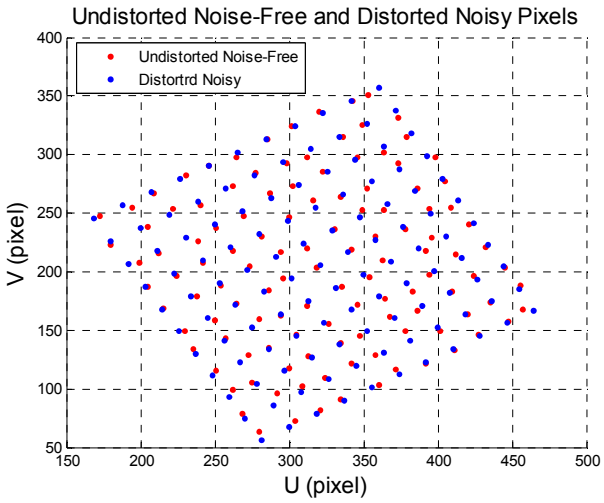

Figure 2. Undistorted noise-free and distorted noisy pixels from the first Monte Carlo run. The standard deviation of each Euler angle is three degrees.

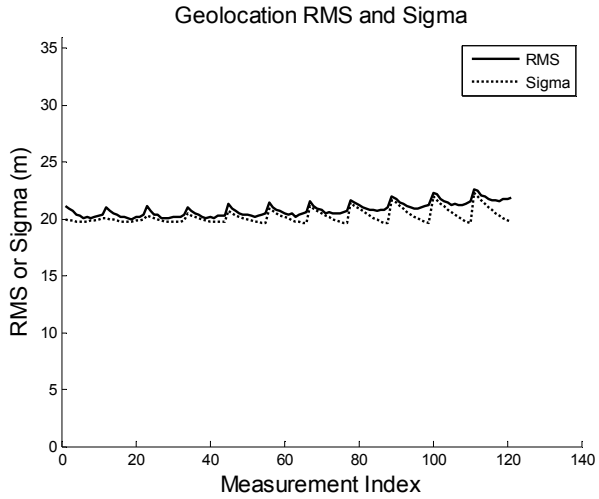

Figure 3. RMS geolocation error and square root of the trace of the average geolocation covariance (Sigma). Standard deviation of each Euler angle is three degrees.

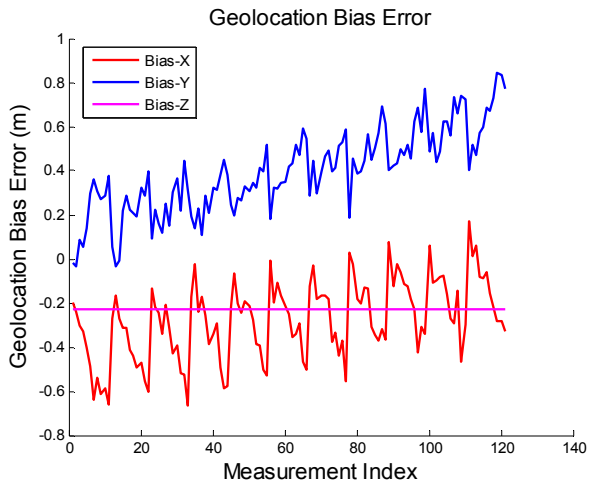

Figure 4. Bias in geolocation estimate.

In order to determine the contribution of each error source to the geolocation covariance in (4-36), we calculated individual terms in (4-36) in addition to the total contribution. Figure 5 presents the square root of the trace of the average covariance matrix for each error source and all error sources, termed as the geolocation standard deviation, for all object points. We observe that the sensor attitude error is the most dominant error source to the geolocation covariance and the pixel location error is the least contributing source. Figure 6 shows the RMS geolocation error with variation in the Euler angle standard deviation. The RMS geolocation error changes from about 10 meters to 35 meters as the Euler angle standard deviation changes from one degree to five degrees.

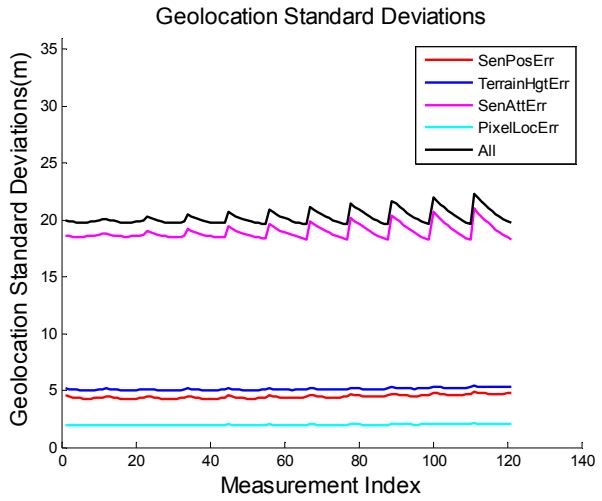

Figure 5. Square root of the trace of the average covariance matrix (Geolocation Standard Deviation) for individual and total contributions. The standard deviation of each Euler angle is three degrees for this case.

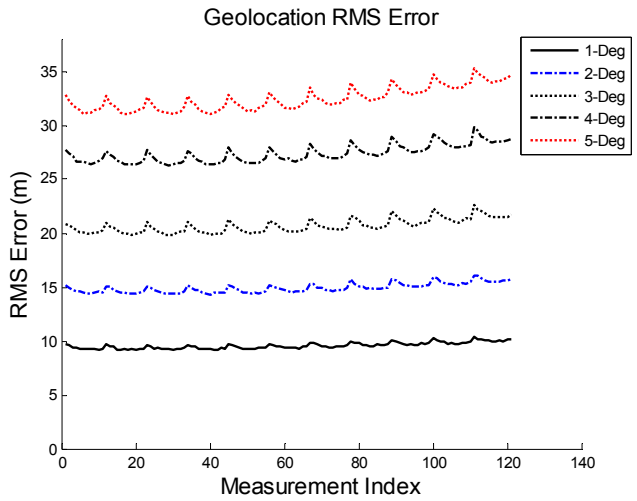

Figure 6. Variation of geolocation RMS error with Euler angle standard deviation.

6 Conclusion

We have developed two geolocation algorithms for UAV video sensor measurements which are based on the perspective transformation and include radial and tangential lens distortions. The first algorithm can be used in scenarios where significant terrain height variations can occur. Both algorithms model errors in sensor position, sensor orientation, terrain height, and measured pixel location. We have presented results using the second algorithm, which is based on the flat earth approximation.

Our results show that the average geolocation covariance for all object points considered is close to the MSEM and is slightly underestimated. Therefore, the geolocation covariance is representative of the actual error in the geolocation estimate. The errors in the measured Euler angles obtained from the INS are the most dominant sources of error while the pixel location errors are the least contributor. The RMS geolocation error varies from about 10 meters to 35 meters as the standard deviation in each Euler angle changes from one degree to five degrees.
Our future work will include results using actual terrain data and variations of standard deviations of other error sources.

Acknowledgement

The author thanks Nabil Alshurafa of Toyon Research Corporation for providing the estimated intrinsic camera parameters for the video camera used in a Procerus™ UAV.

References

[1] American Society of Photogrammetry, *Manual of Photogrammetry*, Fourth Edition, American Society of Photogrammetry, Falls Church, VA, 1980.

[2] P. Arambel, J. Silver, J. Krant, M. Antone, and T. Stratt, Multiple-hypothesis tracking of multiple ground targets from aerial video with dynamic sensor control, *Proc. SPIE, Signal Processing, Sensor Fusion, and Target Recognition XIII,* Vol. 5429, pp. 23-32, Orlando FL, 12-14 April 2004.

[3] Y. Bar-Shalom, X. R. Li and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, Wiley & Sons, 2001.

[4] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House, 1999.

[5] D.C. Brown, Close-Range Camera Calibration**,** *Photogrammetric Engineering*, pp. 855-866, Vol. 37, No. 8, 1971.

[6] A. P. Brown, K. J. Sullivan, and D. J. Miller, Feature-aided multiple target tracking in the image plane, *Proc. SPIE Defense and Security Symposium*, Orlando, FL, April 17-21, 2006.

[7] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision,* Second Edition, Cambridge University Press 2004.

[8] J. Heikkilä and O. Silvén,, A Four-step Camera Calibration Procedure with Implicit Image Correction*, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97),* San Juan, Puerto Rico, p. 1106-1112, 1997.

[9] D. Kahaner, C. Moler, S. Nash, *Numerical Methods and Software*, Prentice Hall, 1988.

[10] R. Kumar, H. Sawhney, S. Samarasekera, S. Hsu, H. Tao, Y. Guo, K. Hanna, A. Pope, R. Wildes, D. Hirvonen, M. Hansen, and P. Burt, Aerial video surveillance and exploitation, *Proc. IEEE*, Vol. 89, pp. 1518 – 1539, 2001.

[11] Q.-T. Luong and O. D. Faugeras, Self-calibration of a moving camera from point correspondences and fundamental matrices, *International Journal of Computer Vision,* Vol. 1, pp. 5-40, 1997.

[12] M. Mallick , T. Kirubarajan, and S. Arulampalam, "Comparison of Nonlinear Filtering Algorithms in Ground Moving Target Indicator (GMTI) Tracking," in *Proc. 2001 International Conference on Information Fusion*, August 2001, Montreal, Canada.

[13] M. Mallick, "Maximum Likelihood Geolocation using a Ground Moving Target Indicator (GMTI) Report," *Proc. 2002 IEEE Aerospace Conference*, March 2002, Big Sky MT, USA.

[14] M. Mallick and S. Arulampalam, "Comparison of Nonlinear Filtering Algorithms in Ground Moving Target Indicator (GMTI) Target Tracking," *Proc. Signal and Data Processing of Small Targets*, San Diego, CA, August 4-7, 2003.

[15] Procerus™ Technologies, http://procerusuav.com.

[16] J. R. Wertz, (Ed.), *Spacecraft Attitude Determination and Control*, D. Reidel Publising Company, Boston, 1978.

[17] http://www.vision.caltech.edu/bouguetj/calib_doc

[18] Z. Zhang, Flexible Camera Calibration By Viewing a Plane From Unknown Orientations, *Proc. International Conference on Computer Vision* (ICCV), pp. 666-673, Corfu, Greece, September 1999.