Unbiased and Consistent Electro-Optical Camera Angular Measurements With Cross-Correlated Errors and Their Fusion

J. K. Y. GOH
Y. BAR-SHALOM
R. YANG

Electro-optical (EO) camera systems are commonly used in target detection and tracking applications. Such camera systems typically comprise a suite of sensors, such as narrow/wide field of view (FOV) cameras, that provide target-originated angular measurements. To estimate the position of a point target in Cartesian space, existing techniques in literature employ the non-linear measurement mapping from the focal-plane array (FPA) to azimuth and elevation space. A common assumption made in using this conversion is that azimuth and elevation measurement errors have the same standard deviation, are uncorrelated, and are uniform across the camera's FOV. This paper presents an approach to derive the azimuth and elevation statistics, including the cross-correlation of their errors. This approach converts the raw target measurements and their covariance in the image space (FPA) to the angular space for subsequent use in Cartesian state filtering. This conversion has been validated to be unbiased and consistent, and results show that the line-of-sight (LOS) angle error variances and their correlations are in fact variable, with magnitudes dependent on the target's location in the FPA. The correct LOS angle covariance matrices should be used in Cartesian state estimation and fusion rather than the assumed constant angle variances and uncorrelated errors between the azimuth and elevation. We demonstrate a multi-sensor fusion case where the LOS angle covariance matrices of our proposed approach are used to derive the final composite target position estimate and its corresponding error covariances. The composite estimates produced from our proposed approach are proven to be statistically efficient. Compared to the use of the uncorrelated and constant LOS angle covariances, there is significant improvement to the error modeling of the fused Cartesian position covariance.

Manuscript received January 5, 2024; revised June 3, 2024; released for publication October 14, 2024

Refereeing of this contribution was handled by Stefano Coraluppi and Ramona Georgescu.

J. K. Y. Goh and R. Yang are with the DSO National Laboratories, Singapore 118225 (e-mail: gkoonyan@dso.org.sg, yrong@dso.org.sg).

Y. Bar-Shalom is with the Department of ECE, University of Connecticut, Storrs, CT 06269, USA (e-mail: yaakov.bar-shalom@uconn.edu).

1557-6418/2024/$17.00

I. INTRODUCTION

The use of electro-optical (EO) camera sensors in target detection and tracking applications has garnered much attention and intensive research in recent years. Their wide range of applications include the deployment of ground-based camera systems for aerial surveillance [12], camera sensors on-board aircraft and unmanned aerial platforms for ground target tracking [1], [6], or even for self-navigation and collision avoidance in the case of unmanned platforms [8], [9]. The many operational advantages of using camera sensors include their covertness due to them being passive, high accuracy of the angular measurements provided down to milliradians, and high data rate up to 30 Hz [7].

In video tracking, a sequence of images captures a moving target across time frames in the camera's field of view (FOV). Image processing algorithms are applied to each image to extract the centroid location of a small target in the focal-plane array (FPA) [6]. For the purpose of this paper, we focus on the conversion from FPA locations to 3D line-of-sight (LOS) angles (azimuth and elevation) in Cartesian coordinates centered at the camera. The target position and velocity can be estimated from these angular (azimuth/elevation) measurements derived from a series of coordinate transformations. Additional information is usually necessary (e.g., range or additional cameras) for the target state to be completely observable.

To track a target's position and velocity in Cartesian space, typical systems pass the angular (azimuth/elevation) measurements derived from the camera(s) and their covariances to a tracking filter. A common assumption made in the literature when using these measurements in a filter is that the azimuth and elevation measurement errors share a constant standard deviation, are uncorrelated, and are uniform across the camera's FOV. The above assumptions have been used for camera system designs that can be broadly classified into three categories: (i) a single camera sensor on a manoeuvring platform [1], [7], (ii) a camera suite comprising multiple (at least two) stationary camera sensors [11], and (iii) a sensor suite comprising a camera system with other sensor types such as radars [10]. For the sake of tracking filters performing at their best, it is essential that the measurements they use be unbiased and have correct covariances. In the case of the multisensor suite [10], data fusion is performed to derive the target's position and velocity, based on the combination of range information provided by other sensor types, such as radar, and the

highly precise angular information provided by the camera sensors.

Recent works have considered the modeling of correlation in measurement errors for camera sensor measurements, applied to the geolocation of targets. The authors in [6] formulated a camera measurement model as a nonlinear transformation function converting pixel location measurements in the image space to a 3D position in Cartesian space, based on the perspective transformation and including radial and tangential lens distortions. The covariance of the geolocation error was modeled to account for errors in pixel coordinates, intrinsic (lens distortion, skew, focal length) and extrinsic (sensor position, attitude) camera parameters, and terrain height. Results show that the geolocation covariance is representative of the actual error in the geolocation estimate. In addition, a two-camera setup to reconstruct the 3D position of the target using triangulation of the rays from corresponding image points in the left and right cameras was investigated in [4], with the total measurement error of the reconstructed point calculated using error propagation theory. The authors in [4] introduced a covariance to model the effect of the camera's orientation angle errors and the accuracy in detecting its corresponding image point.

While we recognize these approaches, the present paper's contributions are specific to camera sensors that are well calibrated with no radial and tangential distortion, and for tracking applications using angular measurements of point targets for the purpose of accurate fusion and subsequent filtering. Our approach converts the raw FPA target measurements and their covariance to the LOS angular space for subsequent use in fusion via the Maximum Likelihood approach. This conversion is shown to be unbiased and includes the variances and correlations of the measurement errors in azimuth and elevation that are in fact different at various image space (FPA) points, with magnitudes dependent on the target's location relative to the camera center.

The rest of the paper is organized as follows: Section II defines and describes the various coordinate systems used in this paper. Section III develops the conversion method from the image space to the global ENU coordinate space to derive unbiased measurements in azimuth and elevation angles, which are shown to have cross-correlated and location-dependent errors. Section IV considers the one-sensor case and presents a simulation test setup to illustrate the significance of the cross-correlated measurement errors, and proves the unbiasedness of the measurements in angular space and consistency of the derived covariance. Section V considers a two-sensor camera setup, and uses the proposed conversion approach to derive the angular error covariance matrices. The composite position measurement derived from the fusion of the cameras' LOS angles and their angular error covariances is shown to have covariance that meets the Cramer–Rao lower bound (CRLB). Lastly, Section VI presents the conclusions.

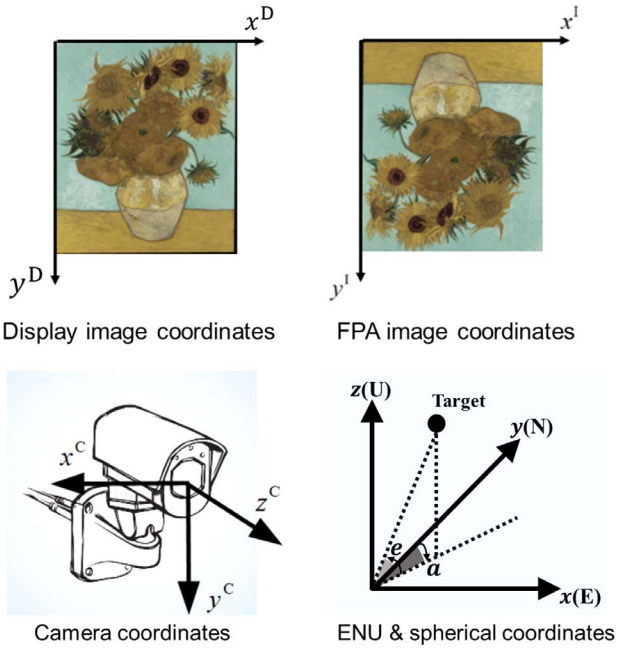

Figure 1. Display image, FPA image, camera, ENU, and spherical coordinate systems [13].

II. COORDINATE SYSTEMS

The following coordinate systems are used in this paper [13]:

- Display image coordinate system with x^D-y^D of the camera's display, shown in Fig. 1.
- Image coordinate system with x^I-y^I of the camera's FPA, shown in Fig. 1.
- Camera coordinate system with x^C-y^C-z^C centered at the camera position, shown in Fig. 1.
- Common coordinate system with x-y-z as east, north, and up (ENU) with origin at camera, with corresponding spherical coordinate system (Fig. 1) comprising:
  - azimuth (a): the clockwise angle in x(E)-y(N) plane from positive y axis to projection of target LOS onto this plane.
  - elevation (e): the angle from x(E)-y(N) plane to the target LOS.

The notations used in the paper are listed in Table I.

In practice, the flower image (on the left of Fig. 1) is shown right side up as in the camera display to correspond to the actual object. It is the resulting image after undergoing an inversion from the FPA, and the FPA image coordinates are related to the display image coordinates by

$$\mathbf{x}^I = \mathbf{h}_1\left(\mathbf{x}^D\right) = \begin{bmatrix} x^I \\ y^I \end{bmatrix} = \begin{bmatrix} P_x - x^D \\ P_y - y^D \end{bmatrix},$$
$$x^I = 0, 1, \ldots, P_x - 1, \quad y^I = 0, 1, \ldots, P_y - 1, \quad (1)$$

with P_x and P_y the (even) number of pixels in the x^I and y^I coordinates, respectively.

Table I
Notation

$\mathbf{x}^{\mathrm{A}}$	$[a \quad e]'$, the corresponding azimuth and elevation angles of a point in the common (ENU) coordinate system.
$\mathbf{x}$	$[x \quad y \quad z]'$, a point in the common (ENU) coordinate system.
$\mathbf{x}^{\mathrm{C}}$	$[x^{\mathrm{C}} \quad y^{\mathrm{C}} \quad z^{\mathrm{C}}]'$, a point in the camera coordinate system.
$\mathbf{x}^{\mathrm{I}}$	$[x^{\mathrm{I}} \quad y^{\mathrm{I}}]'$, a point in the image coordinate system in pixel units.
$\mathbf{x}^{\mathrm{D}}$	$[x^{\mathrm{D}} \quad y^{\mathrm{D}}]'$, a point in the display image coordinate system in pixel units.
$\mathbf{x}^{\mathrm{S}}$	$[x^{\mathrm{S}} \quad y^{\mathrm{S}} \quad z^{\mathrm{S}}]'$, the camera (sensor) position in ENU coordinates.
α	camera pointing azimuth or yaw (clockwise from N).
ϵ	camera pointing elevation or pitch (up from horizontal).
ρ	camera roll, clockwise around the center of the frame (ideally zero).

Next, considering the x^{C}-z^{C} plane and y^{C}-z^{C} plane given in Figs. 2 and 3, respectively, we have

$$\tan\left(\frac{\Theta_x}{2}\right) = \frac{\frac{P_x}{2}}{f} \quad \text{and} \quad \tan\left(\frac{\Theta_y}{2}\right) = \frac{\frac{P_y}{2}}{f}, \tag{2}$$

i.e.,

$$f = \frac{P_x}{2\tan\frac{\Theta_x}{2}} = \frac{P_y}{2\tan\frac{\Theta_y}{2}}, \tag{3}$$

where f is the focal length with units of measure pixel (assumed square); Θ_x and Θ_y are the FOV—angular spans—in x^{I} and y^{I}, respectively.

By similar triangles, from Fig. 2, one has

$$\frac{\frac{P_x}{2} - x^{\mathrm{I}}}{f} = \frac{x^{\mathrm{C}}}{z^{\mathrm{C}}} \quad \Rightarrow \quad x^{\mathrm{C}} = \frac{z^{\mathrm{C}}}{f}\left(\frac{P_x}{2} - x^{\mathrm{I}}\right), \tag{4}$$

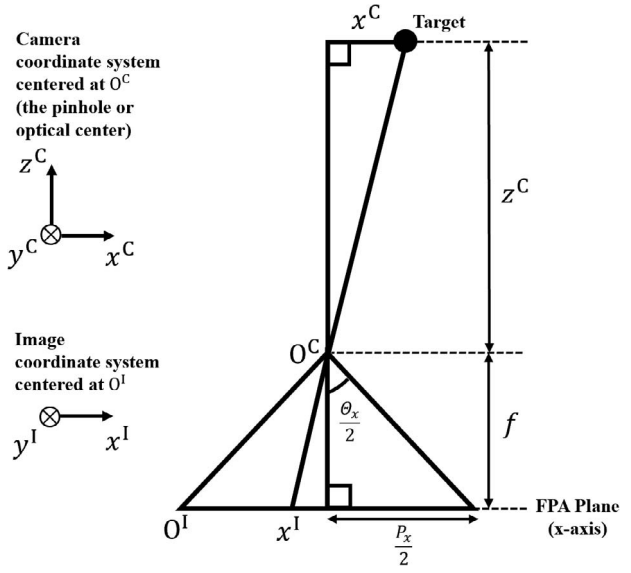

Figure 2. Diagram relating camera coordinate system to image coordinate system in x^{C}-z^{C} plane (top-down view so the FPA appears as a line); O^{C}, O^{I} denote the origins of the camera coordinate system and image coordinate system, respectively; $\otimes$ denotes the tail end of a coordinate system axis vector.

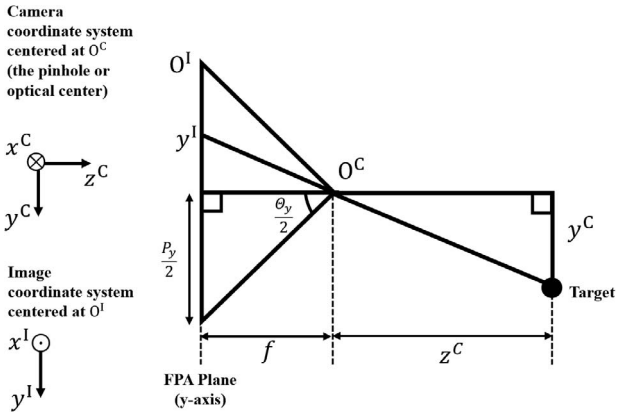

Figure 3. Diagram relating camera coordinate system to image coordinate system in y^{C}-z^{C} plane (side view so the FPA appears as a line); O^{C}, O^{I} denote the origins of the camera coordinate system and image coordinate system, respectively; $\otimes$, $\odot$ denote the tail and tip ends of a coordinate system axis vector, respectively.

and from Fig. 3,

$$\frac{\frac{P_y}{2} - y^{\mathrm{I}}}{f} = \frac{y^{\mathrm{C}}}{z^{\mathrm{C}}} \quad \Rightarrow \quad y^{C} = \frac{z^{\mathrm{C}}}{f}\left(\frac{P_y}{2} - y^{\mathrm{I}}\right). \tag{5}$$

By principle of the pinhole camera model, a point in the FPA corresponds to the projection of the 3D target point coordinates (via a line going through O^{C} as shown in Figs. 2 and 3) onto the image plane, and is the inverted image point of the target.

We further assume for convenience that

$$z^{\mathrm{C}} = f. \tag{6}$$

The LOS is then from O^{C} to the point $\left[x^{\mathrm{C}}, y^{\mathrm{C}}, z^{\mathrm{C}}\right]$ with z^{C} arbitrary since it does not affect the LOS. Then using appropriate notations, to be defined later, the coordinates of this point in the ENU system will be obtained, and they will yield the LOS angles in the ENU system.

Thus, combining equations (4)–(6), the conversion of $\mathbf{x}^{\mathrm{I}}$ to $\mathbf{x}^{\mathrm{C}}$ is given by

$$\mathbf{x}^{\mathrm{C}} = \mathbf{h}_2\left(\mathbf{x}^{\mathrm{I}}\right) = \begin{bmatrix} \frac{z^{\mathrm{C}}}{f}\left(\frac{P_x}{2} - x^{\mathrm{I}}\right) \\ \frac{z^{\mathrm{C}}}{f}\left(\frac{P_y}{2} - y^{\mathrm{I}}\right) \\ z^{\mathrm{C}} \end{bmatrix} = \begin{bmatrix} \frac{P_x}{2} - x^{\mathrm{I}} \\ \frac{P_y}{2} - y^{\mathrm{I}} \\ f \end{bmatrix}. \tag{7}$$

Next, the conversion of $\mathbf{x}^{\mathrm{C}}$ to $\mathbf{x} - \mathbf{x}^{\mathrm{S}}$ is given by

$$\begin{aligned} \mathbf{x} - \mathbf{x}^{\mathrm{S}} &= \mathbf{T}\left(\alpha, \epsilon, \rho\right)\mathbf{x}^{\mathrm{C}} \\ &= \mathbf{T}^{z^{\mathrm{C}}}(\alpha)\mathbf{T}^{x^{\mathrm{C}}}(90^{\circ} - \epsilon)\mathbf{T}^{z^{\mathrm{C}}}(-\rho)\mathbf{x}^{\mathrm{C}}, \end{aligned} \tag{8}$$

where α, ϵ, and ρ refer to the camera's yaw, pitch and roll angles, respectively, and we use the following mnemonic notations for rotations between 3D Cartesian systems

$$\mathbf{T}^{x^{\mathrm{C}}}(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix} \tag{9}$$

for rotation around the x^{C}-axis by ϕ from y^{C} toward z^{C}, and

$$\mathbf{T}^{z^{\mathrm{C}}}(\phi) = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{10}$$

for rotation around the z^{C}-axis by ϕ from x^{C} towards y^{C}. The rotation around the y^{C}-axis is not necessary as $\mathbf{T}^{x^{\mathrm{C}}}(90^\circ - \epsilon)$ replaces the y^{C}-axis by the z^{C}-axis, so that the rotation around the z^{C}-axis occurs twice. The combined rotation in (8) is

$$\mathbf{T}(\alpha, \epsilon, \rho) = \begin{bmatrix} s_\alpha s_\epsilon s_\rho + c_\alpha c_\rho & s_\alpha s_\epsilon c_\rho - c_\alpha s_\rho & s_\alpha c_\epsilon \\ c_\alpha s_\epsilon s_\rho - s_\alpha c_\rho & c_\alpha s_\epsilon c_\rho + s_\alpha s_\rho & c_\alpha c_\epsilon \\ -c_\epsilon s_\rho & -c_\epsilon c_\rho & s_\epsilon \end{bmatrix} \tag{11}$$

where

$$s_\alpha = \sin\alpha, \quad s_\epsilon = \sin\epsilon, \quad s_\rho = \sin\rho, \tag{12}$$

$$c_\alpha = \cos\alpha, \quad c_\epsilon = \cos\epsilon, \quad c_\rho = \cos\rho, \tag{13}$$

For a point $\mathbf{x}$ in the common ENU coordinate system, the conversion of $\mathbf{x} - \mathbf{x}^{\mathrm{S}}$ to $\mathbf{x}^{\mathrm{A}}$ is given by

$$\mathbf{x}^{\mathrm{A}} = \begin{bmatrix} a \\ e \end{bmatrix} = \mathbf{h}_3\left(\mathbf{x} - \mathbf{x}^{\mathrm{S}}\right) = \begin{bmatrix} \tan^{-1}\left(\frac{x - x^{\mathrm{S}}}{y - y^{\mathrm{S}}}\right) \\ \tan^{-1}\left(\frac{z - z^{\mathrm{S}}}{\sqrt{(x-x^{\mathrm{S}})^2 + (y-y^{\mathrm{S}})^2}}\right) \end{bmatrix}. \tag{14}$$

III. CONVERSION METHOD FOR ANGULAR MEASUREMENT ERRORS

A. Uncorrelated Measurement Error Covariance Assumption

When tracking targets' position and velocity in 3D Cartesian space, angular measurements of the target with respect to the ENU coordinate system are provided to the fuser and then to a tracking filter[1] for dynamic state estimation. The derived measurement vector from a camera sensor at time t_k is

$$\mathbf{x}^{\mathrm{A}}(t_k) = [a(t_k)\ e(t_k)]' \tag{15}$$

where $a(t_k)$ is the measured (noisy) azimuth from true North clockwise, and $e(t_k)$ is the measured elevation up from the horizontal, with reference to the camera position given by $\mathbf{x}^{\mathrm{S}}$. The corresponding measurement error covariance is typically assumed as having uncorrelated errors, i.e.,

$$\mathbf{R}^{\mathrm{A}}_{\mathrm{uncorr}} = \mathrm{diag}(\sigma_a^2, \sigma_e^2) \tag{16}$$

with σ_a and σ_e the measurement error standard deviations of azimuth and elevation, respectively. Practically, all filtering applications assume a constant measurement error standard deviation, same in both azimuth and elevation (i.e., $\sigma_a = \sigma_e$), and with uncorrelated errors between them.

B. The Measurement Error Covariance Matrix With Location Dependence and Cross-Correlation

We investigate the validity of the above assumption, whether the variances of the measurement errors are constant, and if measurement error correlation between azimuth and elevation is present. We present an approach to derive the errors of the angular measurements of the target LOS with respect to the ENU coordinate system. This involves calculating the azimuth and elevation measurements of the target (i.e., $\hat{\mathbf{x}}^{\mathrm{A}}$) from the raw target pixel measurements in the display image space (i.e., $\hat{\mathbf{x}}^{\mathrm{D}}$) through the series of transformations defined in Section II. The corresponding measurement error covariance matrix with location dependence and cross-correlated errors in azimuth-elevation space, denoted as $\mathbf{R}^{\mathrm{A}}_{\mathrm{corr}}$, is derived from the error covariance in the image space using the linearized transformation function. The validity of the first-order linearization is ascertained by confirming the unbiasedness of LOS angle errors and its covariance consistency [2], i.e., that the calculated covariance matrix is statistically compatible with the actual errors.

The target angular measurement $\hat{\mathbf{x}}^{\mathrm{A}}$ can be obtained by the global transformation function $\mathbf{h}(\cdot)$

$$\hat{\mathbf{x}}^{\mathrm{A}} = \mathbf{h}\left(\hat{\mathbf{x}}^{\mathrm{D}}\right) \triangleq \mathbf{h}_3\left[\mathbf{T}(\alpha, \epsilon, \rho)\, \mathbf{h}_2\left(\mathbf{h}_1(\hat{\mathbf{x}}^{\mathrm{D}})\right)\right], \tag{17}$$

with $\mathbf{h}_3(\cdot)$, $\mathbf{T}(\alpha, \epsilon, \rho)$, $\mathbf{h}_2(\cdot)$, and $\mathbf{h}_1(\cdot)$ defined in (14), (8), (7), and (1), respectively. Its corresponding covariance matrix (including correlations) in azimuth-elevation space is, based on first-order linearization (see, e.g., [2]), given by

$$\mathbf{R}^{\mathrm{A}}_{\mathrm{corr}} = HP^{\mathrm{D}}H' = H\begin{bmatrix} \sigma^2_{P_x} & 0 \\ 0 & \sigma^2_{P_y} \end{bmatrix} H', \tag{18}$$

where σ_{P_x} and σ_{P_y} are the measurement error standard deviations (in pixels) in the x^{D} and y^{D} axes, respectively[2], and H is the global linearized function

$$H = H_3 H_{\mathbf{T}(\alpha,\epsilon,\rho)} H_2 H_1. \tag{19}$$

as a composite of the Jacobians of the functions $\mathbf{h}_3(\cdot)$, $\mathbf{T}(\alpha, \epsilon, \rho)$, $\mathbf{h}_2(\cdot)$, and $\mathbf{h}_1(\cdot)$ defined in (14), (8), (7), and (1), respectively. Specifically,

$$\begin{aligned} H_3 = \frac{\partial \mathbf{x}^{\mathrm{A}}}{\partial(\mathbf{x} - \mathbf{x}^{\mathrm{S}})} &= \begin{bmatrix} \frac{\partial a}{\partial(x-x^{\mathrm{S}})} & \frac{\partial a}{\partial(y-y^{\mathrm{S}})} & \frac{\partial a}{\partial(z-z^{\mathrm{S}})} \\ \frac{\partial e}{\partial(x-x^{\mathrm{S}})} & \frac{\partial e}{\partial(y-y^{\mathrm{S}})} & \frac{\partial e}{\partial(z-z^{\mathrm{S}})} \end{bmatrix} \\ &= \begin{bmatrix} \frac{y-y^{\mathrm{s}}}{r_{xy}^2} & -\frac{x-x^{\mathrm{s}}}{r_{xy}^2} & 0 \\ -\frac{(x-x^{\mathrm{s}})(z-z^{\mathrm{s}})}{r_{xy}r^2} & -\frac{(y-y^{\mathrm{s}})(z-z^{\mathrm{s}})}{r_{xy}r^2} & \frac{r_{xy}}{r^2} \end{bmatrix}, \end{aligned} \tag{20}$$

where

$$r_{xy} \triangleq \sqrt{(x - x^{\mathrm{s}})^2 + (y - y^{\mathrm{s}})^2} \tag{21}$$

[1] See [3] for the various configurations in multisensor tracking.

[2] These are the same as the errors in (inverted image coordinates) x^{I}, y^{I}.

$$r \triangleq \sqrt{(x-x^{\mathrm{S}})^2+(y-y^{\mathrm{S}})^2+(z-z^{\mathrm{S}})^2}, \qquad (22)$$

and

$$H_{\mathbf{T}(\alpha,\epsilon,\rho)} = \frac{\partial\left(\mathbf{x}-\mathbf{x}^{\mathrm{S}}\right)}{\partial \mathbf{x}^{\mathrm{C}}} = \begin{bmatrix} \frac{\partial(x-x^{\mathrm{S}})}{dx^{\mathrm{C}}} & \frac{\partial(x-x^{\mathrm{S}})}{dy^{\mathrm{C}}} & \frac{\partial(x-x^{\mathrm{S}})}{dz^{\mathrm{C}}} \\ \frac{\partial(y-y^{\mathrm{S}})}{dx^{\mathrm{C}}} & \frac{\partial(y-y^{\mathrm{S}})}{dy^{\mathrm{C}}} & \frac{\partial(y-y^{\mathrm{S}})}{dz^{\mathrm{C}}} \\ \frac{\partial(z-z^{\mathrm{S}})}{dx^{\mathrm{C}}} & \frac{\partial(z-z^{\mathrm{S}})}{dy^{\mathrm{C}}} & \frac{\partial(z-z^{\mathrm{S}})}{dz^{\mathrm{C}}} \end{bmatrix} = \mathbf{T}(\alpha,\epsilon,\rho), \qquad (23)$$

as given in (11), and

$$H_2 = \frac{\partial \mathbf{x}^{\mathrm{C}}}{\partial \mathbf{x}^{\mathrm{I}}} = \begin{bmatrix} \frac{\partial x^{\mathrm{C}}}{\partial x^{\mathrm{I}}} & \frac{\partial x^{\mathrm{C}}}{\partial y^{\mathrm{I}}} \\ \frac{\partial y^{\mathrm{C}}}{\partial x^{\mathrm{I}}} & \frac{\partial y^{\mathrm{C}}}{\partial y^{\mathrm{I}}} \\ \frac{\partial z^{\mathrm{C}}}{\partial x^{\mathrm{I}}} & \frac{\partial z^{\mathrm{C}}}{\partial y^{\mathrm{I}}} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix} \qquad (24)$$

and

$$H_1 = \frac{\partial \mathbf{x}^{\mathrm{I}}}{\partial \mathbf{x}^{\mathrm{D}}} = \begin{bmatrix} \frac{\partial x^{\mathrm{I}}}{\partial x^{\mathrm{D}}} & \frac{\partial x^{\mathrm{I}}}{\partial y^{\mathrm{D}}} \\ \frac{\partial y^{\mathrm{I}}}{\partial x^{\mathrm{D}}} & \frac{\partial y^{\mathrm{I}}}{\partial y^{\mathrm{D}}} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}. \qquad (25)$$

Note the location dependence of the elements of (20).

IV. ONE-SENSOR CASE: SIMULATION RESULTS

Following the conversion (17) and (18), we evaluate the angular measurement error covariance matrix. Next we present the scenarios considered to illustrate the location dependence and the correlation of the errors at various points in the FPA, and the tests whether the conversion method is unbiased and has a consistent covariance [2].

A. Simulation Scenario

We consider a camera positioned at the origin in global 3D Cartesian space, i.e., $\mathbf{x}^{\mathrm{S}} = [0,0,0]'$, with the orientation angles of the camera $(\alpha,\epsilon,\rho) = (0,0,0)$. This is for the sake of illustration; for other orientations, (23) will differ. The aspect ratio of the camera is 16:9, which is typical of commercial camera sensors. This translates to a horizontal and vertical FOV of $\Theta_x = 60^\circ$ and $\Theta_y = 37.5^\circ$, respectively, and an image with $P_x = 1920$ pixels by $P_y = 1080$ pixels for a 2 megapixel (MP) camera, and $P_x = 3840$ pixels by $P_y = 2160$ pixels for a 8MP camera. The camera records measurements when the target is in its FOV with FPA measurement error standard deviations $\sigma_{P_x} = \sigma_{P_y} = 1$ pixel[3]. This error is equivalent to angles of 0.031° (or 0.541 mrad) and 0.016° (or 0.279 mrad) on average for a 2MP and 8MP camera, respectively.

[3] As shown in [5] using a physics-based model, the optimal measurement extractor for a point target can reach a s.d. of $\frac{1}{2}$ pixel. We took a more conservative approach here.

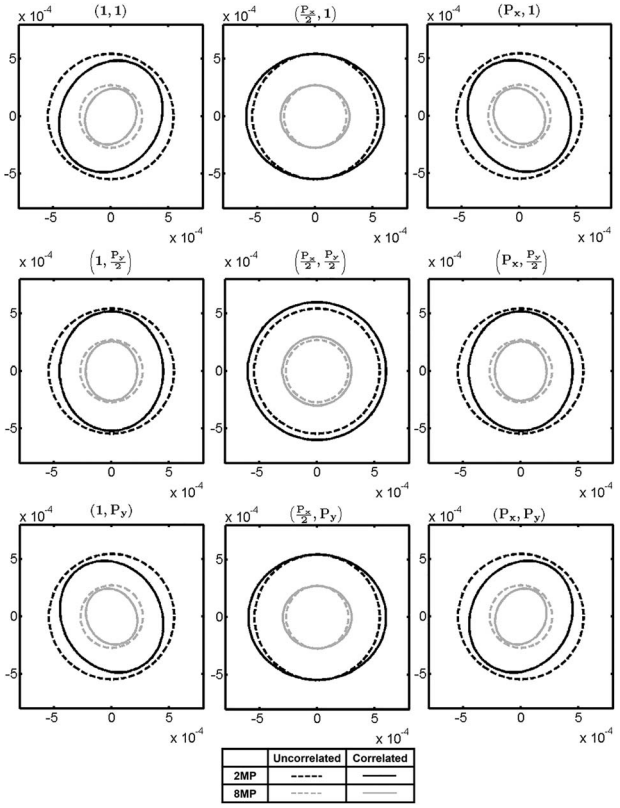

Figure 4. Azimuth-elevation error ellipse plots with assumed uncorrelated errors (dotted) and actual correlated errors (solid) at each of the chosen image points $(x_i^{\mathrm{D}}, y_i^{\mathrm{D}})$, $i = 1,\ldots,9$, from exhaustive combination pairs of $x_i^{\mathrm{D}} \in \{1, \frac{P_x}{2}, P_x\}$ and $y_i^{\mathrm{D}} \in \{1, \frac{P_y}{2}, P_y\}$, for a 2MP camera (black) and 8MP camera (gray). The locations are according to display image coordinates in Fig. 1.

B. Significance of Correlation Errors

For the given camera specifications, Fig. 4 shows measurement error ellipses in azimuth-elevation that were generated for each of nine selected points in the image plane spanning left to right of x^{D} axis and top to bottom of y^{D} axis. The nine selected points of the display image plane include: (i) top-left $(1,1)$, (ii) top-center $\left(\frac{P_x}{2},1\right)$, (iii) top-right $(P_x,1)$, (iv) mid-left $\left(1,\frac{P_y}{2}\right)$, (v) center $\left(\frac{P_x}{2},\frac{P_y}{2}\right)$, (vi) mid-right $\left(P_x,\frac{P_y}{2}\right)$, (vii) bottom-left $(1,P_y)$, (viii) bottom-center $\left(\frac{P_x}{2},P_y\right)$, and (ix) bottom-right (P_x,P_y). A baseline uncorrelated measurement error circle is also shown (in dotted line) to represent the uncorrelated error covariance $\mathbf{R}^{\mathrm{A}}_{\mathrm{uncorr}}$, with $\sigma_a = \sigma_e = \sigma_{P_x}\frac{\Theta_x}{P_x}$, for comparison with the correlated and location-dependent measurement error ellipses $\mathbf{R}^{\mathrm{A}}_{\mathrm{corr}}$ (in solid line) derived from the proposed conversion method. Error ellipses of the 2MP and 8MP camera are colored black andgray, respectively.

To determine if the measurement correlation error is significant, the percentage difference between the areas of the correlated and uncorrelated error ellipses is

Table II
Percentage Difference Between Areas of Correlated Error Ellipses and Uncorrelated Error Circles at Each of Nine Image Points $(x^{\mathrm{D}}, y^{\mathrm{D}})$ for 2MP and 8MP Camera

MP	y^{D} \ x^{D}	1	$\frac{P_x}{2}$	P_x
2	1	−26.8%	10.0%	−26.8%
	$\frac{P_y}{2}$	−21.0%	21.6%	−21.0%
	P_y	−26.8%	10.0%	−26.8%
8	1	−26.8%	10.0%	−26.8%
	$\frac{P_y}{2}$	−21.0%	21.6%	−21.0%
	P_y	−26.8%	10.0%	−26.8%

calculated as follows:

$$\mathrm{d} = \frac{\pi\left(\sqrt{\det\left(\mathbf{R}^{\mathrm{A}}_{\mathrm{corr}}\right)} - \sqrt{\det\left(\mathbf{R}^{\mathrm{A}}_{\mathrm{uncorr}}\right)}\right)}{\pi\sqrt{\det\left(\mathbf{R}^{\mathrm{A}}_{\mathrm{uncorr}}\right)}} \times 100\%, \quad (26)$$

where $\det(\mathbf{A})$ is the determinant of matrix $\mathbf{A}$. A magnitude of $|\mathrm{d}| \geq 10\%$ will imply that the correlated measurement error is significant. A positive percentage difference value implies the error ellipse is larger in area than the baseline uncorrelated error circle, while a negative percentage difference value implies the error ellipse is smaller in area than the baseline uncorrelated error circle. Table II shows the percentage difference between the areas of the correlated and uncorrelated error ellipses for each of the nine selected image points.

With reference to Fig. 4, the error ellipses of the 8MP camera (in gray) are 4 times smaller in area than those of the 2MP camera (in black), due to the larger number of pixels in the image space for the 8MP camera. The ellipses' shapes and orientations of both correlated and uncorrelated error covariances are similar for both 2MP and 8MP cameras, due to both cameras sharing the same aspect ratio. From Table II, the percentage difference values are the same for both cases of the 2MP and 8MP cameras. Given the similarities in observations in both Fig. 4 and Table II, we can, for simplicity, consider the results (of both figure and result table) of either camera for subsequent analysis and discussion in the rest of the paper.

From Table II, since all percentage difference values have magnitude greater than or equal to 10%, the correlations of azimuth and elevation measurement errors are considered significant at all nine selected points of the image.

The most significant differences are at the corner points of the image, where the correct ellipse (with the error correlations) is 26.8% smaller in area than the circles corresponding to the (assumed) uncorrelated errors at these points. The next most significant difference is at the center point of the image, with the correct error ellipse area 21.6% larger than the (assumed) uncorrelated error circle at this point. This is followed by the mid-left and mid-right points of the image, with the correlated error ellipse 21.0% larger in area than the uncorrelated error circle at this point. The least significant difference is at the top-center and bottom-center points of the image, with the correlated error ellipse 10.0% larger in area than the uncorrelated error circle at these points.

C. Unbiasedness and Covariance Consistency

We next assess whether the proposed conversion method is unbiased and that the derived covariance matrix $\mathbf{R}^{\mathrm{A}}_{\mathrm{corr}}$ is consistent.

As the global transformation function $\mathbf{h}(\cdot)$ defined in (17) mapping points in the image space to azimuth-elevation space is nonlinear, a set of $N = 10\,000$ random points with Gaussian noise was generated for each selected image point modeling the probability density distributions in image space, i.e., for each of nine selected image points, separate realizations of N points were generated according to the distribution

$$\left\{\mathbf{x}^{\mathrm{D}}_{i,k}\right\}_{k=1}^{N} = \left\{\begin{bmatrix} x^{\mathrm{D}}_{i,k} \\ y^{\mathrm{D}}_{i,k} \end{bmatrix}\right\}_{k=1}^{N} \sim \mathcal{N}\left(\begin{bmatrix} x^{\mathrm{D}}_{i} \\ y^{\mathrm{D}}_{i} \end{bmatrix}, \begin{bmatrix} \sigma^2_{P_x} & 0 \\ 0 & \sigma^2_{P_y} \end{bmatrix}\right) \quad i = 1, \ldots, 9, \quad (27)$$

where k is the index of the Monte–Carlo runs. Each set of points undergoes the non-linear transformation $\mathbf{h}(\cdot)$ to derive a distribution of points

$$\left\{\mathbf{x}^{\mathrm{A}}_{i,k}\right\}_{k=1}^{N} = \left\{\begin{bmatrix} a_{i,k} \\ e_{i,k} \end{bmatrix}\right\}_{k=1}^{N} \qquad i = 1, \ldots, 9 \quad (28)$$

in azimuth-elevation space corresponding to the ENU coordinate system. The set (28) is based on FPA random points specified by (27). The sample mean of the angles' distribution is

$$\bar{\mathbf{x}}^{\mathrm{A}}_{i} = \begin{bmatrix} \bar{a}^{\mathrm{A}}_{i} \\ \bar{e}^{\mathrm{A}}_{i} \end{bmatrix} = \frac{1}{N}\sum_{k=1}^{N} \mathbf{x}^{\mathrm{A}}_{i,k} \qquad i = 1, \ldots, 9, \quad (29)$$

and the sample covariance is

$$\bar{\mathbf{R}}^{\mathrm{A}}_{\mathrm{corr},i} = \frac{1}{N}\sum_{k=1}^{N}\left(\mathbf{x}^{\mathrm{A}}_{i,k} - \bar{\mathbf{x}}^{\mathrm{A}}_{i}\right)\left(\mathbf{x}^{\mathrm{A}}_{i,k} - \bar{\mathbf{x}}^{\mathrm{A}}_{i}\right)' \quad i = 1, \ldots, 9. \quad (30)$$

Next, each of the nine selected image points $\mathbf{x}^{\mathrm{I}}_{i}$, $i = 1, \ldots, 9$ undergoes the global transformation function $\mathbf{h}(\cdot)$ defined in (17) to obtain the converted angular state $\hat{\mathbf{x}}^{\mathrm{A}}_{i}$ comprising the LOS angles, $\hat{a}^{\mathrm{A}}_{i}$ and $\hat{e}^{\mathrm{A}}_{i}$.

The ratio of the bias error to the baseline measurement error standard deviation is calculated in the azimuth and elevation components separately

$$r_{a,i} = \frac{\hat{a}^{\mathrm{A}}_{i} - \bar{a}^{\mathrm{A}}_{i}}{\sigma_{a,i}} \quad \text{and} \quad r_{e,i} = \frac{\hat{e}^{\mathrm{A}}_{i} - \bar{e}^{\mathrm{A}}_{i}}{\sigma_{e,i}}, \quad (31)$$

where $\sigma_{a,i}$ and $\sigma_{e,i}$ are the square roots of the diagonal terms of (30). These are $\mathcal{N}\left(0, \frac{1}{N}\right)$ distributed (see, e.g., [2]), so they have to be in the 95% region: $\left[-\frac{2}{\sqrt{N}}, \frac{2}{\sqrt{N}}\right] =$

Table III
Ratio r_a of the Bias to the Baseline Measurement Error Standard Deviation for Azimuth Component of 2MP and 8MP Cameras

MP	y^{D} \ x^{D}	1	$\frac{P_x}{2}$	P_x
2	1	0.0101	−0.0022	0.0051
	$\frac{P_y}{2}$	0.0063	0.0048	−0.0051
	P_y	0.0106	−0.0063	0.0032
8	1	−0.0004	0.0036	−0.0022
	$\frac{P_y}{2}$	−0.0106	−0.0142	0.0116
	P_y	0.0008	−0.0066	0.0111

Table IV
Ratio r_e of the Bias to the Baseline Measurement Error Standard Deviation for Elevation Component of 2MP and 8MP Cameras

MP	y^{D} \ x^{D}	1	$\frac{P_x}{2}$	P_x
2	1	0.0051	−0.0147	0.0062
	$\frac{P_y}{2}$	−0.0092	0.0160	0.0038
	P_y	−0.0009	0.0064	−0.0013
8	1	0.0025	−0.0128	0.0054
	$\frac{P_y}{2}$	0.0138	−0.0137	−0.0060
	P_y	−0.0048	0.0124	0.0081

$[-0.02, 0.02]$ to confirm the unbiasedness of the proposed conversion method. Tables III and IV show the results of the ratio for the azimuth (r_a) and elevation (r_e) components, respectively, for both the 2MP and 8MP cameras considered. It can be seen that all ratio values are within the 95% region of ± 0.02. Thus, the proposed conversion method is unbiased.

Lastly, the following covariance consistency test [2] is used:

$$\kappa_i \triangleq \frac{1}{N}\sum_{k=1}^{N}\left(\mathbf{x}_{i,k}^{\mathrm{A}}-\hat{\mathbf{x}}_i^{\mathrm{A}}\right)'\left(\bar{\mathbf{R}}_{\mathrm{corr},i}^{\mathrm{A}}\right)^{-1}\left(\mathbf{x}_{i,k}^{\mathrm{A}}-\hat{\mathbf{x}}_i^{\mathrm{A}}\right) \leq \frac{1}{N}\chi_{2N}^2(1-\alpha) \quad i=1,\ldots,9, \tag{32}$$

where $\chi_{2N}^2(1-\alpha)$ is α-tail probability of the chi-square distribution with $2N$ degrees of freedom. If satisfied, these tests confirm the unbiasedness of the conversion and the correctness of the covariance $\mathbf{R}_{\mathrm{corr},i}^{\mathrm{A}}$, including the correlation between azimuth and elevation errors, for each of the nine selected image points [2]. Table V shows the results of the consistency test values κ_i for both the 2MP and 8MP cameras considered. It can be seen that that all covariance test statistics are below the $1-\alpha = 95\%$ bound of $\frac{1}{10\,000}\chi_{20\,000}^2(0.95) = 2.0330$.

Thus, the proposed conversion method that yields location-dependent variances and correlations is unbiased and consistent.

Table V
Test for Covariance Matrix Consistency of 2MP and 8MP Cameras

MP	y^{D} \ x^{D}	1	$\frac{P_x}{2}$	P_x
2	1	2.0002	2.0002	2.0001
	$\frac{P_y}{2}$	2.0001	2.0002	2.0001
	P_y	2.0002	2.0001	2.0000
8	1	2.0000	2.0002	2.0000
	$\frac{P_y}{2}$	2.0004	2.0003	2.0002
	P_y	2.0000	2.0002	2.0003

V. TWO-SENSOR CASE: MULTI-SENSOR FUSION AND SIMULATION RESULTS

In this section, we derive the composite target position measurement in 3D Cartesian space, by fusing LOS, i.e., 2D, target angle measurements of a two-camera setup. This derivation can be divided into two stages:

Stage 1. Apply our proposed conversion method in Section III-B to derive, for each camera, the target LOS angle measurements with their corresponding angular error covariances with location-dependent variances and cross-correlated errors.

Stage 2. The maximum likelihood (ML) estimate of the 3D composite target position measurement is obtained from fusing[4] the LOS angle measurements of the two cameras using a numerical search performed via the iterated least squares (ILS) technique [2]. The derived ML estimate is proven to be statistically efficient, and as such, the covariance matrix from the CRLB can be used as the measurement error covariance matrix for the resulting composite target 3D position measurement [11] in subsequent track filtering.

It is envisioned that the more appropriate angular error covariance of our approach in Stage 1, compared to an uncorrelated angular error covariance baseline, would translate to improvements to the final composite target position measurement and its Cartesian measurement error covariance matrix in Stage 2.

A. Derivation of Composite (Fused) Position Measurement and Its Corresponding Cartesian Measurement Error Covariance

Consider 2 camera sensors with known sensor positions $\mathbf{x}^{\mathrm{s}_j} = [x^{\mathrm{s}_j}, y^{\mathrm{s}_j}, z^{\mathrm{s}_j}]'$, $j = 1, 2$.

Stage 1. Following the conversion (17), (18), we have the sensor-specific (s_j) LOS measurements and their corresponding covariances denoted $\mathbf{x}_{\mathrm{s}_j}^{\mathrm{A}}$ and $\mathbf{R}_{\mathrm{s}_j,\mathrm{corr}}^{\mathrm{A}}$, $j = 1, 2$, respectively.

Stage 2. The LOS measurements to the target from Stage 1 are fused to obtain a composite position measurement (estimate) $\hat{\mathbf{x}}^{\mathrm{F}} = \left[x^{\mathrm{F}}, y^{\mathrm{F}}, z^{\mathrm{F}}\right]'$ in ENU space [11].

[4]This is Type 3 fusion [3] with the two cameras assumed synchronized.

The measurement $\mathbf{x}^{\mathrm{A}}_{\mathrm{s}_j}$ from sensor s_j relates to the 3D position according to

$$\mathbf{x}^{\mathrm{A}}_{\mathrm{s}_j} = \mathbf{g}_{\mathrm{s}_j}\left(\mathbf{x}^{\mathrm{F}}, \mathbf{x}^{\mathrm{s}_j}\right) + \mathbf{w}_{\mathrm{s}_j}, \qquad j = 1, 2, \tag{33}$$

where

$$\mathbf{w}_{\mathrm{s}_j} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}^{\mathrm{A}}_{\mathrm{s}_j,\mathrm{corr}}) \tag{34}$$

are the angle measurement noises, and

$$\mathbf{g}_{\mathrm{s}_j}\left(\mathbf{x}^{\mathrm{F}}, \mathbf{x}^{\mathrm{s}_j}\right) = \mathbf{h}_3\left(\mathbf{x}^{\mathrm{F}} - \mathbf{x}^{\mathrm{s}_j}\right) \tag{35}$$

with $\mathbf{x}^{\mathrm{A}} = \mathbf{x}^{\mathrm{A}}_{\mathrm{s}_j}$, $\mathbf{x} = \mathbf{x}^{\mathrm{F}}$ and $\mathbf{x}^{\mathrm{S}} = \mathbf{x}^{\mathrm{s}_j}$ in (14).

The composite position measurement $\hat{\mathbf{x}}^{\mathrm{F}}$ in ENU space is obtained by maximizing the likelihood function of $\mathbf{x}^{\mathrm{F}}$ (based on $\mathbf{x}^{\mathrm{A}}_{\mathrm{s}_1}, \mathbf{x}^{\mathrm{A}}_{\mathrm{s}_2}$, not shown for simplicity)

$$\mathbf{\Lambda}\left(\mathbf{x}^{\mathrm{F}}\right) = \prod_{j=1}^{2} p(\mathbf{x}^{\mathrm{A}}_{\mathrm{s}_j}|\mathbf{x}^{\mathrm{F}}), \tag{36}$$

where

$$p(\mathbf{x}^{\mathrm{A}}_{\mathrm{s}_j}|\mathbf{x}^{\mathrm{F}}) = |2\pi\mathbf{R}^{\mathrm{A}}_{\mathrm{s}_j,\mathrm{corr}}|^{-\frac{1}{2}} \cdot \exp\Big(-\frac{1}{2}\left[\mathbf{x}^{\mathrm{A}}_{\mathrm{s}_j} - \mathbf{g}_{\mathrm{s}_j}\left(\mathbf{x}^{\mathrm{F}}, \mathbf{x}^{\mathrm{s}_j}\right)\right]' \left(\mathbf{R}^{\mathrm{A}}_{\mathrm{s}_j,\mathrm{corr}}\right)^{-1}\left[\mathbf{x}^{\mathrm{A}}_{\mathrm{s}_j} - \mathbf{g}_{\mathrm{s}_j}\left(\mathbf{x}^{\mathrm{F}}, \mathbf{x}^{\mathrm{s}_j}\right)\right]\Big). \tag{37}$$

The ML estimate (the composite measurement) is then

$$\hat{\mathbf{x}}^{\mathrm{F}}_{\mathrm{ML}} = \underset{\mathbf{x}^{\mathrm{F}}}{\operatorname{argmax}}\, \mathbf{\Lambda}\left(\mathbf{x}^{\mathrm{F}}\right). \tag{38}$$

Finding the ML estimate in this case is equivalent to a nonlinear least squares problem. The numerical search will be performed via the ILS technique [2].

The ILS estimate after $(m+1)$ iterations is

$$\hat{\mathbf{x}}^{\mathrm{F}(m+1)}_{\mathrm{ILS}} = \hat{\mathbf{x}}^{\mathrm{F}(m)}_{\mathrm{ILS}} + \left[\left(\mathbf{G}^{(m)}\right)' \mathbf{R}^{-1}_{\mathrm{corr}} \mathbf{G}^{(m)}\right]^{-1} \cdot \left(\mathbf{G}^{(m)}\right)' \mathbf{R}^{-1}_{\mathrm{corr}}\left[\mathbf{z} - \mathbf{g}\left(\hat{\mathbf{x}}^{\mathrm{F}(m)}_{\mathrm{ILS}}, \mathbf{s}\right)\right], \tag{39}$$

with

$$\mathbf{z} = \left[\mathbf{x}^{\mathrm{A}}_{\mathrm{s}_1}, \mathbf{x}^{\mathrm{A}}_{\mathrm{s}_2}\right]', \tag{40}$$

$$\mathbf{s} = \left[\mathbf{x}^{\mathrm{s}_1}, \mathbf{x}^{\mathrm{s}_2}\right]', \tag{41}$$

$$\mathbf{g}\left(\hat{\mathbf{x}}^{\mathrm{F}(m)}_{\mathrm{ILS}}, \mathbf{s}\right) = [a_{\mathrm{s}_1}, e_{\mathrm{s}_1}, a_{\mathrm{s}_2}, e_{\mathrm{s}_2}], \tag{42}$$

$$\mathbf{R}_{\mathrm{corr}} = \begin{bmatrix} \mathbf{R}^{\mathrm{A}}_{\mathrm{s}_1,\mathrm{corr}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}^{\mathrm{A}}_{\mathrm{s}_2,\mathrm{corr}} \end{bmatrix}, \tag{43}$$

and

$$\mathbf{G}^{(m)} = \left.\frac{\partial \mathbf{g}\left(\hat{\mathbf{x}}^{\mathrm{F}(m)}_{\mathrm{ILS}}, \mathbf{s}\right)}{\partial \mathbf{x}^{\mathrm{F}}}\right|_{\mathbf{x}^{\mathrm{F}} = \hat{\mathbf{x}}^{\mathrm{F}(m)}_{\mathrm{ILS}}}, \tag{44}$$

is the Jacobian matrix of the stacked measurement vector evaluated at the ILS estimate from the previous iteration. The Jacobian matrix is given by

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{s_1} \\ \mathbf{G}_{s_2} \end{bmatrix}, \tag{45}$$

with $\mathbf{G}_{\mathrm{s}_j} = H_3$, $j = 1, 2$ as in (20).

To perform the numerical search via ILS, an initial estimate $\hat{\mathbf{x}}^{\mathrm{F}(0)}_{\mathrm{ILS}}$ is required. The initial Cartesian position is solved using the LOS measurements a_{s_1}, a_{s_2}, and a_{e_1} of the first iteration, and the known sensor positions $\mathbf{x}^{\mathrm{s}_1}$ and $\mathbf{x}^{\mathrm{s}_2}$ as follows:

$$x^{\mathrm{F}(0)}_{\mathrm{ILS}} = x^{\mathrm{s}_1} + \left(y^{\mathrm{F}(0)}_{\mathrm{ILS}} - y^{\mathrm{s}_1}\right)\tan a_{\mathrm{s}_1}, \tag{46}$$

$$y^{\mathrm{F}(0)}_{\mathrm{ILS}} = \frac{(x^{\mathrm{s}_2} - x^{\mathrm{s}_1}) + (y^{\mathrm{s}_1}\tan a_{\mathrm{s}_1} - y^{\mathrm{s}_2}\tan a_{\mathrm{s}_2})}{\tan a_{\mathrm{s}_1} - \tan a_{\mathrm{s}_2}}, \tag{47}$$

$$z^{\mathrm{F}(0)}_{\mathrm{ILS}} = z^{\mathrm{s}_1} + \tan e_{\mathrm{s}_1}\left|\frac{(x^{\mathrm{s}_1} - x^{\mathrm{s}_2})\cos a_{\mathrm{s}_2} + (y^{\mathrm{s}_1} - y^{\mathrm{s}_2})\sin a_{\mathrm{s}_2}}{\sin(a_{\mathrm{s}_1} - a_{\mathrm{s}_2})}\right|. \tag{48}$$

For the error covariance, the CRLB provides a lower bound on the estimation error obtainable from an unbiased estimator

$$\mathbb{E}\left[(\mathbf{x}^{\mathrm{F}} - \hat{\mathbf{x}}^{\mathrm{F}})(\mathbf{x}^{\mathrm{F}} - \hat{\mathbf{x}}^{\mathrm{F}})'\right] \geq \mathbf{J}^{-1}, \tag{49}$$

where $\mathbf{J}$ is the Fisher Information Matrix (FIM), $\mathbf{x}^{\mathrm{F}}$ is the true value to be estimated, and $\hat{\mathbf{x}}^{\mathrm{F}}$ is the estimate [2], [11]. The FIM is given by

$$\mathbf{J} = \mathbb{E}\left\{\left[\nabla_{\mathbf{x}^{\mathrm{F}}} \ln \mathbf{\Lambda}\left(\mathbf{x}^{\mathrm{F}}\right)\right]\left[\nabla_{\mathbf{x}^{\mathrm{F}}} \ln \mathbf{\Lambda}\left(\mathbf{x}^{\mathrm{F}}\right)\right]'\right\}\Big|_{\mathbf{x}^{\mathrm{F}} = \mathbf{x}^{\mathrm{F}}_{\mathrm{true}}}. \tag{50}$$

The gradient of the log-likelihood function is

$$\nabla_{\mathbf{x}^{\mathrm{F}}} \boldsymbol{\lambda}(\mathbf{x}^{\mathrm{F}}) = \sum_{j=1}^{2} \mathbf{G}'_{\mathrm{s}_j}\left(\mathbf{R}^{\mathrm{A}}_{\mathrm{s}_j,\mathrm{corr}}\right)^{-1}\left(\mathbf{x}^{\mathrm{A}}_{\mathrm{s}_j} - \mathbf{g}_{\mathrm{s}_j}\left(\mathbf{x}^{\mathrm{F}}, \mathbf{x}^{\mathrm{s}_j}\right)\right), \tag{51}$$

which, when plugged into, (50) gives

$$\mathbf{J} = \sum_{j=1}^{2} \mathbf{G}'_{\mathrm{s}_j}\left(\mathbf{R}^{\mathrm{A}}_{\mathrm{s}_j,\mathrm{corr}}\right)^{-1} \mathbf{G}_{\mathrm{s}_j}\Big|_{\mathbf{x}^{\mathrm{F}} = \mathbf{x}^{\mathrm{F}}_{\mathrm{true}}} = \mathbf{G}'\mathbf{R}^{-1}_{\mathrm{corr}}\mathbf{G}\Big|_{\mathbf{x}^{\mathrm{F}} = \mathbf{x}^{\mathrm{F}}_{\mathrm{true}}}. \tag{52}$$

B. Simulation Scenario

In the two-sensor setup, we have two cameras s_1 and s_2 spaced 1 km apart in global 3D cartesian space, with sensor positions $\mathbf{x}^{\mathrm{s}_1} = [x^{\mathrm{s}_1}, y^{\mathrm{s}_1}, z^{\mathrm{s}_1}]' = [-500\text{ m}, 0\text{ m}, 0\text{ m}]'$ and $\mathbf{x}^{\mathrm{s}_2} = [x^{\mathrm{s}_2}, y^{\mathrm{s}_2}, z^{\mathrm{s}_2}]' = [500\text{ m}, 0\text{ m}, 0\text{ m}]'$. Unlike the single-sensor case in Section IV-A, we consider the more realistic non-zero orientation angles of cameras 1 and 2 measuring $(\alpha_{\mathrm{s}_1}, \epsilon_{\mathrm{s}_1}, \rho_{\mathrm{s}_1}) = (24.5^\circ, 2.1^\circ, 4.5^\circ)$ and $(\alpha_{\mathrm{s}_2}, \epsilon_{\mathrm{s}_2}, \rho_{\mathrm{s}_2}) = (-2.6^\circ, -3.4^\circ, 2.8^\circ)$, respectively. It is worth noting that for every nonzero camera orientation, it will map to a reference case of our generic

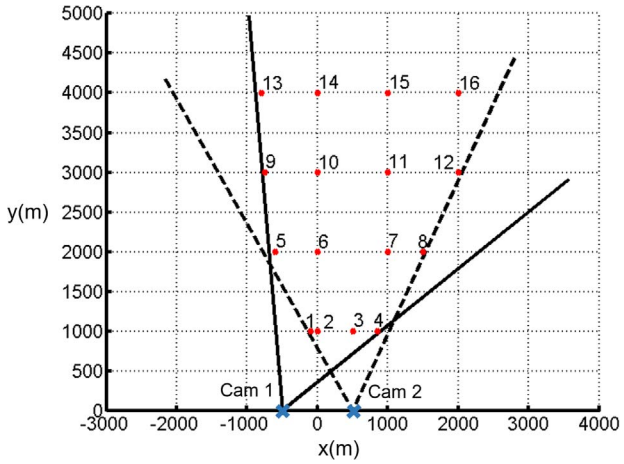

Figure 5. 2-camera setup with 16 targets in ENU space.

azimuth-elevation error ellipse characterization results we derived with camera orientation angles of zero values in Section IV-B. For this experiment, both cameras are 2-MP cameras, with the remaining camera sensor parameters and measurement error standard deviations the same as defined for the 2-MP variation of the one-sensor case in Section IV-A.

A total of 16 point targets at various ranges and heights in ENU space were considered. The 2D plot in Fig. 5 shows the camera sensor locations as crosses, with Camera 1 and 2 fields-of-view (FOV) in solid and dotted lines, respectively, and the 16 labeled targets as solid circles that are all within the FOV of both sensors. For each target point in ENU space, we plot their corresponding positions in the display image space of Camera 1 (left) and Camera 2 (right) in pixel coordinates, as shown in Fig. 6. The dotted lines segment the image space into nine general reference regions in which we have previously characterized its azimuth-elevation error ellipses in Section IV-A. Based on Fig. 6, the distribution of the targets in the image spaces of both cameras is sufficiently comprehensive to cover all scenarios for fusion.

C. Evaluation of the Composite Position Measurements and Their Error Covariances

We envisaged that the angular error covariance of our approach used in Stage 1, compared to an uncorrelated angular error covariance baseline, would translate

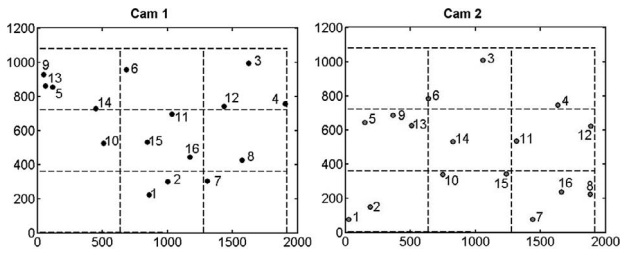

Figure 6. 16 targets in display image space of Camera 1 (left) and Camera 2 (right).

to improvements of the final composite target position measurement and its error covariance matrix. As such, we compared the composite position measurements and their corresponding Cartesian error covariances derived using

(i) the location-dependent angular error covariances, derived from the method described in Section V-A, versus
(ii) the constant uncorrelated angular error covariance baseline,

with the following evaluation criteria:

1) 3D position error improvement
2) 3D position algorithm-calculated accuracy
3) Statistical consistency of the estimates.

For each true target location in ENU space based on Fig. 5, denoted as $\check{\mathbf{x}}_i^{\mathrm{F}}, i = 1, \ldots, 16$, we get the corresponding positions in display image space of Camera 1 and Camera 2 based on Fig. 6, denoted as $\check{\mathbf{x}}_{\mathrm{s}_1,i}^{\mathrm{D}}$ and $\check{\mathbf{x}}_{\mathrm{s}_2,i}^{\mathrm{D}}, i = 1, \ldots, 16$, respectively, using the conversion

$$\check{\mathbf{x}}_{\mathrm{s}_j,i}^{\mathrm{D}} = \mathbf{h}_1^{-1}\left(\mathbf{h}_2^{-1}\left(\mathbf{T}_{\mathrm{s}_j}^{-1}\left(\alpha_{\mathrm{s}_j}, \epsilon_{\mathrm{s}_j}, \rho_{\mathrm{s}_j}\right)\left(\check{\mathbf{x}}_i^{\mathrm{F}}\right)\right)\right) \tag{53}$$

with $\mathbf{h}_1^{-1}(\cdot)$, $\mathbf{h}_2^{-1}(\cdot)$ and $\mathbf{T}_{\mathrm{s}_j}^{-1}(\cdot)$, $j = 1, 2$ the inverses of $\mathbf{h}_1(\cdot)$, $\mathbf{h}_2(\cdot)$ and $\mathbf{T}(\cdot)$ given by (1), (7), and (8), respectively. We assume the corresponding covariances in display image coordinates are given by

$$\mathbf{P}_{\mathrm{s}_j,i}^{\mathrm{D}} = \begin{bmatrix} \sigma_{P_x}^2 & 0 \\ 0 & \sigma_{P_y}^2 \end{bmatrix}, \quad j = 1, 2, \quad i = 1, \ldots, 16 \tag{54}$$

similar to what was assumed in (18).

We now derive the 3D composite target position estimates and their corresponding Cartesian error covariances of the following variations:

Case (i) - $\left(\check{\mathbf{x}}_{\mathrm{corr},i}^{\mathrm{F}}, \mathbf{P}_{\mathrm{corr},i}^{\mathrm{F}}\right)$. Following the derivation steps in Section V-A, Stage 1 is first executed by applying the conversion equations (17) and (18) with display image inputs (53) and (54) of the jth camera across $i = 1, \ldots, 16$ target points. We thus derive the camera-specific target LOS angles and the corresponding angular error covariances with location dependent variances and cross-correlated errors for each of the 16 target points, i.e., $\left(\check{\mathbf{x}}_{\mathrm{s}_j,i}^{\mathrm{A}}, \mathbf{R}_{\mathrm{s}_j,\mathrm{corr},i}^{\mathrm{A}}\right), i = 1, \ldots, 16$ and $j = 1, 2$. Stage 2 is then executed to derive the composite target position estimates (fused measurements) $\check{\mathbf{x}}_{\mathrm{corr},i}^{\mathrm{F}}$ obtained from (33) to (48) and the CRLB $\mathbf{P}_{\mathrm{corr},i}^{\mathrm{F}}$ obtained from (52) evaluated at $\mathbf{x}^{\mathrm{F}} = \check{\mathbf{x}}_{\mathrm{corr},i}^{\mathrm{F}}$ for the ith target point, $i = 1, \ldots, 16$.

The covariances are evaluated at the actual (noiseless) LOS angles (like the CRLB) and will be used to compare the uncertainties of the proposed and baseline methods.

Case (ii) - $\left(\check{\mathbf{x}}_{\mathrm{uncorr},i}^{\mathrm{F}}, \mathbf{P}_{\mathrm{uncorr},i}^{\mathrm{F}}\right)$. First, the camera-specific target LOS angles for each of the $i = 1, \ldots, 16$ target points $\check{\mathbf{x}}_{\mathrm{s}_j,i}^{\mathrm{A}}$ were derived following (17), as in

Table VI
Percentage Difference Between Volumes of the Correlated and Uncorrelated Variants of the Cartesian Measurement Error Ellipses at Each of 16 Target Points $\breve{\mathbf{x}}_i^{\mathrm{F}}$, $i = 1, \ldots, 16$

Tgt	1	2	3	4	5	6	7	8
Vol diff. (%)	−8.9	1.7	8.0	**−19.9**	**−25.0**	**18.2**	7.7	**−21.3**
Tgt	9	10	11	12	13	14	15	16
Vol diff. (%)	**−20.4**	**19.2**	**25.9**	**−13.9**	**−13.5**	**17.6**	**27.2**	3.7

Case (i). The baseline uncorrelated angular error covariances $\mathbf{R}^{\mathrm{A}}_{\mathrm{s}_j,\mathrm{uncorr},i}$, $i = 1, \ldots, 16$ and $j = 1, 2$ follow according to (16) with[5] $\sigma_a = \sigma_e = \sigma_{P_x} \frac{\Theta_x}{P_x}$. Similarly, the composite target position estimates $\breve{\mathbf{x}}^{\mathrm{F}}_{\mathrm{uncorr},i}$ are obtained from (33) to (48) and the CRLB $\mathbf{P}^{\mathrm{F}}_{\mathrm{uncorr},i}$ is obtained from (52) evaluated at $\mathbf{x}^{\mathrm{F}} = \breve{\mathbf{x}}^{\mathrm{F}}_{\mathrm{uncorr},i}$ for the ith target point, $i = 1, \ldots, 16$, with the correlated covariances $\mathbf{R}^{\mathrm{A}}_{\mathrm{s}_j,\mathrm{corr},i}$ in the equations replaced with their uncorrelated counterpart $\mathbf{R}^{\mathrm{A}}_{\mathrm{s}_j,\mathrm{uncorr},i}$.

We now compare and evaluate our proposed approach according to each of the three criteria listed above.

1) 3D position error improvement: The percentage difference between the volumes of the correlated and uncorrelated variants of the Cartesian measurement error ellipses is calculated by

$$\mathrm{d} = \frac{\pi\left(\sqrt{\det\left(\mathbf{P}^{\mathrm{F}}_{\mathrm{corr}}\right)} - \sqrt{\det\left(\mathbf{P}^{\mathrm{F}}_{\mathrm{uncorr}}\right)}\right)}{\pi\sqrt{\det\left(\mathbf{P}^{\mathrm{F}}_{\mathrm{uncorr}}\right)}} \times 100\%, \quad (55)$$

where $\det(\mathbf{A})$ is the determinant of matrix $\mathbf{A}$, and the results are tabulated in Table VI for each of the 16 selected target points. A magnitude of $|\mathrm{d}| \geq 10\%$ will imply that the difference between the correlated and uncorrelated variants of the Cartesian measurement error ellipses is significant, and these values are highlighted in bold in Table VI.

It is observed that the majority of target points have significant differences in volume between the correlated and uncorrelated variants of the Cartesian measurement error ellipses, which demonstrates the importance of using our proposed conversion method for improved error modeling. Furthermore, our approach can yield up to a maximum of approximately 27% difference in volume of the correlated Cartesian measurement error ellipse relative to the baseline uncorrelated Cartesian measurement error ellipse.

2) 3D position accuracy improvement: At this point, we generate a set of $N = 1000$ random points with Gaussian noise in the display image space of Cameras 1 and 2 for each selected target point, which models the probability distributions in the display image space

$$\left\{\mathbf{x}^{\mathrm{D}(k)}_{\mathrm{s}_j,i}\right\}_{k=1}^{N} \sim \mathcal{N}\left(\breve{\mathbf{x}}^{\mathrm{D}}_{\mathrm{s}_j,i}, \mathbf{R}^{\mathrm{D}}_{\mathrm{s}_j,i}\right) \quad j = 1, 2, \quad i = 1, \ldots, 16. \quad (56)$$

Each set of points undergoes the non-linear transformation $\mathbf{h}(\cdot)$ defined in (17) to derive a distribution of points

$$\left\{\mathbf{x}^{\mathrm{A}(k)}_{\mathrm{s}_j,i}\right\}_{k=1}^{N} \quad j = 1, 2, \quad i = 1, \ldots, 16, \quad (57)$$

that model the probability distribution in azimuth-elevation space corresponding to the ENU coordinate system. For *Case (i)*, the corresponding set of correlated angular error covariances

$$\left\{\mathbf{R}^{\mathrm{A}(k)}_{\mathrm{s}_j,\mathrm{corr},i}\right\}_{k=1}^{N} \quad j = 1, 2, \quad i = 1, \ldots, 16, \quad (58)$$

are derived from the transformation (18) based on the random angular state points specified by (57). For *Case (ii)*, the corresponding set of baseline uncorrelated angular error covariances

$$\left\{\mathbf{R}^{\mathrm{A}(k)}_{\mathrm{s}_j,\mathrm{uncorr},i}\right\}_{k=1}^{N} \quad j = 1, 2, \quad i = 1, \ldots, 16, \quad (59)$$

assume fixed values as specified by (16).

Subsequently, we apply equations (33)–(48) and (52) to derive the composite target estimates

$$\left\{\hat{\mathbf{x}}^{\mathrm{F}(k)}_{\mathrm{corr},i}, \mathbf{P}^{\mathrm{F}(k)}_{\mathrm{corr},i}\right\}_{k=1}^{N} \quad i = 1, \ldots, 16, \quad (60)$$

for *Case (i)*. The kth composite position estimate of the i^{th} target $\hat{\mathbf{x}}^{\mathrm{F}(k)}_{\mathrm{corr},i}$ is based on the fusion of the kth random angular points, and their respective covariances from camera sensors s_1 and s_2, $\left(\mathbf{x}^{\mathrm{A}(k)}_{\mathrm{s}_1,i}, \mathbf{R}^{\mathrm{A}(k)}_{\mathrm{s}_1,\mathrm{corr},i}\right)$ and $\left(\mathbf{x}^{\mathrm{A}(k)}_{\mathrm{s}_2,i}, \mathbf{R}^{\mathrm{A}(k)}_{\mathrm{s}_2,\mathrm{corr},i}\right)$, specified in (57) and (58). The corresponding kth Cartesian error covariance estimate of the ith target $\mathbf{P}^{\mathrm{F}(k)}_{\mathrm{corr},i}$ follows from (52) evaluated at $\mathbf{x}^{\mathrm{F}} = \hat{\mathbf{x}}^{\mathrm{F}(k)}_{\mathrm{corr},i}$.

In a similar manner, we apply equations (33)–(48) and (52) to derive the composite target estimates and their covariances

$$\left\{\hat{\mathbf{x}}^{\mathrm{F}(k)}_{\mathrm{uncorr},i}, \mathbf{P}^{\mathrm{F}(k)}_{\mathrm{uncorr},i}\right\}_{k=1}^{N} \quad i = 1, \ldots, 16 \quad (61)$$

for *Case (ii)*, based on the random angular point sets specified by (57) and (59), and with (52) evaluated at $\mathbf{x}^{\mathrm{F}} = \hat{\mathbf{x}}^{\mathrm{F}(k)}_{\mathrm{uncorr},i}$.

We calculate the differences of the true composite target state with the correlated variant of the composite position estimate

$$\Delta\mathbf{x}^{\mathrm{F}(k)}_{\mathrm{corr},i} = \left(\hat{\mathbf{x}}^{\mathrm{F}(k)}_{\mathrm{corr},i} - \breve{\mathbf{x}}^{\mathrm{F}}_i\right)' \left(\hat{\mathbf{x}}^{\mathrm{F}(k)}_{\mathrm{corr},i} - \breve{\mathbf{x}}^{\mathrm{F}}_i\right) \quad i = 1, \ldots, 16, \quad (62)$$

and the uncorrelated variant of the composite position estimate

$$\Delta\mathbf{x}^{\mathrm{F}(k)}_{\mathrm{uncorr},i} = \left(\hat{\mathbf{x}}^{\mathrm{F}(k)}_{\mathrm{uncorr},i} - \breve{\mathbf{x}}^{\mathrm{F}}_i\right)' \left(\hat{\mathbf{x}}^{\mathrm{F}(k)}_{\mathrm{uncorr},i} - \breve{\mathbf{x}}^{\mathrm{F}}_i\right) \quad i = 1, \ldots, 16, \quad (63)$$

[5]These are the "simplistic" constant measurement accuracies.

where $\check{\mathbf{x}}_i^{\mathrm{F}}$ are the true target locations in ENU space.

Following the above, both correlated and uncorrelated variants of the overall root mean square error (RMSE) across $N = 1000$ sample runs are calculated in (64) and (65), respectively

$$\mathrm{RMSE}_{\mathrm{corr},i} = \sqrt{\frac{1}{N}\sum_{k=1}^{N} \Delta\mathbf{x}_{\mathrm{corr},i}^{\mathrm{F}(k)}} \qquad i = 1, \ldots, 16, \quad (64)$$

$$\mathrm{RMSE}_{\mathrm{uncorr},i} = \sqrt{\frac{1}{N}\sum_{k=1}^{N} \Delta\mathbf{x}_{\mathrm{uncorr},i}^{\mathrm{F}(k)}} \qquad i = 1, \ldots, 16. \quad (65)$$

Additionally, the percentage difference between the overall RMSE values of the correlated and uncorrelated variants of the 3D composite position estimates is calculated by

$$\mathrm{RMSE\_diff} = \frac{\mathrm{RMSE}_{\mathrm{corr}} - \mathrm{RMSE}_{\mathrm{uncorr}}}{\mathrm{RMSE}_{\mathrm{uncorr}}} \times 100\%, \quad (66)$$

and the results are tabulated in Table VII for each of the 16 selected target points.

It is observed that our approach using the more appropriate correlated angular error covariance matrices yields up to a maximum of approximately 0.7% improvement in 3D positional accuracy relative to the composite position estimate derived from the baseline uncorrelated angular error ellipses. Across all targets, the 3D positional accuracy differences are insignificant. This is because in both correlated and uncorrelated cases, we used the same transformation in (17) to convert measurements from the display image space to the azimuth-elevation space. The azimuth and elevation values for each point set were thus identical in both cases. As a result, the triangulation estimates of the two methods are similar. Although the measurement error covariances in (58) and (59) differ, their contributions do not significantly impact the RMSE. While Table VII showed little differences, Table VI showed significant differences for the algorithm-calculated covariances (which are used in recursive state estimation).

3) Statistical Consistency: To test the statistical consistency of the correlated variant of the composite position estimate, the normalized estimation error squared (NEES) [2] is used, with the CRLB $\mathbf{P}_{\mathrm{corr},i}^{\mathrm{F}}$ as the covariance matrix, i.e., the inverse of the FIM in (52) evaluated at $\mathbf{x}^{\mathrm{F}} = \check{\mathbf{x}}_{\mathrm{corr},i}^{\mathrm{F}}$ for the ith target point, $i = 1, \ldots, 16$. Along with definition (62), we have that

$$\epsilon_{\mathrm{corr},i}^{\mathrm{F}(k)} = \left(\Delta\mathbf{x}_{\mathrm{corr},i}^{\mathrm{F}(k)}\right)' \mathbf{P}_{\mathrm{corr},i}^{\mathrm{F}\ -1} \left(\Delta\mathbf{x}_{\mathrm{corr},i}^{\mathrm{F}(k)}\right) \sim \chi_{n_x}^2, \quad (67)$$

and thus the chi-squared test statistic

$$\bar{\epsilon}_{\mathrm{corr},i}^{\mathrm{F}} = \frac{1}{N}\sum_{k=1}^{N} \epsilon_{\mathrm{corr},i}^{\mathrm{F}(k)} \sim \frac{1}{N}\chi_{Nn_x}^2, \quad (68)$$

With $N = 1000, n_x = 3$ and $\alpha = 0.95$, the two-tailed 95% interval is given by

$$\bar{\epsilon}_{\mathrm{corr},i}^{\mathrm{F}} \in [r_1, r_2] \quad i = 1, \ldots, 16, \quad (69)$$

where $[r_1, r_2] = \left[\frac{1}{1000}\chi_{3000}^2(0.025), \frac{1}{1000}\chi_{3000}^2(0.975)\right] = [2.8501, 3.1537]$.

Similarly, we repeat steps (67)–(69) to test the statistical consistency of the baseline uncorrelated variant of the composite position estimate. We still evaluate this uncorrelated variant of the estimate at the same CRLB $\mathbf{P}_{\mathrm{corr},i}^{\mathrm{F}}$, and replace definition (62) with (63) to derive the uncorrelated versions $\epsilon_{\mathrm{uncorr},i}^{\mathrm{F}(k)}$ and $\bar{\epsilon}_{\mathrm{uncorr},i}^{\mathrm{F}}$. Note that $\bar{\epsilon}_{\mathrm{uncorr},i}^{\mathrm{F}}$ shares the same two-tailed 95% test interval as $\bar{\epsilon}_{\mathrm{corr},i}^{\mathrm{F}}$.

Figure 7 shows the plot of the sample average NEES from $N = 1000$ runs of the correlated (red-filled circles) and uncorrelated (black-filled triangles) variants of the composite position estimate, using the same appropriate CRLB as the covariance matrix, for the 16 selected targets. The dotted lines within the plot indicate the 95% two-tailed interval calculated as per (69). Since all chi-squared test statistics across all 16 targets fall within the 95% two-tailed interval, the composite position estimates derived from both our correlated and the uncorrelated baseline approaches are statistically consistent.

Thus, the benefit of the proposed method of fusion is in the correctly calculated uncertainty region volumes presented in Table VI, even though the estimates of the

Table VII
Absolute RMSE Values and Percentage Difference between RMSE Values of the Correlated and Uncorrelated Variants of the Composite 3D Cartesian Position Estimates Evaluated at Each of 16 True Target Points $\check{\mathbf{x}}_i^{\mathrm{F}}, i = 1, \ldots, 16$

Tgt	1	2	3	4	5	6	7	8
RMSE [corr] (m)	1.127	1.191	1.463	1.745	3.534	3.650	4.957	5.945
RMSE [uncorr] (m)	1.131	1.192	1.466	1.758	3.536	3.656	4.968	5.958
RMSE diff. (%)	0.332	0.134	0.209	0.714	0.045	0.179	0.221	0.219
Tgt	9	10	11	12	13	14	15	16
RMSE [corr] (m)	7.495	7.814	9.210	12.136	13.017	13.731	15.285	18.008
RMSE [uncorr] (m)	7.484	7.820	9.209	12.146	13.025	13.720	15.280	18.019
RMSE diff. (%)	− 0.150	0.065	− 0.005	0.082	0.058	− 0.076	− 0.031	0.062

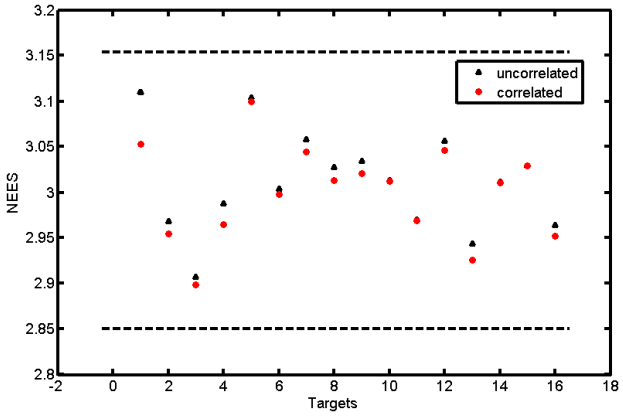

Figure 7. Sample average NEES for correlated and uncorrelated variants of the composite position estimate for targets $i = 1, \ldots, 16$.

proposed and baseline approaches are not significantly different.

VI. CONCLUSIONS

In this paper, we devised an approach to convert the raw target measurements and their covariance in the image space of EO cameras to the azimuth and elevation angle measurements in the global ENU coordinate system, which is validated to be unbiased, with FPA-location-dependent variances and cross-correlated errors. Typical tracking applications usually assume that azimuth and elevation measurement errors have the same standard deviation are uncorrelated and uniform across the camera's FOV, which is not the case. Simulations were performed for nine selected points on the FPA coming from both 2MP and 8MP commercially available camera sensors, with comparisons of the bias error and correlated covariances to an uncorrelated baseline error measure. Results show that cross-correlated errors are present and significant enough to be considered. The magnitudes of variances and correlations are dependent on the target's location with respect to the camera center. In addition, the angular estimate derived from the conversion has been shown to be unbiased, and the derived covariance matrix with cross-correlated errors has been shown to be consistent, based on the statistical tests conducted.

With these findings, we envisioned an unbiased converted angular measurement and a more appropriate measurement error covariance accounting for cross-correlated errors should be used as input to the 3D Cartesian state filtering equations and multisensor fusion for improved tracking performance. To demonstrate this, we considered a multisensor fusion application to derive a composite target position measurement and its corresponding Cartesian measurement error covariance in 3D Cartesian space. This was done for a two-camera setup, using LOS target angle measurements with the more appropriate measurement error covariances of our method. Simulations were performed for 16 selected target points in ENU space. Results show that the composite 3D target position and its corresponding Cartesian measurement error covariance estimates derived using the LOS angular measurements and angular measurement error covariance of our approach are statistically efficient. Compared to the uncorrelated LOS angular error covariances, our approach in accounting for location-dependent variances and cross-correlated errors in the LOS angular error covariances translated to a significant improvement to the error modeling of the Cartesian measurement error ellipse, but minimal improvement to the 3D Cartesian position accuracy of the target. Since in filtering the gain depends on the measurement covariance, the proposed approach will yield superior performance. The impact of using the correlated covariance for the measurements on the filtering performance is a topic of future research. It is envisaged that the correct error covariance is crucial in real-world applications, particularly for tracking in cluttered environments or closely spaced targets. Accurate measurement error covariance enhances track-to-measurement association accuracy. The additional computational cost, which is not significant, is well worth it.

REFERENCES

[1] B. Balaji, R. Sithiravel, Z. Daya, and T. Kirubarajan "Aspects of detection and tracking of ground targets from an airborne EO/IR sensor," *Proc. SPIE*, vol. 9474, Art. no. 947403.

[2] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan *Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software*. New York, NY, USA: Wiley, 2001.

[3] Y. Bar-Shalom, P. Willett, and X. Tian *Tracking and Data Fusion: A. Handbook of Algorithms*. Tulsa, OK, USA: YBS Publishing, 2011.

[4] X. Liu, W. Chen, H. Madhusudanan, L. Du, and Y. Sun "Camera orientation optimization in stereo vision systems for low measurement error," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 2, pp. 1178–1182, Apr. 2021.

[5] Q. Lu, Y. Bar-Shalom, P. Willett, and B. Balasingam "Measurement extraction for a point target from an optical sensor," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 6, pp. 2735–2745, Dec. 2018.

[6] M. Mallick "Geolocation using video sensor measurements," *Proc. 10th Int. Conf. Inf. Fusion*, Quebec, QC, Canada, 2007, pp. 1–8.

[7] M. Mallick, X. Tian, Y. Zhu, and M. Morelande "Angle-only filtering of a maneuvering target in 3D," *Sensors*, vol. 22, no. 4, Art. no. 1422, 2022.

[8] L. Mejias, J. J. Ford, and J. Lai "Towards the implementation of vision-based UAS sense-and-avoid system," Computational Statistics & Data Analysis, *Proc. 27th Int. Congr. Aeronautical Sci., Acropolis Conference Centre*, Nice, France, Sep. 2010.

[9] A. Miller and B. Miller "Tracking of the UAV trajectory on the basis of bearing-only observations,"

in *Proc. 53rd IEEE Conf. Decis. Control*, Los Angeles, CA, USA, 2014, pp. 4178–4184.

[10] S. Omar and S. Winberg
"Multisensor data fusion: Target tracking with a Doppler radar and an electro-optic camera,"
Proc. 2011 IEEE Int. Conf. Control System, Computing and Engineering, Penang, Malaysia, 2011, pp. 210–215.

[11] R. W. Osborne and Y. Bar-Shalom
"Statistical efficiency of composite position measurements from passive sensors,"
IEEE Trans. Aerosp. Electron. Syst., vol. 49, no. 4, pp. 2799–2806, Oct. 2013.

[12] R. Yang and Y. Bar-Shalom
"Full state information transfer across adjacent cameras in a network using gauss helmert filters,"
J. Adv. Inf. Fusion, 17(1), pp. 14–28, Jun. 2022.

[13] R. Yang, Y. Bar-Shalom, and H. A. J. Huang
"Camera calibration using inaccurate and asynchronous discrete GPS trajectory from drones,"
J. Adv. Inf. Fusion, vol. 18, no. 1, pp. 3–15, Jun. 2023.

Jessica Koon Yan Goh received the B.Sc. and Master of Mathematics and Statistics degree from the University of Warwick, Coventry, UK, in 2017. She is currently a Senior Member of Technical Staff at DSO National Laboratories, Singapore. Her research interests include target tracking and multisensor data fusion.

Yaakov Bar-Shalom (F'84) received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion in 1963, Haifa, Israel and 1967, respectively, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 1970. He is currently a Board of Trustees Distinguished Professor with the ECE Dept. and Marianne E. Klewin Professor at the University of Connecticut, Storrs, CT, USA. His current research interests are in estimation theory, target tracking, and data fusion. He has published more than 650 papers and book chapters. He coauthored/edited eight books, including *Tracking and Data Fusion* (YBS Publishing, 2011). He has been elected Fellow of IEEE for "contributions to the theory of stochastic systems and of multitarget tracking." He served as an Associate Editor for the Transactions on Automatic Control and Automatica. He was General Chairman of the 1985 ACC, General Chairman of FUSION 2000, President of ISIF in 2000 and 2002 and Vice President for Publications from 2004 to 2013. Since 1995, he has been a Distinguished Lecturer of the IEEE AESS. He is a corecipient of the M. Barry Carlton Award for the best paper in the IEEE TAE Systems in 1995 and 2000. In 2002, he received the J. Mignona Data Fusion Award from the DoD JDL Data Fusion Group. He is a member of the Connecticut Academy of Science and Engineering. In 2008, he was awarded the IEEE Dennis J. Picard Medal for Radar Technologies and Applications, and in 2022, the IEEE AESS Pioneer Award. He has been listed by academic.research.microsoft (top authors in engineering) as #1 among the researchers in aerospace engineering based on the citations of his work. He is the recipient of the 2015 ISIF Award for a Lifetime of Excellence in Information Fusion. This award has been renamed in 2016 as the Yaakov Bar-Shalom Award for a Lifetime of Excellence in Information Fusion. He has the following Wikipedia page: https://en.wikipedia.org/wiki/Yaakov Bar-Shalom.

Rong Yang received her B.E. degree in information and control from Xi'an Jiao Tong University, Xi'An China, in 1986, M.Sc. degree in electrical engineering from National University of Singapore, Singapore, in 2000, and Ph.D. degree in electrical engineering from Nanyang Technological University, Singapore, in 2012. She is currently a Principal Member of Technical Staff at DSO National Laboratories, Singapore. Her research interests include passive tracking, low observable target tracking, GMTI tracking, hybrid dynamic estimation, and data fusion. She was Publicity and Publication Chair of FUSION 2012 and received the FUSION 2014 Best Paper Award (first runner up).