ASYMPTOTICALLY SMALL GEOPOSITIONING ERROR FROM LARGE SETS OF IMAGERY

Reuben Settergren
Seth Merickel
rjs@jhu.edu
sethmerickel@gmail.com

ABSTRACT

The geolocation error resulting from least-squares Multi-Image Geopositioning (MIG) with n images is shown theoretically and empirically to decrease as $1/\sqrt{n}$, using a synthetically augmented set of 1000 overlapping Worldview-1 images. A novel heuristic approach called 'Hourglassing' is introduced, which requires no understanding of apriori covariance, but implicitly infers uncertainty from the distribution of the ray bundle itself. Hourglassing geolocation and error estimation is compared to MIG using the 1000-image testbed.

1. Introduction

Least-squares Multi-Image Geopositioning (hereafter MIG) is well-understood in photogrammetry[11], [6], [10], providing geolocation and output covariance that are as accurate as the apriori projection and covariance data that are input into the process. However, for practical reasons, the number of images that have been used in MIG calculations have been rather small. Even after the advent of computing enabled the automatic processing of many images, there still remained the practical problem of *having* very many images that observed the same ground feature (as well as meaningful apriori covariance models). A typical airborne collection with a mapping camera and 60/40 forward/side-lap configuration would yield at most 6 views. A collection with large satellite images is usually designed to have minimally sufficient overlap, yielding only 1- or 2-deep coverage almost everywhere (depending on a mono or stereo collect), with imagery being 2- or 4-deep at the overlaps only to ensure there are no gaps.

Some authors have applied MIG with more extensive collections of imagery. In [9], Jeong, Yang, and Kim analyze MIG with one pair of imagery from each of three different satellites: IKONOS, QuickBird, and KOMPSAT-2. In [2], R. T. Collins geopositions with 7 images simultaneously (although not with MIG, but a technique called 'Space-Sweep' – which is similar to the Hourglassing method introduced in section 3). In [3], a team from the Chinese Academy of Sciences investigates MIG accuracy with 2–9 images of the moon from the Lunar Reconnaissance Orbiter. Wonnacott [16] analyzes MIG accuracy with SAR, using 32 image 'sets' (each set being a 'three-aperture path sequence').

These image collections, however, represent only a fraction of what is currently possible, given the ubiquity of high-resolution commercial satellite imaging from companies like Digital Globe, Airbus, SPOT, Pléiades, KOMPSAT, COSMO-SkyMed, TerraSAR-X, etc. A search of the Digital Globe online imagery catalog [4] reveals well over 100 images for many large cities. And the current wave of SmallSat flocks (Planet Labs, RapidEye, etc.) with their rapid revisit times, promise to provide floods of repeat imagery. Although SmallSats may be inferior in resolution and accuracy at the moment, technology will improve, and it is the goal of this paper to show that these deficiencies can be overcome by the power of large amounts of statistically independent information from huge numbers of images.

2. Least-Squares MIG With Many Images

2.1. MIG Solution.

We begin with a derivation of least-squares MIG. For further details of this procedure see the discussion of the Gauss-Helmert model in [11], [6], [10].

In general MIG can determine the locations and joint covariance of many ground points, but here we are concerned with just one point. Also, MIG is capable of appropriately handling known correlations, but for our purposes we assume full independence between all images and measurements.

The goal is therefore to determine the ground location that minimizes the weighted sum of image measurement and sensor parameter deviations using non-linear least squares. Let $\boldsymbol{X} = [X, Y, Z]'$ be an apriori estimate of the position of the ground point being geolocated. $\boldsymbol{x_i} = [x_i, y_i]'$ is the measurement of the

Thanks to Digital Globe for use of the 10 Worldview-1 images from which the testbed in section 2.3 was synthesized.
Special thanks also to Tim Nagy, whose original idea behind 'Hourglassing' he encouraged us to develop.

ground point in the ith image. The sensor parameters for the ith image are $\boldsymbol{p_i} = [p_{i1}, p_{i2}, \ldots]'$. The ground point, image measurements, and sensor parameters are related through the sensor's ground to image function $\boldsymbol{x_i} = \boldsymbol{g_i}(\boldsymbol{p_i}, \boldsymbol{X})$. Let $\boldsymbol{F_i}$ be the measurement residual function for the ith image, which is the difference between the projected ground point and image measurement:

$$\boldsymbol{F_i}(\hat{\boldsymbol{p}}_{\boldsymbol{i}}, \hat{\boldsymbol{X}}, \hat{\boldsymbol{x}}_{\boldsymbol{i}}) = \boldsymbol{g_i}(\hat{\boldsymbol{p}}_{\boldsymbol{i}}, \hat{\boldsymbol{X}}) - \hat{\boldsymbol{x}}_{\boldsymbol{i}},$$

where $\hat{\boldsymbol{x}}_{\boldsymbol{i}}$, $\hat{\boldsymbol{p}}_{\boldsymbol{i}}$, and $\hat{\boldsymbol{X}}$ are measured quantities related to the predicted quantities by:

$$\hat{\boldsymbol{x}}_{\boldsymbol{i}} = \boldsymbol{x_i} + \boldsymbol{\nu_i^m}$$
$$\hat{\boldsymbol{p}}_{\boldsymbol{i}} = \boldsymbol{p_i} + \boldsymbol{\nu_i^p}$$
$$\hat{\boldsymbol{X}} = \boldsymbol{X} + \boldsymbol{\Delta}$$

$\boldsymbol{\nu_i^m}$ and $\boldsymbol{\nu_i^p}$ are differences in the predicted and measured image coordinates and sensor parameters respectively. $\boldsymbol{\Delta}$ will be the correction to the ground point, for which we are solving.

Taylor expanding $\boldsymbol{F_i}$ about the predicted values gives

$$\boldsymbol{F_i}(\hat{\boldsymbol{p}}_{\boldsymbol{i}}, \hat{\boldsymbol{X}}, \hat{\boldsymbol{x}}_{\boldsymbol{i}}) = \boldsymbol{F_i}(\boldsymbol{p_i}, \boldsymbol{X}, \boldsymbol{x_i}) + \frac{\partial \boldsymbol{F_i}}{\partial \boldsymbol{x_i}} \boldsymbol{\nu_i^m} + \frac{\partial \boldsymbol{F_i}}{\partial \boldsymbol{p_i}} \boldsymbol{\nu_i^p} + \frac{\partial \boldsymbol{F_i}}{\partial \boldsymbol{X}} \boldsymbol{\Delta} = 0 \tag{1}$$

Sensor parameter partials are denoted by $A_i^p = \frac{\partial \boldsymbol{F_i}}{\partial \boldsymbol{p_i}}$, and the sensor ground partials by $B_i = \frac{\partial \boldsymbol{F_i}}{\partial \boldsymbol{X}}$. The partials with respect to image coordinates are $\frac{\partial \boldsymbol{F_i}}{\partial \boldsymbol{x_i}} = -I_{2\times 2}$, the negative of the 2×2 identity matrix. Substituting A_i and B_i into (1) and rearranging terms gives:

$$-\boldsymbol{\nu_i^m} + A_i \boldsymbol{\nu_i^p} + B_i \boldsymbol{\Delta} = \boldsymbol{x_i} - \boldsymbol{g_i}(\boldsymbol{p_i}, \boldsymbol{X}) = \boldsymbol{f_i} \tag{2}$$

Stacking up equations for each measurement into one vector equation gives

$$-\boldsymbol{\nu^x} + A\boldsymbol{\nu^p} + B\boldsymbol{\Delta} = \boldsymbol{f} \tag{3}$$

The function to be minimized is the weighted sum of the measurement deviations and sensor deviations given by $\Phi = \boldsymbol{\nu^{x\prime}} W^x \boldsymbol{\nu^x} + \boldsymbol{\nu^{p\prime}} W^p \boldsymbol{\nu^p}$ subject to the constraints in (3). W^x and W^p are inverses of the measurement covariance matrix and sensor parameter covariance matrix respectively. Minimizing Φ using Lagrange multipliers to account for the constraint specified in (3) results in the following normal equations:

$$B'WB\boldsymbol{\Delta} = B'W\boldsymbol{f} \tag{4}$$

where W is the sensor parameter and image measurement error combined into a single weight matrix:

$$W = (W^{x-1} + AW^{p-1}A')^{-1} \tag{5}$$

This linear system can be solved for Δ, which is used to correct the ground estimate $\boldsymbol{X}$. To the extent that partial derivative matrices A and B evaluated at $\boldsymbol{X}$ vs $\boldsymbol{X} + \Delta$ differ from their values at $\boldsymbol{X}$, this is a non-linear problem, so iteration is used to converge to a solution. After each iteration the ground point locations from the previous iteration are updated $\boldsymbol{X^k} = \boldsymbol{X^{k-1}} + \boldsymbol{\Delta^k}$, and the partial derivative matrices A and B are evaluated at the new ground point location for the next iteration. Convergence is typically achieved after only a few iterations. (However for the experiments in section 2.3, the partial derivatives were so stable, only one iteration was necessary; the 2nd iteration yielded only negligible correction.)

2.2. **MIG Covariance.** From the normal equations (4) it is straightforward to show that the covariance of the ground point estimate is:

$$C^g = (B'WB)^{-1} \tag{6}$$

To understand how C^g depends on the number of images n, we establish an upper bound on C^g and show that the variance of each component decays as $1/N$ where N is the number of images. As a metric for comparing covariance matrices, we use the volume of the standard ellipsoid $(X - \bar{X})'C^g(X - \bar{X}) = 1$. Note that we do not need to differentiate between identically-sized ellipsoids that differ only by rotation. The volume of an ellipsoid is $\pi/4 \cdot abc$, where a, b, and c are the lengths of the three axes of the ellipsoid, which are the eigenvalues λ_1, λ_2, and λ_3 of the covariance matrix (We use the convention $\lambda_1 \leq \lambda_2 \leq \lambda_3$, and since covariance matrices are positive definite, $\lambda_1 > 0$). Using the fact that the determinant of a positive definite matrix is the product of its eigenvalues we define: $C_1 \geq C_2 \leftrightarrow |C_1| \geq |C_2|$

Using (6), and the properties of the determinant, we have $|C| = |B'WB|^{-1}$. Since we want an upper bound on $|C|$, we need to demonstrate a lower bound for $|B'WB|$. Recall the initial assumption of independence of all image measurements, and sensor parameters for different images. This means the image measurement covariance matrix C^m and sensor parameter covariance matrix C^p are block diagonal. Thus we have:

$$\begin{aligned} B'WB &= \sum_i B_i'(C_i^m + A_i C_i^p A_i')^{-1} B_i \\ &= \sum_i B_i' W_i B_i \end{aligned} \tag{7}$$

where $W_i = (C_i^m + A_i C_i^p A_i')^{-1}$.

If A is an $n \times n$ matrix then let $\lambda_i(A)$ be the ith smallest eigenvalue of A (and in particular $\lambda_1(A)$ is the smallest eigenvalue). In (7) each image contributes a term of $B_i' W_i B_i$ to the sum. Because W_i is 2×2 and B_i is 2×3, each $B_i' W_i B_i$ is a rank-deficient 3×3. The two eigenvectors of $B_i' W_i B_i$ with nonzero eigenvalues are the rows of B_i (the ground-space projections of the line and sample directions in image space), and the eigenvector corresponding to the eigenvalue $\lambda_1(B_i' W_i B_i) = 0$ is their cross-product, the line of sight ray projecting from the image measurement, out of the sensor (or perhaps into it, sign is not important).

Assume two images and measurements i and j such that the rays from those measurements are identical/parallel. Call that vector l. Then

$$\begin{aligned} l'(B_i' W_i B_i + B_j' W_j B_j) l &= l' B_i' W_i B_i l + l' B_j' W_j B_j l \\ &= 0 + 0 \end{aligned}$$

Thus l is also a trivial eigenvector of $B_i' W_i B_i + B_j' W_j B_j$, and $\lambda_1(B_i' W_i B_i + B_j' W_j B_j) = 0$.

However, since in reality no two images will be collected with perfectly parallel image rays, we can safely assume that for all $i \neq j \in 1 \ldots n$, $\lambda_1(B_i' W_i B_i + B_j' W_j B_j) > 0$. Let

$$\delta = \min_i \lambda_1(B_i' W_i B_i + B_{i+1}' W_{i+1} B_{i+1}) > 0 \tag{8}$$

be the smallest λ_1 of the sum of *consecutive* $B_i' W_i B_i$. Note that even for an infinite collection of images, it is possible for $\delta > 0$, if care is taken ordering the images. For instance, consider a fixed stereo collection geometry, with arbitrarily many images collected from the exact position and orientation of the left and right camera stations. As long as the images are ordered alternating left, right, then

$$\begin{aligned} \delta &= \min_i \lambda_1(B_i' W_i B_i + B_{i+1}' W_{i+1} B_{i+1}) \\ &= \lambda_1(B_1' W_1 B_1 + B_2' W_2 B_2) > 0 \end{aligned}$$

Horn and Johnson in [8] give a useful property of eigenvalues of positive semidefinite matrices A and B: for $l \leq k, \lambda_{k-l+1}(A) + \lambda_l(B) \leq \lambda_k(A + B)$. In particular, for $k = l = 1$,

$$\lambda_1(B_i' W_i B_i) + \lambda_1(B_j' W_j B_j) \leq \lambda_1(B_i' W_i B_i + B_j' W_j B_j) \tag{9}$$

Using (9), we can prove the following theorem:

Theorem 1. *For $\delta > 0$, if $\lambda_1(B_i' W_i B_i + B_{i+1}' W_{i+1} B_{i+1}) \geq \delta$ for all $i = 1 \ldots n-1$, then $\lambda_1(\sum_i B_i' W_i B_i) \geq \frac{n\delta}{2}$ for n even and $\geq \frac{(n-1)\delta}{2}$ for n odd.*

This can be proven using induction. A few base cases make the rest of the proof clear. For $n = 2$:

$$\lambda_1(B_1' W_1 B_1 + B_2' W_2 B_2) \geq \delta = \frac{n\delta}{2}$$

For $n = 3$:

$$\begin{aligned} \lambda_1(B_1' W_1 B_1 + B_2' W_2 B_2 + B_3' W_3 B_3) &\geq \lambda_1(B_1' W_1 B_1 + B_2' W_2 B_2) + \lambda_1(B_3' W_3 B_3) \\ &\geq \delta + 0 \\ &= \frac{(n-1)\delta}{2} \end{aligned}$$

The pattern above illustrates the nature of the inductive step. For n even,

$$\begin{aligned}\lambda_1(\sum_{i=1}^{n} B_i' W_i B_i) &= \lambda_1(B_1' W_1 B_1 + B_2' W_2 B_2 + \sum_{i=3}^{n} B_i' W_i B_i) \\ &\geq \lambda_1(B_1' W_1 B_1 + B_2' W_2 B_2) + \lambda_1(\sum_{i=3}^{n} B_i' W_i B_i) \\ &\geq \delta + \frac{(n-2)\delta}{2} \\ &= \frac{n\delta}{2}\end{aligned}$$

The proof for odd n is similar. □

Since Theorem 1 guarantees that the smallest eigenvalue $\lambda_1(\sum_i B_i' W_i B_i) \geq {}^{(n-1)\delta}/_2$, the other eigenvalues are still larger: $\lambda_3 \geq \lambda_2 \geq \lambda_1 \geq {}^{(n-1)\delta}/_2$. Recall ground covariance $C^g = (\sum_i B_i' W_i B_i)^{-1}$, so its eigenvalues are reciprocal, thus

$$\lambda_i(C^g) \leq \frac{2}{(n-1)\delta} \tag{10}$$

These eigenvalues are the variances of ground uncertainty in the directions of the three axes (eigenvectors). (10) demonstrates that the variance of ground error is bounded above by $O(1/n)$, and decreases like $1/n$ as n grows. This is equivalent to standard deviation decreasing like $O({}^1/_{\sqrt{n}})$, a result which is consistent with the sweepingly general applicability of the Law of Large Numbers [1].

2.3. **Simulation with Synthetic Imagery.** In this section we construct a large testbed of simulated/synthetic imagery, and evaluate the accuracy of MIG geolocation and error estimation, on samples of randomly-selected subsets of images, of size ranging from 4 to 1000.

To develop the test set, we start with a large as possible set of real sensor models which have common overlap. (Note 'sensor models', not 'images.' For this simulated experiment, no pixels are needed, only sensor models. Actual scene content is irrelevant.) From [5] we have sensor models from 10 Worldview-1 images which all view the ground location 36N 117.5W 1700mHAE (WGS84). This point is set as Ground Truth. (The sensor model used in this study is the SOCET GXP [12] Worldview sensor model.) We then synthetically extend the set of real sensor models by randomly perturbing position and orientation adjustable parameters with very large corrections. Using this random perturbation technique, we generate 99 perturbations for each of the 10 original images, for a total of 1000 sensor models.

The simulated set of 1000 images thus constructed is considered the Truth set of sensor models. The sensor model's ground to image function is used to project the truth point into image coordinates for each sensor model. These image measurements are retained as truth also, and used for all the following experimental ray-intersections. The ray bundle emanating from these truth image points intersects perfectly (to within computational precision of the image to ground function) at the truth point in ground space. Let this idealized set of images and measurements be $\mathcal{I}$.

Once the ideal bundle $\mathcal{I}$ is assembled, an image bundle with realistic errors can be constructed by perturbing sensor model parameters from $\mathcal{I}$ with a controlled amount of error (randomly sampled from a known distribution), to obtain perturbed image set $\mathcal{P}$. Nine position/velocity/acceleration parametrs are given sigmas of $\sigma = 1\text{m}, 0.1\text{m/s}, 0.01\text{m}^2/\text{s}^2$, and 9 attitude/rate/acceleration parameters are given sigmas of $\sigma = 5\mu\text{rad}, 0.5\mu\text{rad/s}, 0.05\mu\text{rad/s}^2$. This model yields a CE90 of about 6-8m for the original 10 images, depending on obliquity. Any subset of images of $\mathcal{P}$ (along with the truth measurements on those images) can be used as a realistic input to least-squares MIG. Apriori covariance input to the MIG is the same error model that was used to perturb $\mathcal{I}$ into $\mathcal{P}$.

Let $\mathcal{P}$ be the set of 1000 perturbed images. For each $n \in \{4, 5, \ldots\ 100, 105, \ldots 995\}$, we randomly sample $k = 100$ subsets of size n images from $\mathcal{P}$. Least-squares MIG is used to estimate the geolocation of the truth point using each n-image subset. Because we know the truth ground point (the trivial intersection of $\mathcal{I}$), we can compute the actual error of each of these MIGs, and observe the distribution of error as n increases.

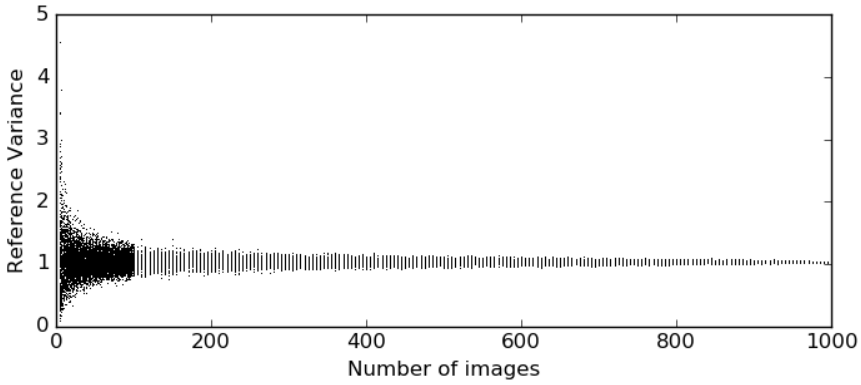

FIGURE 1. Reference variances for every MIG intersection. For every n, each of the $k = 100$ MIG generates one pixel in the graph.

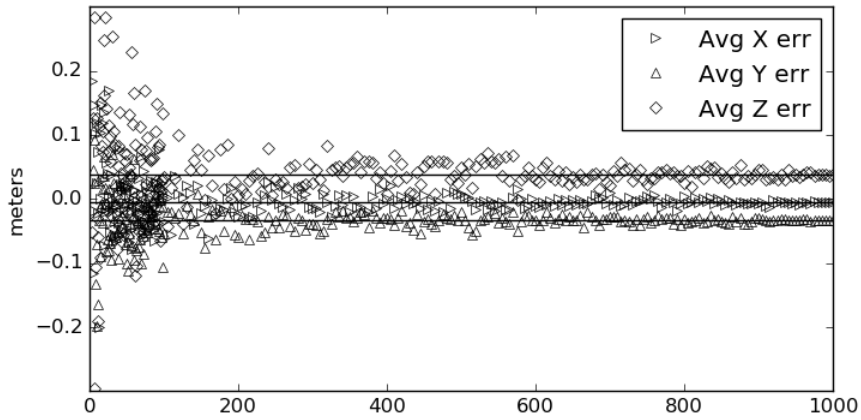

FIGURE 2. Average Errors in X, Y, and Z. The horizontal lines represent the MIG solution for all $n = 1000$ images of $\mathcal{P}$: (-0.0049, -0.0330, 0.0378), to which the subset MIGs converge.

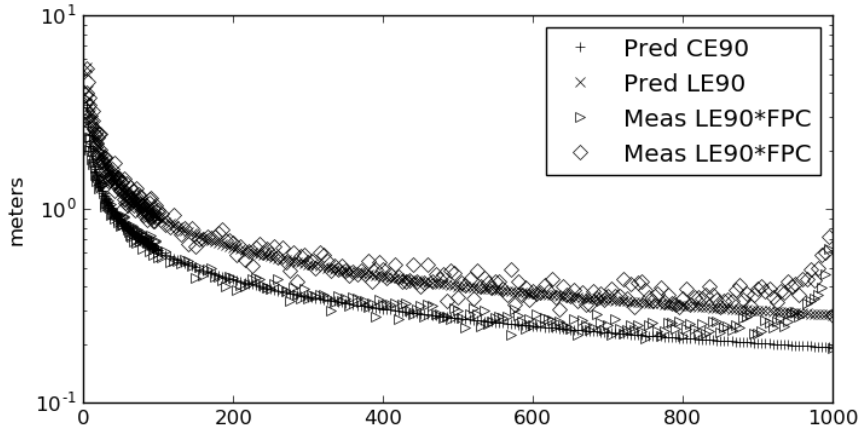

FIGURE 3. Predicted vs measured CE90/LE90 (log scale). Measured values are adjusted with a finite population correction.

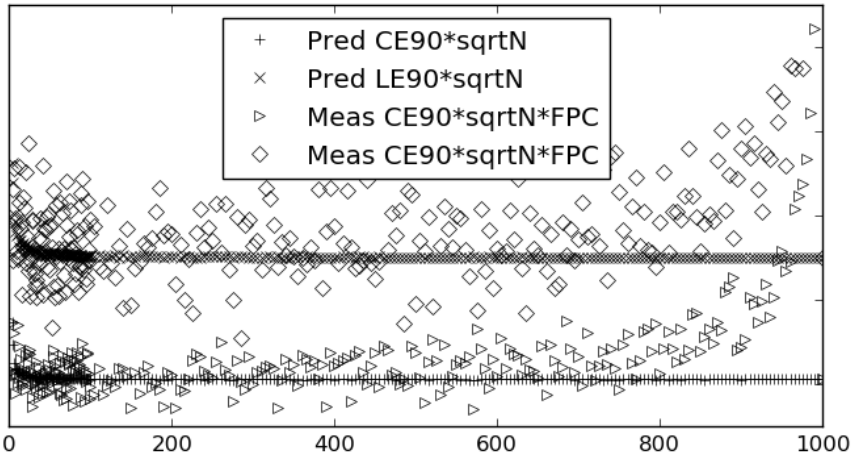

FIGURE 4. Predicted vs measured CE90/LE90, scaled by $\sqrt{n}$.

Figure 1 shows the reference variances of every MIG from this experiment (a total of 27600 individual MIG calculations). Reference variance is tightly clustered around 1, which demonstrates that the use of apriori covariance in the MIG calculations is consistent with the errors involved in the perturbation of $\mathcal{P}$.

Figure 2 shows the average of MIG errors in the X, Y, and Z dimensions. Each point represents an average over $k = 100$ MIGs. Because errors in positive and negative directions cancel each other out, this graph shows the amount of bias in the MIG calculations across n-subsets. For all $n = 1000$ images of $\mathcal{P}$, the MIG solution differs from ground truth by (-0.0049, -0.0330, 0.0378) meters. Horizontal lines at those values indicate that as $n \to 1000$, bias across the $k = 100$ MIGs for each n converges to that error inherent to $\mathcal{P}$.

To avoid positive and negative errors canceling each other out, we switch to positive quantities. Figure 3a shows the average predicted and measured CE90/LE90 for each n across the $k = 100$ MIGs for that n. To clarify the rate of convergence and separation of the curves, a logarithmic scale is used. (The measured errors are adjusted with a finite population correction of $\sqrt{(1000-1)/(1000-n)}$ [15]. The deviation of the measured errors from prediction near 1000 is believed to be due to limitations of the finite population correction when applied to a 90% statistic rather than a 1-sigma standard deviation)

In order to demonstrate the $1/\sqrt{n}$ behavior of MIG error, Figure 3b scales each point from Figure 3a by a factor of $\sqrt{n}$. The flatness of these curves demonstrates the thesis of this paper, that MIG follows the Law of Large Numbers, and exhibits error that decreases as $1/\sqrt{n}$.

In an earlier study [13], MIG experiments were also run in a variety of variable conditions: images having different error models rather than all the same, increasing amounts of measurement error, restricting image subsets to certain convergence angle/obliquity, and in all cases, the MIG responded to the variations in intuitive ways, and $1/\sqrt{N}$ behavior was evident.

3. Hourglassing

3.1. **Motivation.** Because commonly-available sensor models (such as RPC) often lack a meaningful error model to input to MIG, and for computational simplicity, we introduce a heuristic approach as an alternative to rigorous least-squares MIG. Although when visualized 'up close' to the true answer, any particular tangle of rays may seem unclear about where the ground point should be localized, from a distant perspective, the ray bundle will be something like a cone, widening to the cluster of satellite (or airborne) perspective centers, narrowing at the ground point, and widening again beyond the ground point. We seek the answer at the narrowest point of the cone. In an ideal case, the ray bundle will intersect at a single point, which has cross-sectional area of 0. In a real-world case, the bundle will appear as a cone with a 'fat' intersection, or an hourglass. Thus the name 'hourglassing' to motivate the heuristic technique.

In 'plane-sweep stereo' [14], a plane is swept through a ground scene to find stereo focus depth, where conjugate rays are closest. The technique is expanded for multi-view 3D scene reconstruction in [2]. We adopt a similar approach here. Given a bundle of rays, we can intersect the bundle with planes of various heights, compute the collection of intersections of the ray bundle with each height plane, measure the spread

of the 2-D distribution of points, and choose the plane with the least spread to be the solution for the height of the desired ground point. For the horizontal location of the ground point, the natural choice is the mean of the intersection points in the optimal (spread-minimizing) plane.

Rather than attacking this problem with a brute force search for the spread-minimizing height by computing intersection sets at very many heights, and slicing height space sufficiently thin to achieve a desired vertical resolution, we attempt some theoretical underpinnings for this technique that will allow a more efficient and precise solution.

3.2. **Computing height of minimum spread.** Assume two heights $z_+ > z_-$, and assume a set of n 3D lines L_i, none of which is horizontal (without loss of generality, we use horizontal planes. If necessary, the ray bundle can be rotated into a coordinate system so that the 'center' of the ray bundle is vertical). Specify the lines by their intersections with the planes of heights $z_\pm$ at points (x^i_+, y^i_+, z_+) and (x^i_-, y^i_-, z_-), for $i = 1 \dots n$.

Define

$$x^i(\lambda) = \lambda x^i_+ + (1-\lambda)x^i_-$$
$$y^i(\lambda) = \lambda y^i_+ + (1-\lambda)y^i_-$$
$$z(\lambda) = \lambda z_+ + (1-\lambda)z_-$$

For any λ, $(x^i(\lambda), y^i(\lambda), z(\lambda))$ is a point on line L_i. Whether the point is an interpolation between or extrapolation beyond the $z_\pm$ planes depends on whether $0 \le \lambda \le 1$ (in either case, we will just use the term interpolated). Together, all the points $(x^i(\lambda), y^i(\lambda), z(\lambda))$ for $i = 1 \dots n$ represent the intersection of lines L_i with the plane with height $z(\lambda)$.

Note that the set of all means $(\bar{x}(\lambda), \bar{y}(\lambda), z(\lambda))$ comprise a line. For the plane at height $z(\lambda)$ that yields the smallest spread of points $(x^i(\lambda), y^i(\lambda))$ will be the point on that line which we choose as our answer. (If desired, this mean line between $(\bar{x}(1), \bar{y}(1), z_+)$ and $(\bar{x}(0), \bar{y}(0), z_-)$ can be computed first, and the entire bundle rotated so the mean line becomes vertical.)

The two-dimensional spread of the intersection set at a particular $z(\lambda)$ is a symmetric, positive definite, 2x2 covariance matrix:

$$M(\lambda) = \begin{bmatrix} var(x^i(\lambda)) & covar(x^i(\lambda), y^i(\lambda)) \\ - & var(y^i(\lambda)) \end{bmatrix}$$

It can be shown that the covariance of the interpolated points $(x^i(\lambda), y^i(\lambda))$ can be expressed in terms of variances of and covariances between the four elementary datasets $x^i_+, x^i_-, y^i_+, y^i_-$ as follows:

$$var(x^i(\lambda)) = \lambda^2\sigma^2_{x+x+} + \lambda(1-\lambda)(\sigma^2_{x+x-} + \sigma^2_{x-x+}) + (1-\lambda)^2\sigma^2_{x-x-} \tag{11}$$

$$var(y^i(\lambda)) = \lambda^2\sigma^2_{y+y+} + \lambda(1-\lambda)(\sigma^2_{y+y-} + \sigma^2_{y-y+}) + (1-\lambda)^2\sigma^2_{y-y-} \tag{12}$$

$$covar(x^i(\lambda), y^i(\lambda)) = \lambda^2\sigma^2_{x+y+} + \lambda(1-\lambda)(\sigma^2_{x+y-} + \sigma^2_{x-y+}) + (1-\lambda)^2\sigma^2_{x-y-} \tag{13}$$

where

$$\sigma^2_{x+x+} = var(x^i_+),$$
$$\sigma^2_{x+y-} = covar(x^i_+, y^i_-),$$

etc. Note in particular that, since all the σ^2 are functions only of constants $x^i_\pm, y^i_\pm$, and n, each of the 4 elements of $M(\lambda)$ is a quadratic function of λ.

If we denote the determinant of a matrix with $|\cdot|$, the area of the 1-sigma error ellipse of $M(\lambda)$ is

$$a(\lambda) = \pi\sqrt{|M(\lambda)|},$$

i.e. the square root of a quartic (4th degree) polynomial.

We seek to minimize $a(\lambda)$, which is equivalent to minimizing the quartic polynomial

$$\begin{aligned} d(\lambda) &= |M(\lambda)| \\ &= var(x^i(\lambda))var(y^i(\lambda)) - covar(x^i(\lambda), y^i(\lambda))^2 \end{aligned}$$

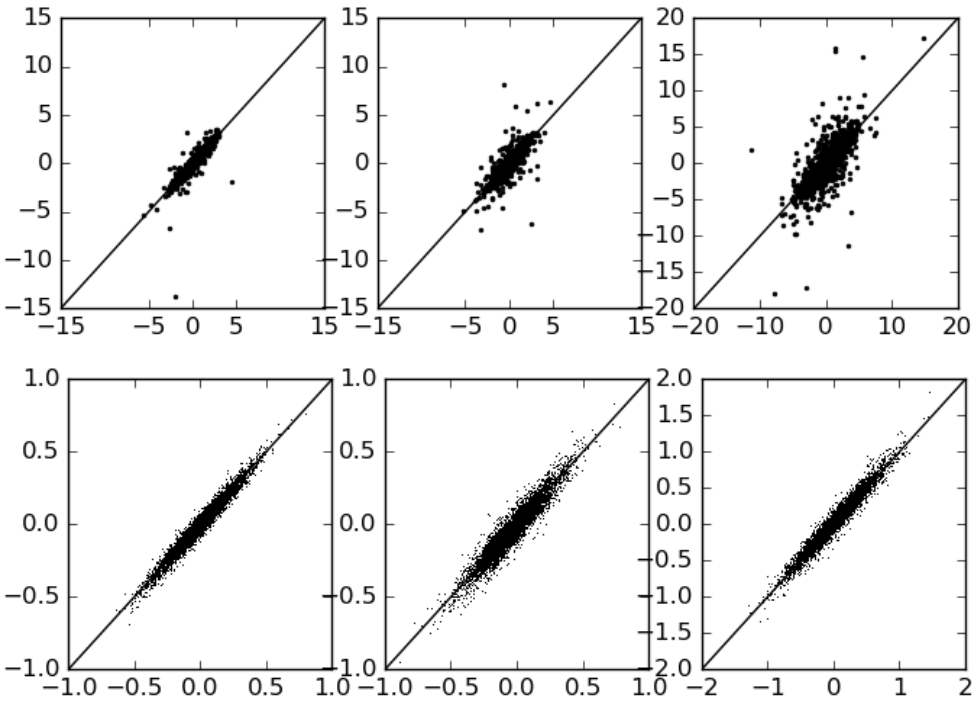

FIGURE 5. MIG vs Hourglass errors. All scales are meters. Each row is one graph each for X, Y, and Z errors. Above, 700 points are plotted from intersections with $4 \leq n \leq 100$. Below, 26900 points are plotted for $10 < n < 1000$.

Note that, for efficient computation, the quadratic, linear, and constant coefficients of equations (11-13) can be computed to restate more simply as:

$$var(x^i(\lambda)) = a_x\lambda^2 + b_x\lambda + c_x \tag{14}$$

$$var(y^i(\lambda)) = a_y\lambda^2 + b_y\lambda + c_y \tag{15}$$

$$covar(x^i(\lambda), y^i(\lambda)) = a_{xy}\lambda^2 + b_{xy}\lambda + c_{xy} \tag{16}$$

Given coefficients computed in equations (14-16), $d(\lambda)$ can then be expressed as

$$\begin{aligned} d(\lambda) &= (a_x\lambda^2 + b_x\lambda + c_x)(a_y\lambda^2 + b_y\lambda + c_y) - (a_{xy}\lambda^2 + b_{xy}\lambda + c_{xy})^2 \\ &= (a_xa_y - a_{xy}^2)\lambda^4 + (a_xb_y + b_za_y - 2a_{xy}b_{xy})\lambda^3 + \\ &\quad (a_xc_y + b_xb_y + c_xa_y - 2a_{xy}c_{xy} - b_{xy}^2)\lambda^2 + \\ &\quad (b_xc_y + c_xb_y - 2b_{xy}c_{xy})\lambda + (c_xc_y - c_{xy}^2) \end{aligned} \tag{17}$$

The value λ_{min} which minimizes $d(\lambda)$ will also determine the height $z(\lambda_{min})$ of the output ground point, and then λ_{min} can be used to interpolate the intersection set $(x^i(\lambda_{min}), y^i(\lambda_{min}))$, of which the mean $(\bar{x}(\lambda_{min}), \bar{y}(\lambda_{min}))$ constitutes the horizontal component of the output ground point.

3.3. **Empirical Equivalence with MIG.** For every MIG calculation that went into Figures 1–3, an Hourglass geolocation calculation was also performed. Figure 5 shows the actual errors in the X, Y, and Z directions from a MIG calculation (horizontal axis) versus an Hourglass calculation (vertical axis), for all of those calculations. The scatter follows the unity line $y = x$ very closely, showing that Hourglassing yields very closely the same geolocation as MIG.

3.4. **Non-uniqueness.** It is desirable that this quartic polynomial $d(\lambda)$ have no local minima, but only a single, global minimum. Equivalently, the cubic derivative should have a single real root and two complex roots. Unfortunately, there can be degenerate arrangements of image rays with three real roots of $d'(\lambda)$ and multiple minima for $d(\lambda)$.

At least we can compute the exact form of $d(\lambda)$, and with standard techniques understand clearly whether such a degenerate situation were ever to present itself. In our empirical testing in section 3.6, we very seldom encountered such a degenerate case. In the 27600 Hourglassing computations, 26 degenerate cases were encountered: 16 for $n = 4$, 7 for $n = 5$, and 1 each for $n = 6, 7, 8$. In each case, the outer two extrema are local minima, and the middle is a local maximum, and the ambiguity of the degeneracy was computed as the difference between the local minimum values of $d(\lambda)$. The ambiguities ranged from 0.15 to 65. In any case, these are very small numbers of images, within the scope of this study. Numbers as small as $n = 4$ were included only to provide context and continuity to the results for large numbers of images. As demonstrated by this study, the chance of degeneracy for a large number of images is negligible, and in any case it can be detected for manual review.

3.5. **Hourglass Error Estimation: Self-Projected Covariance.** Review of Figures 3–4 suggests an empirical method of error estimation for Hourglassing. Predicted error decreases very predictably as $1/\sqrt{n}$. So an alternative method for estimating the error of a MIG with n images would be to subsample $m \ll n$ images many times, compute the sample covariance C_m of the m-MIGs, and because of the $1/n$ behavior proven in section 2.2, and using a finite population correction, predict the error of the n-MIG to be

$$C_n = C_m \frac{m}{n} \cdot \frac{n-1}{n-m} \tag{18}$$

This allows estimation of error apart from the output covariance of the n-MIG (or even of any of the m-MIG). There's not much point to this for MIG, since theoretically correct output covariance is a natural by-product of the MIG algorithm. But for Hourglassing, this provides a method of error estimation which requires no apriori error information, but only calculation of sample covariance of additional sub-sampled Hourglass calculations. We call this error estimation techniqe 'Self-Projected Covariance'.

The fundamental principle behind this error estimation technique, is that apriori error information is not needed for each image in the bundle, because the bundle *is* the distribution of error. For typical collections of only 2 or 4 images, the sample size is too small to reliably know whether the ray separation is a true representation of the amount of error in the system, or whether it may be concidentally large (or small). Thus a traditional MIG with apriori and aposteriori covariance, and reference variance as a consistency check, is most appropriate. But for large enough collections of images, the distribution of image rays is a reliable summary of the error of the process.

3.6. **Self-Projected Covariance Simulated Results.** To evaluate this empirical method of error estimation for Hourglassing, the following experiment was conducted: for the basic iteration, a subset of n images from $\mathcal{P}$ were chosen, and an Hourglass-geolocation computed for them. Subsets of size $m = n/4$ were sampled from the n-set, $k = 100$ times, and an Hourglass-geolocation is computed for each m-subset. 3x3 sample covariance was computed for those k ground points, and then projected into an error estimate for the n images using (18). For reference, the n-subset was also input to MIG, and the m-sets were also input to MIG to provide another self-projected covariance estimate. This basic process was repeated $r = 100$ times each, for $n = 100, 200$ and 400.

Figure 6 compares self-projected covariance from MIG and Hourglass points on the m-subsamples. The left graph shows x variance, and the right shows z variance (y is omitted from the horizontal graph because it is virtually identical to x and just makes the graph harder to read). Self-projected variance from MIG and Hourglass points are remarkably similar, with Hourglass self-projected variance trending a litle higher (above the identity line $y = x$). MIG variances from 2.2 are much less variable; the vertical lines show the magnitude of MIG variance for the n-sets. Although self-projected covariance has more variability, it is clustered around the correct value.

4. Conclusion

We have shown with both theory and simulation with real/augmented data, that least-squares Multi-Image Geopositioning (MIG) yields output ground covariance which diminishes as $1/n$ (and linear error

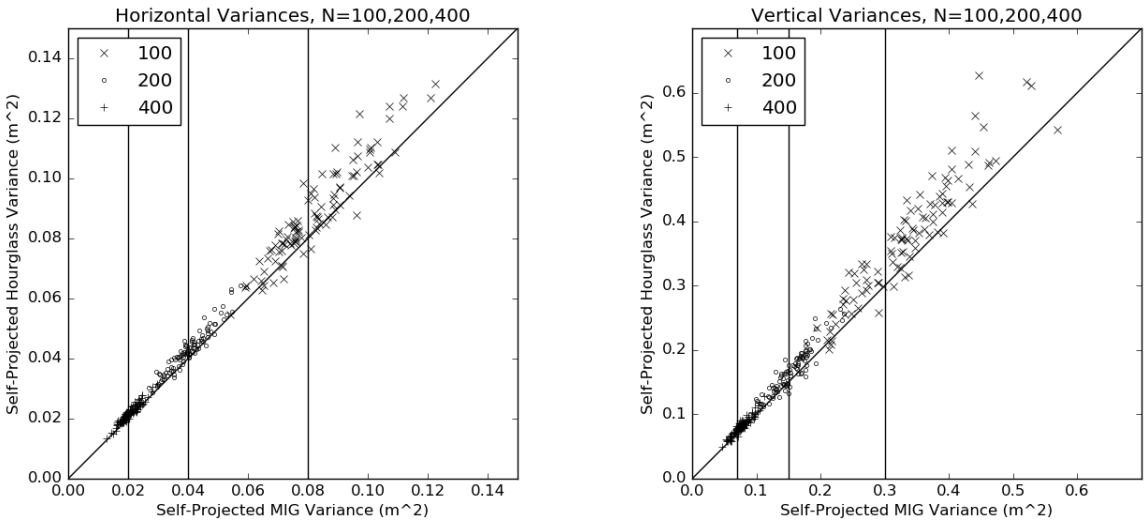

FIGURE 6. MIG vs Hourglass estimated variances. From left to right, the clusters represent $n = 400, 200$ and 100 images. X variance only is presented on the left. Y variance is omitted, because it is so similar to X, it would just make the graph more difficult to read. The vertical lines represent the magnitude of variance from a traditional (non-self-projected) MIG

estimates such as CE90/LE90 diminish as $1/\sqrt{n}$), for increasing numbers of images n. An empirical geolocation technique called Hourglassing is introduced, which needs no apriori sensor covariance, and which is shown to behave comparably to MIG, including an error estimation procedure that captures uncertainty from the distribution of image rays itself.

Using techniques such as these, ever-growing stores of satellite imagery can be exploited to yield increasingly high accuracy geolocation.

REFERENCES

[1] Casella, George, Roger L. Berger, 1990. *Statistical Inference*, Brooks/Cole, Belmont CA, pp. 214-20.
[2] Collins, R. T., 1996. A space-sweep approach to true multi-image matching, In: *Proc. CVPR*, 1996.
[3] Di, K., B. Xu, B. Liu, M. Jia, Z. Liu, 2016. Geopositioning precision analysis of multiple image triangulation using LRO NAC Lunar Images, In: *Proc. ISPRS*, 2016.
[4] Digital Globe Image Finder, `http://browse.digitalglobe.com`, accessed 2017-02-24.
[5] Dolloff, John, and Reuben Settergren, 2010. Worldview-1 Stereo Extraction Accuracy With and Without MIN Processing, In: *Proc. ASPRS*, 2010.
[6] Förstner, Wolfgang, and Bernhard P. Wrobel, 2016. *Photogrammetric Computer Vision*, Springer.
[7] `https://github.com/sethmerickel/Multi-Image-Geopositioning`. The implementation of Hourglassing described in section 3.2 is available, as well as sample data.
[8] Horn, Roger, and Charles R. Johnson, 2012. *Matrix Analysis*, 2nd ed. Cambridge University Press, p. 239.
[9] Jeong, Jaehoon, Chansu Yang, Taejung Kim, 2015. Geo-positioning accuracy using multiple-satellite images: IKONOS, QuickBird, and KOMPSAT-2 Stereo Images, *Remote Sensing* **7**(4), pp. 4549-64.
[10] McGlone, J. Chris (Ed.), 2013. *Manual of Photogrammetry*, 6th ed.
[11] Mikhail, Edward M., James S. Bethel, J. Chris McGlone, 2001. *Modern Photogrammetry*, Wiley, New York.
[12] SOCET GXP, `http://geospatialexploitationproducts.com`
[13] Settergren, R., S. Merickel and S. Walker, 2017. Reducing geopositioning error with large numbers of images, 2017 Geomatics Engineering Conference, Fresno, CA, 27 Jan 2017. 79 slides, available online [7]
[14] Szeliski, Richard, 2011. *Computer Vision*, Springer, London, p. 474.
[15] Thompson, Steven K., 2012. *Sampling*, 3rd ed. Wiley, Hoboken, p. 15.
[16] Wonnacott, W. M., 2008. *Geolocation with Error Analysis Using Imagery from an Experimental Spotlight SAR*, PhD Thesis, Purdue.