Cohen, E.A.K.; Ober, R.J., "Image registration error analysis with applications in single molecule microscopy," in *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on* , vol., no., pp.996-999, 2-5 May 2012
doi: 10.1109/ISBI.2012.6235725
keywords: {bio-optics;biomolecular effects of radiation;fluorescence;image registration;least squares approximations;measurement errors;medical image processing;regression analysis;errors-in-variable problem;generalized least squares method;image registration error analysis;localization error;measurement error;monochromatic fluorescence microscopy;multivariate generalized least squares estimator;multivariate linear regression problem;photon count;single molecule microscopy;Covariance matrix;Image registration;Measurement errors;Microscopy;Noise;Photonics;Vectors;Image registration;Microscopy;Total least squares methods},
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6235725&isnumber=6235468

IMAGE REGISTRATION ERROR ANALYSIS WITH APPLICATIONS IN SINGLE MOLECULE MICROSCOPY

E. A. K. Cohen and R. J. Ober

Eric Jonsson School of Electrical Engineering and Computer Science,
University of Texas at Dallas, Richardson, TX 75083-0688 USA.

Department of Immunology,
University of Texas Southwestern Medical Center, Dallas, TX 75235-8576 USA.

ABSTRACT

This paper is concerned with assessing localization errors emanating from the image registration of two monochromatic fluorescence microscopy images. Assuming an affine transform exists between images, registration in this setting typically involves using control points to solve a multivariate linear regression problem; however with measurement errors existing in both sets of variables the use of linear least squares is inappropriate. It is shown that image registration is an errors-in-variable problem and as such the correct method is to use generalized least squares. Traditionally this requires the measurement errors to be independent and identically distributed (iid); an assumption that is rarely satisfied in practical situations. An extension of the multivariate generalized least squares estimator that allows non-iid noise is applied. The distributional properties of the estimators are used to derive localization errors emanating from the image registration process in terms of photon counts and experimental parameters.

***Index Terms*—** Image registration, Microscopy, Total least squares methods

1. INTRODUCTION

In its most general form, image registration is the process of transforming different data sets into a single coordinate system. It is of use in areas as diverse as computer vision, target tracking, medical imaging and in fluorescence microscopy, the setting for this paper.

Recent advances in fluorescence microscopy have made it possible to detect single molecules in a cellular environment, e.g. [1]. A typical goal of a fluorescence microscopy experimental set-up is to observe the relative positions of a collection of two different protein molecules, for example to see whether they colocalize e.g. [2]. Two different fluorescents that emit at different wavelengths are used to tag the proteins, one for each type of molecule. The two types of molecule are imaged separately, but simultaneously using a pair of filters suitable for the fluorescents. A pair of monochromatic images are formed. It is often the case that the pair of images can not be considered to use the same coordinate system and the goal is to register the images so that the positions of the molecules can be viewed in unison, e.g. [3]. To do so we consider an image to be a subset of the space $\mathbb{R}^2$. Suppose we have two images, $\mathcal{I}_1$ and $\mathcal{I}_2$, say, and there is some mapping $T : \mathcal{I}_1 \to \mathcal{I}_2$ between the pair of images; the type of mapping assumed here is an *affine transformation*.

An affine transformation between two vector spaces is a linear transformation followed by a translation. When both vector spaces are $\mathbb{R}^d$ then $T : \mathbb{R}^d \to \mathbb{R}^d$ is affine if there exists a matrix parameter $A \in \mathbb{R}^{d\times d}$ (the linear transformation), and a vector parameter $t \in \mathbb{R}^d$ (the translation) such that for $x \in \mathbb{R}^d$, $T : x \to Ax + t$. In this situation we are concerned with the case $d = 2$.

In practice the affine transformation that exists between a pair of images is not known and needs to be estimated experimentally. To do this it is necessary to have points of reference (control points) visible in both images. In fluorescence microscopy this is achieved through fiduciary markers such as small but bright fluorescent beads [2]. These objects emit photons across the visible spectrum resulting in them being observed through both filters and appearing in both images.

Suppose K control points are located in $\mathcal{I}_1$ at true locations $\{x_k^{(1)}, k = 1, ..., K\}$, and in $\mathcal{I}_2$ at true locations $\{x_k^{(2)}, k = 1, ..., K\}$, such that $x_k^{(2)} = T(x_k^{(1)}) = Ax_k^{(1)} + t$. In reality the position of these beads cannot be known exactly and must instead be measured. Consequently the true coordinates are not directly observed, but instead we observe coordinates that have been perturbed by some random additive noise. That is to say we observe $\{y_k^{(1)}, k = 1, ..., K\}$ and $\{y_k^{(2)}, k = 1, ..., K\}$ where $y_k^{(j)} = x_k^{(j)} + \epsilon_k^{(j)}$, $k = 1, ..., K$, $j = 1, 2$, where $\epsilon_k^{(j)} \in \mathbb{R}^2$ is a random vector known as the measurement error with zero mean and covariance $\Sigma_k^{(j)}$.

Supported in part by the National Institute of Health grant R01 GM085575.

Through the work of Ober and Ram [4] [5], in practice it is possible to determine the error covariance matrices $\{\Sigma_k^{(1)}, k = 1, ..., K\}$ and $\{\Sigma_k^{(2)}, k = 1, ..., K\}$, with dependency on the microscopy set up and the intensity of the objects being imaged. It is assumed throughout this paper that the localization errors are normally distributed with mean zero [6] [7].

Given that we are able to construct estimators $\hat{A}$ and $\hat{t}$ of the transformation parameters A and t, registration of the two images can commence. In the molecular fluorescence microscopy setting we observe a molecule in $\mathcal{I}_1$ at point $y_a^{(1)}$, where $y_a^{(1)} = x_a^{(1)} + \epsilon_a^{(1)}$ (again the covariance of $\epsilon_a^{(1)}$ is estimable from [4]). $x_a^{(1)}$ is the true positional vector of the molecule. We estimate its new coordinate (the registered position) in $\mathcal{I}_2$ as $\hat{A}y_a^{(1)} + \hat{t}$. A key question is; what is the error associated with the registered position in $\mathcal{I}_2$? Or equivalently to determine the covariance of the image registration error (IRE)

$$\nu_a \equiv x_a^{(2)} - \hat{A}y_a^{(1)} - \hat{t},$$

where $x_a^{(2)} = Ax_a^{(1)} + t$.

In Section 2 we formulate the image registration problem and show that parameter estimation is consistent with a multivariate errors-in-variable framework. This is followed by a discussion on the image registration error. In Section 3 is a discussion on the distributional properties of the measurement errors. Traditional linear least squares methods are biased for errors-in-variable models so it is necessary to use the generalized least squares method to estimate the affine transform parameters. These estimators and their associated distributional properties are presented in Section 4 and are used to determine the image registration error in Section 5.

2. THE IMAGE REGISTRATION PROBLEM

We define the following $\mathbb{R}^{2\times K}$ matrices for $j = 1, 2$

$$\begin{aligned} X^{(j)} &\equiv \left[x_1^{(j)}, ..., x_K^{(j)}\right] \\ Y^{(j)} &\equiv \left[y_1^{(j)}, ..., y_K^{(j)}\right] \\ E^{(j)} &\equiv \left[\epsilon_1^{(j)}, ..., \epsilon_K^{(j)}\right], \end{aligned}$$

and further define the $\mathbb{R}^{4\times K}$ matrices

$$X = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix}, \quad Y = \begin{bmatrix} Y^{(1)} \\ Y^{(2)} \end{bmatrix}, \quad E = \begin{bmatrix} E^{(1)} \\ E^{(2)} \end{bmatrix}.$$

With this notation the system of equations can be conveniently represented as the single matrix equation

$$Y = \alpha \mathbf{1}_K^T + \Lambda X^{(1)} + E, \tag{1}$$

where $\alpha = [\mathbf{0}^T, t^T]^T$, $\Lambda = [I_2, A^T]^T$ and $\mathbf{1}_K$ is a column vector of length K with every element taking the value 1. The columns of $X^{(1)}$ are known as the *independent* variables and the columns of $X^{(2)}$ are the *dependent* variables. Linear regression problems where both the independent and dependent variables contain measurement errors are known as *errors-in-variable* (EIV) models. Let E_k denote the kth column of E, then E_k is 4-D multivariate normally distributed with mean zero and covariance

$$\Sigma_k = \begin{bmatrix} \Sigma_k^{(1)} & 0 \\ 0 & \Sigma_k^{(2)} \end{bmatrix},$$

for which we write $E_k \overset{\mathrm{d}}{=} N_2(\mathbf{0}, \Sigma_k)$.

2.1. Image registration error

Suppose we image a molecule in $\mathcal{I}_1$ with true position $x_a^{(1)} \in \mathcal{I}_1$. The true position of the molecule in $\mathcal{I}_2$, $x_a^{(2)} \in \mathcal{I}_2$ will be related to $x_a^{(1)}$ by the same affine transform, namely $x_a^{(2)} = Ax_a^{(1)} + t$. Let $y_a^{(1)} = x_a^{(1)} + \epsilon_a^{(1)}$ be the measured position of the molecule in $\mathcal{I}_1$, where $\epsilon_a^{(1)} \overset{\mathrm{d}}{=} N_2(\mathbf{0}, \Sigma_a^{(1)})$.

Definition 2.1. *The image registration error (IRE) $\nu_a = [\nu_{a,1}, \nu_{a,2}]^T$ associated with registering the molecule in $\mathcal{I}_2$ is defined as*

$$\nu_a \equiv x_a^{(2)} - \hat{A}y_a^{(1)} - \hat{t}, \tag{2}$$

where $\hat{A}$ and $\hat{t}$ are estimators of the matrix A and vector t, respectively.

Let us define the difference between the true and estimated values of the transform parameters as $\Delta A \equiv \hat{A} - A$ and $\Delta t \equiv \hat{t} - t$. It can be shown that

$$\nu_a = -A\epsilon_a^{(1)} - \Delta t - \Delta A x_a^{(1)} - \Delta A \epsilon_a^{(1)},$$

and hence

$$E\{\nu_a\} = -E\{\Delta t\} - E\{\Delta A\}x_a^{(1)},$$

where $E\{\cdot\}$ is the expectation operator. Let $\mathrm{cov}\{v\}$ denote the covariance matrix of a vector v.

Lemma 2.2. *If $\hat{A}$ and $\hat{t}$ are unbiased estimators and $\epsilon_a^{(1)}$ is independent of both ΔA and Δt, then $E\{\nu_a\} = \mathbf{0}$ and*

$$\begin{aligned} \mathrm{cov}\{\nu_a\} = A\Sigma_a^{(1)}A^T + E\{(\Delta t + \Delta A x_a^{(1)})(\Delta t + \Delta A x_a^{(1)})^T\} \\ + E\{\Delta A \epsilon^{(1)} \epsilon^{(1)T} \Delta A^T\}. \end{aligned} \tag{3}$$

To derive an expression for the second order moments of the IRE it is necessary to understand the distributional properties of the estimators ΔA and Δt.

3. MEASUREMENT ERRORS

In [4] detailed analytical expressions for the localization accuracy of a light emitting point source in a microscopy imaging experiment are derived allowing reliable estimates of the covariance matrices for the measurement errors. The Cramér-Rao lower bound (CRLB) theorem states:

Theorem 3.1. *The covariance matrix of any unbiased estimator $\hat{\theta}$ of an unknown vector parameter θ is bounded from below by the inverse of the Fisher information matrix $I(\theta)$, i.e.* $\mathrm{cov}(\hat{\theta}) \geq I^{-1}(\theta)$.

For the purposes of this paper we assume that the CRLB is attained. We consider the most general and suitable case for application purposes, whereby the object being imaged emits photons as a Poisson process, and there is an additional background Poisson noise process and Gaussian read out noise. Typically the beads have large photon emission rates in comparison to the background and readout noise processes. In such circumstances the Fisher information matrix for the estimator of the location parameter $\theta = (u, v)$ of an object in $\mathcal{I}_j$ can be considered to be of the form $I(\theta) = NZ^{(j)}$ where $N \in \mathbb{Z}^+$ is the photon count associated with the imaged object and $Z^{(j)}$ is a $\mathbb{R}^{2\times 2}$ symmetric positive definite (SPD) matrix of known form.

The photon count associated with the kth bead in $\mathcal{I}_j$ is labeled $N_k^{(j)}$. The noise terms $\epsilon_k^{(j)}$, $k = 1, ..., K$, $j = 1, 2$, have covariance $\Sigma_k^{(j)} = (1/N_k^{(j)})\tilde{\Sigma}_0^{(j)}$ where $\tilde{\Sigma}_0^{(j)} = Z^{(j)-1}$ is a SPD matrix and universal for all imaged points in $\mathcal{I}_j$. We now make the assumption that there is a linear relation between the brightness of the bead in each image, i.e. a bead that is bright in $\mathcal{I}_1$ is also bright in $\mathcal{I}_2$. Mathematically we say $N_k^{(2)} = cN_k^{(1)}$ for all $k = 1, ..., K$, where $c > 0$ is a constant of proportionality, universal for all control points. We define $\bar{N}^{(j)} \equiv (1/K)\sum_{k=1}^K N_k^{(j)}$, and $\Sigma_0^{(j)} \equiv (1/\bar{N}^{(j)})\tilde{\Sigma}_0^{(j)}$, for $j = 1, 2$, then $\bar{N}^{(2)} = c\bar{N}^{(1)}$ and we have the situation where the covariance of the measurement errors are scalar multiples of a common matrix, i.e.

$$\Sigma_k = \eta_k \begin{bmatrix} \Sigma_0^{(1)} & 0 \\ 0 & \Sigma_0^{(2)} \end{bmatrix}$$

where $\eta_k = \bar{N}^{(1)}/N_k^{(1)}$ and $\Sigma_0^{(1)}$ and $\Sigma_0^{(2)}$ are known.

4. GENERALIZED LEAST SQUARES WITH COVARIANCE WEIGHTING

Given the columns of E are iid with common covariance Σ_0 say, EIV models of type (1) are traditionally solved by minimizing the normalized residual sum of squares

$$r = \sum_{k=1}^K v_k^T \Phi^{-1} v_k \tag{4}$$

where $v_k = y_k^{(2)} - t - Ay_k^{(1)}$, $\Phi = U\Sigma_0 U^T$ and $U = [-A, I_2]$. In the statistics literature, the values of A and t that minimize r are known as the *generalized least squares* (GLS) estimators [8]. In the engineering and numerical analysis literature *total least squares* (TLS) is the common approach where a different minimization problem is solved. Both estimators are identical to one another. Here we generalize (4) for the case where the columns of E no longer have common covariance.

Definition 4.1. *The generalized least squares estimators $\hat{A}$ and $\hat{t}$ are the values of A and t, respectively, that minimize the weighted residual sum of squares*

$$r = \sum_{k=1}^K v_k^T \Phi_k^{-1} v_k, \tag{5}$$

where $\Phi_k = U\Sigma_k U^T$.

Consider the multivariate EIV model (1), where the covariance of E_k is given by the SPD matrix $\Sigma_k = \eta_k \Sigma_0$ where $(1/K)\sum_{k=1}^K \eta_k^{-1} = 1$. For the estimators that minimize (5) and their asymptotic statistical properties we state the following two theorems. They are given full treatment in [9].

Theorem 4.2. *Define the matrix $S \equiv (1/K)\sum_{k=1}^K \eta_k^{-1}(y_k - \bar{y})(y_k - \bar{y})^T$ where $\bar{y} \equiv (1/K)\sum_{k=1}^K \eta_k^{-1} y_k$. By considering $V = GDG^T$, the eigenvalue decomposition of matrix $V \equiv S\Sigma_0^{-1}$, with $D = \mathrm{diag}\{d_1, d_2, d_3, d_4\}$ where $d_1 \geq ... \geq d_4$ and $GG^T = G^T G = I_4$, then by making the partition*

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{11} \end{bmatrix},$$

the GLS estimator of A and t are given as

$$\hat{A} = G_{21}G_{11}^{-1}, \quad \hat{t} = \sum_{k=1}^K \eta_k^{-1} U y_k.$$

Theorem 4.3. *Define $\mathbb{R}^2$ vector $\bar{x} \equiv (1/K)\sum_{k=1}^K \eta_k^{-1} x_k^{(1)}$, $\mathbb{R}^{2\times 2}$ matrix $\Xi \equiv (1/K)\sum_{k=1}^K \eta_k^{-1} x_k^{(1)} x_k^{(1)T}$, matrix $\Psi = \Xi - \bar{x}\bar{x}^T$ and matrix $\Theta = \Psi^{-1}(\Lambda^T \Sigma_0^{-1} \Lambda)^{-1}\Psi^{-1} + \Psi^{-1}$, then we have the following identities*

$$\mathrm{cov}\{\Delta t_m, \Delta t_n\} = K^{-1}\left(1 + \bar{x}\Theta\bar{x}^T\right)\Phi_{mn} \tag{6}$$

$$\mathrm{cov}\{\Delta t_l, \Delta a_{mn}\} = K^{-1}[\Theta\bar{x}]_n \Phi_{ml} \tag{7}$$

$$\mathrm{cov}\{\Delta a_{mn}, \Delta a_{m'n'}\} = K^{-1}\Theta_{mm'}\Phi_{nn'}. \tag{8}$$

5. IMAGE REGISTRATION ERROR ANALYSIS

We now consider the asymptotic form of the IRE covariance matrix. Consider a molecule in $\mathcal{I}_1$ at point $y_a^{(1)}$, where $y_a^{(1)} = x_a^{(1)} + \epsilon_a^{(1)}$ with the measurement error having covariance Σ_a estimable from [4], and $x_a^{(1)}$ being the true positional vector. We estimate its new coordinate (the registered position) in $\mathcal{I}_2$ as $\hat{A}y_a^{(1)} + \hat{t}$. The IRE is defined in (2) and its covariance matrix given in (3). The individual elements of $\mathrm{cov}\{\nu_a\}$ can now be calculated using (6), (7) and (8).

5.1. Approximating the image registration error

Let us consider some simplifying assumptions. It is common that the control points (beads) are randomly scattered throughout the image and as such we will model their location as a 2-D Gaussian distribution.

Lemma 5.1. *Let control point positions* $\{x_k^{(1)}, k = 1, ..., K\}$ *be* K *realizations of a random variable* $\mathcal{X} = [\mathcal{X}_1, \mathcal{X}_2]^T \in \mathbb{R}^2$ *where*

$$\mathcal{X} \stackrel{\mathrm{d}}{=} N_2(\mu, \Pi),$$

then asymptotically $\bar{x} = \mu$ *and* $\Psi = \Pi$.

5.2. Assumptions

1. Consider the positions $\{x_k^{(1)}, k = 1, ..., K\}$ to be K realizations of a random variable $\mathcal{X} = [\mathcal{X}_1, \mathcal{X}_2]^T \in \mathbb{R}^2$ where
$$\mathcal{X} \stackrel{\mathrm{d}}{=} N_2(\mathbf{0}, \kappa^2 I_2).$$
From Lemma 5.1 we have $\bar{x} = 0$ and $\Psi = \kappa^2 I_2$.

2. The affine transformation parameter A represents a rotation R combined with a magnification M. Consequently A would be of the form MR where $RR^T = R^T R = I_2$ and $M = mI_2$.

3. We model the measurement errors as $E_k \stackrel{\mathrm{d}}{=} N_4(\mathbf{0}, \Sigma_k)$ where
$$\Sigma_k = \begin{pmatrix} \sigma_{1,k}^2 I_2 & 0 \\ 0 & \sigma_{2,k}^2 I_2 \end{pmatrix} = \frac{\bar{N}^{(1)}}{N_k^{(1)}} \begin{pmatrix} \sigma_1^2 I_2 & 0 \\ 0 & \sigma_2^2 I_2 \end{pmatrix}$$
with the normalized variance terms σ_j^2 inversely proportional to the mean photon count, $\sigma_j^2 = \zeta^{(j)}/\bar{N}^{(j)}$, $j = 1, 2$. The terms $\zeta^{(j)}$ are known functions of the experimental parameters, including the Airy profile and photon wavelength [4].

Assumptions 1 – 3 give the identities $x = 0$, $\Phi = (m^2\sigma_1^2 + \sigma_2^2)I_2$ and $\Theta = \kappa^{-4}(\sigma_1^2 + m^2\sigma_2^2)I_2 + \kappa^{-2}I_2$. We immediately conclude $\mathrm{cov}\{\Delta t_l, \Delta a_{ij}\} = 0$ for $l, i, j = 1, 2$. Further to this $\mathrm{cov}\{\Delta t_i, \Delta t_j\} = K^{-1}(m^2\sigma_1^2 + \sigma_2^2)\delta_{ij}$, where δ_{ij} is the Kronecker delta. $\mathrm{cov}\{\Delta a_{ij}, \Delta a_{i'j'}\} = 0$ if $i \neq i'$ and/or $j \neq j'$. If $i = i'$ and $j = j'$, then given $\kappa^2 \gg (\sigma_1^2 + m^2\sigma_2^2)$ we can treat $\mathrm{cov}\{\Delta a_{ij}, \Delta a_{i'j'}\} \approx 0$. This is equivalent to stating that the spread of the control points is much greater than the individual measurement errors, a valid assumption in a microscopy setting.

Theorem 5.2. *The covariance terms of the IRE can be approximated as*

$$\mathrm{cov}\{\nu_{a,i}, \nu_{a,j}\} \approx \left[A\Sigma_a A^T\right]_{ij} + K^{-1}\left(\sigma_1^2 m^2 \delta_{ij} + \sigma_2^2\right).$$

If the covariance Σ_a *is itself representable as* $\sigma_a^2 I_2$ *then in terms of photon count the covariance terms become*

$$\mathrm{cov}\{\nu_{a,i}, \nu_{a,j}\} \approx m^2\sigma_a^2\delta_{ij} + K^{-1}\left(m^2\frac{\zeta^{(1)}}{\bar{N}^{(1)}}\delta_{ij} + \frac{\zeta^{(2)}}{\bar{N}^{(2)}}\right).$$

6. CONCLUSION

When the control points have measurement error covariances that are known and some scalar multiple of a common SPD matrix, then the second order moments of the IRE have been derived. For microscopy applications we have shown there exists inverse dependencies on the mean photon counts of the control points, however there remains a theoretical lower bound to the variance or the IRE in either dimension of $m^2\sigma_a^2$. Examples and simulations are found in [9].

7. REFERENCES

[1] D. W. Pierce, N. Hom-Booher, and R. D. Vale, "Imaging individual green fluorescent proteins," *Nature*, vol. 388, pp. 338, 1997.

[2] E. Lachmanovich, D. E. Shvartsman, Y. Malka, C. Botvin, Y. I. Henis, and A. M. Weiss., "Co-localization analysis of complex formation among membrane proteins by computerized fluorescence microscopy: application to immunofluorescence co-patching studies," *Journal of Microscopy*, vol. 212, no. 2, pp. 122 — 131, 2003.

[3] C. P. Allier, G. Hiernard, V. Poher, and J. M. Dinten, "Bacteria detection with thin wetting film lensless imaging," *Biomedical Optics Express*, vol. 1, no. 2, pp. 762 — 770, 2010.

[4] R. J. Ober, S. Ram, and E. S. Ward, "Localization accuracy in single-molecule microscopy," *Biophysical Journal*, vol. 86, pp. 1185 — 1200, 2004.

[5] S. Ram, E. S. Ward, and R. J. Ober, "Localization accuracy in single-molecule microscopy," *Proceedings National Academy of Science USA*, vol. 103, pp. 4457 — 4462, 2006.

[6] N. Bobroff, "Position measurement with a resolution and noise limited instrument," *Review Scientific Instruments*, vol. 57, pp. 1152 — 1157, 1986.

[7] M. Goulian and S. M. Simon, "Tracking single proteins within cells," *Biophysical Journal*, vol. 66, pp. 2188 — 2198, 2000.

[8] L. J. Gleser, "Estimation in a multivariate "errors in variable" regression model: large sample results," *The Annals of Statistics*, vol. 9, no. 1, pp. 24 — 44, 1981.

[9] E. A. K. Cohen and R. J. Ober, "Generalized least squares with weighted covariance for image registration error analysis in single molecule microscopy," *To be submitted*, 2011.