



Published in final edited form as:

IEEE Trans Med Imaging. 2011 March ; 30(3): 679–693. doi:10.1109/TMI.2010.2091513.

General Approach to First-Order Error Prediction in Rigid Point Registration

Andrei Danilchenko and

Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37235 USA

J. Michael Fitzpatrick [Fellow, IEEE]

Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37235 USA

Andrei Danilchenko: andrei.v.danilchenko@vanderbilt.edu; J. Michael Fitzpatrick: j.michael.fitzpatrick@vanderbilt.edu

Abstract

A general approach to the first-order analysis of error in rigid point registration is presented that accommodates fiducial localization error (FLE) that may be *inhomogeneous* (varying from point to point) and *anisotropic* (varying with direction) and also accommodates arbitrary weighting that may also be inhomogeneous and anisotropic. Covariances are derived for target registration error (TRE) and for weighted fiducial registration error (FRE) in terms of covariances of FLE, culminating in a simple implementation that encompasses all combinations of weightings and anisotropy. Furthermore, it is shown that for ideal weighting, in which the weighting matrix for each fiducial equals the inverse of the square root of the cross covariance of its two-space FLE, fluctuations of FRE and TRE are mutually independent. These results are validated by comparison with previously published expressions and by simulation. Furthermore, simulations for randomly generated fiducial positions and FLEs are presented that show that correlation is negligible correlation coefficient < 0.1 in the exact case for both ideal and uniform weighting (i.e., no weighting), the latter of which is employed in commercial surgical guidance systems. From these results we conclude that for these weighting schemes, while valid expressions exist relating the covariance of FRE to the covariance of TRE, there are no measures of the goodness of fit of the fiducials for a given registration that give to first order any information about the fluctuation of TRE from its expected value and none that give useful information in the exact case. Therefore, as estimators of registration accuracy, such measures should be approached with extreme caution both by the purveyors of guidance systems and by the practitioners who use them.

Index Terms

Accuracy; anisotropic; error analysis; fiducial; registration; surgical guidance; weighting

I. Introduction

In 2009, a new method of first-order analysis of the point-registration problem was introduced [2]. The method was developed to facilitate the statistical analysis of a phenomenon that had been observed years earlier, first by Steinmeier using phantoms [3] and more recently by Woerdeman *in vivo* [4]. These investigators attempted to establish for several fiducial-based, surgical navigation systems, a correlation between the displayed estimates of accuracy and the true accuracy. Steinmeier found no correlation, and Woerdeman observed a negligible correlation coefficient of only 0.08. Thus, lower than normal estimated accuracies were not accompanied by lower than normal true accuracies, nor were there correlations for higher than normal values. The navigational systems' estimates of accuracy are based on the goodness of fit of the fiducial points after registration, and the surprising result that these estimates seem to be unrelated to truth, suggests that there is an underlying problem with this method of error prediction. The 2009 work [2] investigated this problem for a special case in which expected error in fiducial localization is isotropic and identical for all fiducials, and it proved to first order that fluctuations in registration accuracy are statistically independent of the fluctuations of any estimate of that accuracy based on the goodness of fit. In the present work, we extend this analysis to the general case. We show that, when ideal weighting is used for each fiducial, this independence is maintained, and we show by simulations that, without weighting, the correlation between them is on the order of that low value observed by Woerdeman. In addition, we provide new expressions for error statistics that are more general than those previously provided. We begin in Section II by providing background on, and precise definitions of, error in these navigation systems. In Section III we present our new approach and derive independence and the new statistical expressions. In Sections IV and V we compare our results to previously published results and to results from computer simulations. In Sections VI and VII we discuss our theory and its results and present our conclusions.

II. Background

Fiducial-based navigation has become a ubiquitous adjunct to surgery when rigid anatomy—bone or soft tissue constrained by bone—is involved. The major commercial guidance systems—the StealthStation Surgical Navigation System (Medtronic, Inc., Minneapolis, MN) and the VectorVision System (BrainLAB AG, Feldkirchen, Germany), each provide it as an option in addition to surface-based navigation, and both options incorporate rigid point registration in their algorithms to determine the position of the surgeon's tracked probe or tracked instrument. The registration approach analyzed in the following sections is employed by both of these systems. Fiducials, often called "markers," are provided that typically adhere to the skin, but the selection of anatomical "fiducials" are supported by these systems as well. Both fiducial and surface registration approaches are often called "frameless" stereotaxy, in reference to the cumbersome stereotactic frame that they are designed to replace. These frameless approaches are offered by virtually every surgical center as an adjunct to neurosurgery, spine surgery, and orthopedic surgery, and they are provided as well by radiation oncology centers as an adjunct to radiosurgery.

Because fiducials by their nature are discrete objects, they provide positional information only at isolated regions of the anatomy. Information about the position or movement of anatomy outside these regions can be inferred only if all these regions move as a single rigid object. As a result, fiducial-based navigation systems employ only rigid transformations during the registration process. To the extent that the anatomy is non-rigid, rigid transformations are inappropriate, but for surgical resections involving bone, such as a vertebra or a femur, or tissue and bone that move as a single rigid object, such as the head, rigidity is a reasonable assumption. Failure of the anatomy to remain rigid will cause systems' predictions of fiducial-registration accuracy, such as those studied by Steinmeier and Woerdeman, to fail, but nonrigidity is not the source of failure that we present here. Indeed Steinmeier's experiments involved only rigid phantoms. Instead we confine our attention to those applications in which rigidity is a valid assumption.

For rigid anatomy, the image-guidance problem is reduced to the determination of the rigid transformation that the anatomy has undergone between image acquisition and intervention. That transformation can be estimated via fiducials that are attached to the anatomy before imaging and remain in place through the registration procedure. The position of each fiducial is localized in both image space and physical space, and then a transformation from image space to physical space is usually chosen such that it minimizes the fiducial registration error (FRE), where, in the simplest method, which is employed by commercial systems, FRE is the root-mean-square of distances between corresponding fiducials after the registration, i.e., the distance between the localized position of each fiducial as transformed from image space to physical space and the position of that corresponding fiducial localized in physical space. In the general case, the displacement between the transformed image point and the physical point is multiplied by a "weighting" matrix, which may be different for each fiducial, and FRE is the root-mean-square of the lengths of the resulting vectors. The weighting matrix is used in cases when some information about the nature of fiducials localization procedure is available. The simplest method described above is equivalent to uniform weighting, and it is employed by commercial systems in the absence of such information. Because of fiducial localization error (FLE), the registration is invariably flawed to some degree, and FRE is a readily available measure of registration error that serves as the most commonly used indicator of the goodness of fit of the of the corresponding points.

Sibson in 1979 investigated the statistical relation between FRE and FLE for the case of isotropic FLE and provided an expression that can be used to relate their expected root-mean-square values as follows [5]:

$$\langle \text{FRE}^2 \rangle = \left(1 - \frac{2}{N} \right) \langle \text{FLE}^2 \rangle \quad (1)$$

where $\langle \cdot \rangle$ means "expected value of." Expressions for FRE for each individual fiducial have also been derived [10]. FRE is an easily measured quantity, but it is not as important as target registration error, **TRE**, which is the displacement from its true position of a registered point *not* used as a fiducial¹. The target error **TRE**(*r*) at a point *r* of interest to the surgeon is the true measure of registration error. In 1998 and 2001, the statistical relation

between TRE and FLE was investigated for isotropic FLE and expressions were provided relating their expected root-mean-square values [6], [7]

$$\langle |\mathbf{TRE}(r)|^2 \rangle = \frac{1}{N} \left(1 + \frac{1}{3} \sum_{k=1}^3 \frac{d_k^2}{f_k^2} \right) \langle \mathbf{FLE}^2 \rangle \quad (2)$$

where d_k is the distance of x from principal axis k of the fiducial configuration, f_k^2 is the mean of the squared distances of the fiducials from that axis. Equation (1) can be used to estimate the distribution of FLE from measurements of FRE for many cases. With this estimate available, (2) can then be exploited to predict TRE based on the fiducial configuration when fiducials are used (a) on a patient, (b) on a coordinate reference frame attached to a patient, and (c) on a surgical probe [8]. Equation (2) is not appropriate for FLE that is *anisotropic*, which means that it varies with direction, *inhomogeneous*, which means that it is different for different points. For anisotropic and inhomogeneous FLE, more complex expressions than that of (2) have recently been derived by Wiles [9], [23] and Moghari [11].

Since both $\langle \mathbf{FRE}^2 \rangle$ and $\langle |\mathbf{TRE}(r)|^2 \rangle$ are proportional to $\langle \mathbf{FLE}^2 \rangle$, it might be expected that for a given registration, if FRE is small, then $|\mathbf{TRE}(r)|$ will be small and when one is large the other will be large. While it may seem intuitive to expect the size of TRE to be correlated with that of FRE, it is easy to construct counter-examples, as is shown by Fig. 1. That figure shows two examples. For one of them FRE is zero while $|\mathbf{TRE}(r)|$ is large, and for the other $|\mathbf{TRE}(r)|$ is zero while FRE is large. The key difference is that in the first example the set of FLEs are equivalent to a rigid transformation, while in the second example, they are not. Fig. 1(a) shows image space, while (b) and (d) show physical space. The localizations are assumed to be perfect in image space. The “rigid” set of localization errors occur in (b), while the errors in (d) cannot be represented exactly or even approximately by any rigid transformation.

Fig. 1 shows schematically that, despite the fact that $\langle \mathbf{FRE}^2 \rangle$ and $\langle |\mathbf{TRE}(r)|^2 \rangle$ are related functionally through (1) and (2), TRE is not necessarily related to FRE for specific registration instances. Thus, the fact that their means are related does not imply that their variations around those means are related. It is easy to construct such examples, but no example or set of examples can answer the question, “Are the fluctuations² of FRE and TRE about their means correlated?” In the next sections we will show that, for the case of small FLE, when either of the two most common weighting schemes is employed, the answer to this question is, “No.”

We begin by proving this answer mathematically to first order in localization error for the case of so-called “ideal weighting”. That proof requires a statistical analysis of a linear model of the problem of rigid point registration. Recently a new approach [2] for such an analysis was developed for the special case of isotropic, homogeneous FLE with uniform

¹For FLE, FRE, and TRE, we use boldface (only) when we are denoting a vector quantity

²Also called “deviations”

weighting (i.e., un-weighted). In the next section, we extend that approach to the general case of anisotropic, inhomogeneous FLE with arbitrary weighting, and we use the approach both to investigate the dependence of FRE and TRE and to provide unified formulas for their respective covariances that are more general than any previous expressions.

III. Derivations

Our derivations are based on a linearization of the rigid, point registration problem. We begin with a statement of the registration problem, followed by the linearization of the problem for small FLE.

A. Registration Problem

The fundamental problem of rigid, fiducial registration is to find the rotation matrix R and translation vector t that minimize the expression

$$\text{FRE}^2 = \sum_{i=1}^N |W_i (R(x_i + \Delta x_i) + t - (y_i + \Delta y_i))|^2 \quad (3)$$

where FRE is the weighted fiducial registration error x_i, y_i , are 3×1 vectors representing corresponding points in two spaces, where at least three of the x_i are not collinear, Δx_i and Δy_i represent localization errors in the two spaces, W_i is a 3×3 non-singular weighting matrix, which may be a function of R (see the discussion of “ideal weighting” at the beginning of Section III-G). Clearly the solution, R, t , obtained by minimizing FRE^2 in (3) is unchanged if all the W_i are multiplied by the same factor w , while FRE^2 itself is multiplied by w^2 . Thus, size of FRE as defined by (3) is somewhat arbitrary. For example, the general form for the uniform weighting scheme employed by commercial systems is $W_i = wI$, where I is the 3×3 identity matrix, and for this weighting FRE equals times the root-means square of the distance between fiducials after registration. We can remove this arbitrary factor of w from FRE by stipulating that the W_i be normalized such that

$$\sum_{i=1}^N |W_i|^2 = \sum_{i=1}^N \text{trace}(W_i^t W_i) = 3. \text{ With this stipulation, uniform weighting requires that } W_i = N^{-1/2} I, \text{ and FRE recaptures its meaning as the unweighted root-mean square of the fiducial fitting errors and satisfies (1).}$$

The rotation matrix is restricted to be orthogonal, which means that $R^t R = I$. The points, y_i , arise from the application of the true, but unknown, rigid transformation, $R^{(0)}, t^{(0)}$, to the x_i , as follows:

$$y_i = R^{(0)} x_i + t^{(0)}. \quad (4)$$

$R^{(0)}, t^{(0)}$ is the transformation that is sought by the minimization of FRE in (3), but because of fiducial localization error, the transformation that minimizes FRE, namely R, t , will differ from the true transformation $R^{(0)}, t^{(0)}$. We model the Δx_i and Δy_i as random, independent vectors of fiducial localization error, with each of their three components being normally

distributed with zero means and covariances $\sum_i^{(1)}$ and $\sum_i^{(2)}$, respectively. Thus, the localization errors in the two spaces may be anisotropic and inhomogeneous. While these statistics are known, it is assumed that no other information about the errors themselves is available. If any relationships among the specific errors for a given set of localizations or if any components of any of the errors were known, that information could be employed in the registration process, and the derivations and proofs below would need to be modified. It is helpful to rewrite (3) as

$$\text{FRE}^2 = \sum_{i=1}^N |W_i (Rx_i + t - y_i - \Delta\xi_i)|^2 \quad (5)$$

where $\xi_i = y_i - R x_i$, is the “two-space” localization error vector for point-pair i , and its elements are normally distributed with zero means and covariances $R \sum_i^{(1)} R^t + \sum_i^{(2)}$. We define $\sum_i = R \sum_i^{(1)} R^t + \sum_i^{(2)}$ to be the “two-space” covariance of FLE. Once the rotation and translation that minimize FRE^2 have been found, they can be used to determine the vector **TRE** of target registration error at any point r that is not used as a fiducial point

$$\mathbf{TRE}(r) = Rr + t - (R^{(0)}r + t^{(0)}) \quad (6)$$

B. Linearization

Sibson showed how derivations of error statistics for this problem could be made tractable by (a) assuming that FLE is small, (b) approximating to first-order in FLE, (c) moving the origin of the coordinate system to the centroid of the x_i , (d) reorienting the coordinate system along the principal axes of the x_i , and (e) separating the rotational and translational problems [5]. He used this approach to investigate TRE; we and others have used it to investigate TRE [6], [7], [9], [11] and FRE [10], [19]. For our present investigation, we have developed a new approach. We employ (a) and (b), but we omit (c), (d), and (e). Thus, we solve to first order in FLE, but we retain the laboratory coordinate system, and we solve for both the rotation and the translation simultaneously. Solving the first-order problem allows us to cast the problem as a simple, unconstrained set of $3N$ linear algebraic equations in six unknowns of the form: $Cq = e$, where C is a $3N \times 6$ matrix, e is a $3N \times 1$ vector, and q is a 6×1 vector whose elements are $\theta_1, \theta_2, \theta_3, t_1, t_2, t_3$, where θ_k is an angle of rotation (in radians) about axis k , and t_k is a translation along that axis.

We begin by defining $x_i^{(0)} = R^{(0)}x_i$ and $t = t - t^{(0)}$, which allows us to transform the vector whose squared length appears in the summation in (3) as follows:

$$\begin{aligned}
R(x_i + \Delta x_i) + t - (y_i + \Delta y_i) &= R \left(R^{(0)t} R^{(0)} \right) (x_i + \Delta x_i) + t - R^{(0)} x_i - t^{(0)} - \Delta y_i \\
&= \left(RR^{(0)t} - I \right) R^{(0)} x_i + \Delta t - \left(\Delta y - R^{(0)} \Delta x_i \right) + \left(RR^{(0)t} - I \right) R^{(0)} \Delta x_i \quad (7) \\
&= \Delta R x_i^{(0)} + \Delta t - \Delta \xi_i^{(0)} + \Delta R R^{(0)} \Delta x_i
\end{aligned}$$

where we have used (4) and have defined

$$\Delta R = RR^{(0)t} - I \quad (8)$$

and

$$\Delta \xi_i^{(0)} = \left(\Delta y_i - R^{(0)} \Delta x_i \right). \quad (9)$$

Using (7) in (3) gives us

$$\text{FRE}^2 = \sum_{i=1}^N \left| W_i \left(\Delta R x_i^{(0)} + \Delta t - \Delta \xi_i^{(0)} + \Delta R R^{(0)} \Delta x_i \right) \right|^2. \quad (10)$$

From assumption (a), we know that the components of $\Delta \xi_i^{(0)}$ and x_i will be small. If they were all zero, then (10) tells us that we could achieve a perfect minimum of zero for (3) by setting R and t to zero. For finite but small $\Delta \xi_i^{(0)}$ and x_i , we see from this same equation that, when FRE is minimized, the elements of both R and t will be of first order in $\Delta \xi_i^{(0)}$ and x_i . We note that the elements of $R^t R$ can be expected to be small relative to those of R and that the elements of $RR^{(0)} x_i$ will be small relative to those of $\Delta \xi_i^{(0)}$, because they are all of second degree in the elements of $\Delta \xi_i^{(0)}$ and x_i , which themselves can be expected to be small. Following Sibson, we now make the assumption that the $\Delta \xi_i^{(0)}$ and x_i are small enough to render the elements of $R^t R$ negligible in comparison to the elements of R and the elements of $RR^{(0)} x_i$ negligible in comparison to the elements of $\Delta \xi_i^{(0)}$. We note that

$$\begin{aligned}
\Delta R^t \Delta R &= \left(RR^{(0)t} - I \right)^t \left(RR^{(0)t} - I \right) \\
&= I - R^{(0)} R^t - RR^{(0)t} + I \quad (11) \\
&= -(\Delta R^t + \Delta R)
\end{aligned}$$

where we have used the orthogonality of rotation matrices to go from the first line to the second. From (11), we see that, if the elements of $R^t R$ are indeed negligible in comparison to the elements of R , then $R^t + R \approx 0$, which means that R is antisymmetric, i.e., $R^t = -R$. Thus, the nonlinear restriction $R^t R = I$ is replaced by a linear one.

This relaxed constraint is essential to the original approach of Sibson on the statistics of FRE [5], the approaches of the subsequent work on the statistics of TRE [6], [7], [9], and our

current approach. However, we show in simulations below that our results hold to an excellent approximation even without linearization. The difference being that for the linearized equations we can prove that they hold for all configurations, all $\sum_i^{(1)}$ and all $\sum_i^{(2)}$, whereas for the nonlinear version, we can show agreement only for those cases that are simulated. Using our definitions of R and t in (6) gives

$$\begin{aligned} \mathbf{TRE}(r) &= (RR^{(0)t} - I) R^{(0)} r + t - t^{(0)} \\ &= \Delta R R^{(0)} + \Delta t \end{aligned} \quad (12)$$

where we have defined $r^{(0)} = R^{(0)} r$ (similarly to our definition above of $x_i^{(0)}$). We note finally, that, to first order $\sum_i = R^{(0)} \sum_i^{(1)} R^{(0)t} + \sum_i^{(2)}$.

C. Some Useful Results From Linear Algebra

The minimization of (10), subject to $R^t = -R$ with $RR^{(0)}$ x_i ignored in comparison to $\Delta \xi_i^{(0)}$ is equivalent to finding the least-squares solution to the following over-determined set of $3N$ linear equations in the six unknowns, $\theta_1, \theta_2, \theta_3, t_1, t_2, t_3$

$$(W_i \Delta R x_i^{(0)})_j + (W_i \Delta t)_j = (W_i \Delta \xi_i^{(0)})_j, \quad i=1, 2, \dots, N, \quad j=1, 2, 3, \quad (13)$$

where

$$\Delta R = \begin{bmatrix} 0 & -\Delta\theta_3 & \Delta\theta_2 \\ \Delta\theta_3 & 0 & -\Delta\theta_1 \\ -\Delta\theta_2 & \Delta\theta_1 & 0 \end{bmatrix} \quad (14)$$

and the subscript j enumerates the components of the three-dimensional vectors.

The meaning of R can be understood by adding I to both sides of its definition in (8) and multiplying on the right by $R^{(0)}$. The result, $(I + R)R^{(0)} = R$ shows that the rotation matrix that we seek by minimizing (3) is the equivalent to the true rotation followed by a ‘‘rotation’’ $I + R$. Thus, $I + R$ approximates the rotation away from the truth caused by fiducial localization error. If localization error were zero, then R would also equal zero. Our assumption that localization error is merely small results in the approximation, $R^t = -R$. The displacement r of a point r caused by that small rotation is $r = (I + R)r - r = Rr$, which, because R^t is equal to $-R$, can be written in the form of the familiar cross product for small rotations, $r = Rr = \theta \times r$, where $\theta = (\theta_1, \theta_2, \theta_3)$. This cross product produces the exact expression for r in the limit of small rotation angle

$|\Delta\theta| = (\Delta\theta_1^2 + \Delta\theta_2^2 + \Delta\theta_3^2)^{1/2}$. Equation (14) can also be derived by representing the exact rotation $I + R$ in terms of Euler angles, expanding each sine and each cosine in that representation via a Taylor’s series in the angles $\theta_1, \theta_2,$ and $\theta_3,$ and dropping all but

zeroth- and first-order terms from those expansions. The first-order terms all equal one and give rise to the identity I ; the first-order terms produce the expression given for R in (14).

We now define the $3N \times 6$ matrix C

$$C = \begin{bmatrix} W_1 X_1^{(0)} & W_1 \\ W_2 X_2^{(0)} & W_2 \\ \dots & \dots \\ W_N X_N^{(0)} & W_N \end{bmatrix} \quad (15)$$

where

$$X_i^{(0)} = \begin{bmatrix} 0 & x_{i3}^{(0)} & -x_{i2}^{(0)} \\ -x_{i3}^{(0)} & 0 & x_{i1}^{(0)} \\ x_{i2}^{(0)} & -x_{i1}^{(0)} & 0 \end{bmatrix} \quad (16)$$

and $x_{ik}^{(0)}$ is the k th component of $x_i^{(0)}$. By defining the $3N \times 1$ vector e as follows:

$$e_{3(i-1)+j} = (W_i \Delta \xi_i^{(0)})_j, \quad i=1, \dots, N, \quad j=1, 2, 3 \quad (17)$$

and using our definition of C , we can transform (13) into

$$Cq = e. \quad (18)$$

With (18), we have transformed the problem into a generic linear-algebra problem, and we need to establish some properties of the least-squares solution to that problem. We begin by considering the singular-value decomposition [15] of C

$$C = U \Lambda V^t \quad (19)$$

where U is a $3N \times 3N$ orthogonal matrix, Λ is a $3N \times 6$ diagonal matrix whose diagonal elements Λ_{ii} are the singular values of C , and V is a 6×6 orthogonal matrix³. We will require that the six singular values Λ_{ii} of C all be nonzero. This requirement is equivalent to the requirement that the columns of C be independent. We can prove that they are independent via proof by contradiction. Thus, we assume instead that the columns C of are linearly

dependent, which means that for some set of six constants, a_1, a_2, \dots, a_6 , $\sum_{i=1}^6 a_i C_i = 0$, where each C_i is a column of C . Because each column is of length $3N$, this constraint is equivalent to $3N$ equations, which have the form

³For convenience, we have augmented both U and Λ to the decomposition described in [15] by adding $3N - 6$ columns to U so as to make it orthogonal and adding Λ rows of zeros to U .

$$\begin{aligned} x_{i2}^{(0)} a_3 - x_{i3}^{(0)} a_2 &= a_4 \\ x_{i3}^{(0)} a_1 - x_{i1}^{(0)} a_3 &= a_5 \quad \text{for } i=1, 2, \dots, N. \quad (20) \\ x_{i1}^{(0)} a_2 - x_{i2}^{(0)} a_1 &= a_6 \end{aligned}$$

These $3N$ equations are in turn equivalent to the N vector equations $x_i^{(0)} \times a_{1,2,3} = a_{4,5,6}$, where $a_{\alpha,\beta,\gamma}$ is the 3×1 column vector whose elements are $a_\alpha, a_\beta, a_\gamma$. In order for all N of these cross products to be equal to the same constant vector N , all points $x_i^{(0)}$ must be collinear. We note from Section III-A above that at least three of the points x_i must not be collinear and that rotating and translating these points to produce the points $x_i^{(0)}$ does not change the linearity of the points. Thus, our assumption that the columns of C are dependent has led to a contradiction. Therefore, the columns of C are in fact independent, and therefore the singular values of C are all nonzero.

The least-squares solution of (18) is given as follows [15],

$$q_{\min} = C^+ e \quad (21)$$

where

$$C^+ = V \Lambda^+ U^t \quad (22)$$

and Λ^+ is a $6 \times 3N$ diagonal matrix for which $\Lambda_{ii}^+ = \Lambda_{ii}^{-1}$. We note that Λ_{ii}^+ are well defined because, as we have shown above, the Λ_{ii}^+ are nonzero. The elements of the $3N \times 1$ vector

$$f_{\min} = e - C q_{\min} \quad (23)$$

are the residual errors in the solution and can be rewritten as follows:

$$f_{\min} = (I - CC^+) e \quad (24)$$

where I is the $3N \times 3N$ identity matrix. Using (19) and (22) in (24), we have

$$\begin{aligned} f_{\min} &= (I - CC^+) e \\ &= (I - U \Lambda \Lambda^+ U^t) e \quad (25) \\ &= U (I - \Lambda \Lambda^+) U^t e. \end{aligned}$$

For $i, j = 1, \dots, 3N$, we have $(\Lambda \Lambda^+)_{ij} = \sum_{k=1}^6 \Lambda_{ik} \Lambda_{kj}^+ = \sum_{k=1}^6 \Lambda_{ik} \Lambda_{kj}^{-1}$, from which we find that the first six diagonal elements of $\Lambda \Lambda^+$ are equal to 1 and the last $3N - 6$ are equal to 0. Furthermore, we see that the first six diagonal elements of $I - \Lambda \Lambda^+$ are equal to 0 and the last $3N - 6$ are equal to 1. Using these diagonal elements of $I - \Lambda \Lambda^+$ in the last line of (25), and defining

$$\tilde{e} = U^t e \quad (26)$$

we find that

$$(f_{\min})_k = \sum_{j=1}^6 U_{kj} \cdot 0 \cdot \tilde{e}_j + \sum_{j=7}^{3N} U_{kj} \cdot 1 \cdot \tilde{e}_j = \sum_{j=7}^{3N} U_{kj} \tilde{e}_j. \quad (27)$$

Thus we see that each element of f_{\min} is a function of only the last $3N - 6$ elements of \tilde{e} .

D. FRE in Terms of \tilde{e}

We now return to the registration problem. We note that the elements of f_{\min} in (27) are the components of the individual fiducial registration error vectors, \mathbf{FRE}_i , of those points. Thus, $(f_{\min})_{3(i-1)+j} = \mathbf{FRE}_{ij}$ for $i = 1, \dots, N, j = 1, 2, 3$. Therefore, each component of \mathbf{FRE}_i for each fiducial point is a function of only the last $3N - 6$ elements of \tilde{e} .

E. TRE in Terms of \tilde{e}

Our expression for $\mathbf{TRE}(r)$ in (12) can be transformed into a simple matrix-times-vector multiplication

$$\delta = Dq \quad (28)$$

by defining $\delta_j = \mathbf{TRE}(r)_j, j = 1, 2, 3$ and defining the 3×6 matrix

$$D = \begin{bmatrix} 0 & r_3^{(0)} & -r_2^{(0)} & 1 & 0 & 0 \\ -r_3^{(0)} & 0 & r_1^{(0)} & 0 & 1 & 0 \\ r_2^{(0)} & -r_1^{(0)} & 0 & 0 & 0 & 1 \end{bmatrix} \quad (29)$$

where $r_k^{(0)}$ = the k th component of $r^{(0)}$. In (28), we use the least-squares solution of (18), as given by (21), and the decomposition of as given by (22), to get

$$\delta_{\min} = Dq_{\min} = DC^+ e = DV\Lambda^+ U^t e = DV\Lambda^+ \tilde{e}. \quad (30)$$

Thus,

$$(\delta_{\min})_k = \sum_{j=1}^6 \sum_{l=1}^{3N} (DV)_{kj} \Lambda_{jl}^+ \tilde{e}_l = \sum_{j=1}^6 (DV)_{kj} \Lambda_{jj}^{-1} \tilde{e}_j \quad (31)$$

where, in the second step, we have used the explicit form of Λ^+ given just after (22). From (31) we see that each component of $\mathbf{TRE}(r)$ is a function of only the first 6 elements of \tilde{e} .

F. General Formulation for the Covariances of FRE and TRE

Equation (31) can be employed to derive a general expression for Σ_{FRE} , the covariance of FRE, and Σ_{FRE} , the covariance of TRE. First, we define two $3N \times 3N$ block-diagonal matrices: a covariance matrix Σ for the entire set of FLEs and a weighting matrix W for the entire set of fiducial points. Each matrix is nonzero only within 3×3 blocks on the diagonal. Specifically, $\Sigma_{3(i-1)+j, 3(i-1)+j'} = (\Sigma_i)_{jj'}$ and $W_{3(i-1)+j, 3(i-1)+j'} = (W_i)_{jj'}$, for $i = 1 \dots N$, $j = 1, 2, 3$, $j' = 1, 2, 3$, and all other elements of each matrix equal zero. With these definitions it can be seen from the definition e of given by (17) that

$$\langle ee^t \rangle = W \Sigma W^t. \quad (32)$$

Our expression for Σ_{FRE} can be gotten by noting that $(\Sigma_{\text{FRE}})_{kl} = \langle (f_{\text{min}})_k (f_{\text{min}}^t)_l \rangle$ and then using (27), followed by (26) and (32), to get

$$\begin{aligned} (\Sigma_{\text{FRE}})_{kl} &= \sum_{j, j'=7}^{3N} U_{kj} U_{lj'} \langle \tilde{e}_j \tilde{e}_{j'} \rangle \\ &= \sum_{j, j'=7}^{3N} U_{kj} U_{lj'} (U^t W \Sigma W^t U)_{jj'}. \end{aligned} \quad (33)$$

The summation over three diagonal elements corresponding to the same marker provides an expression for an individual weighted FRE of that marker

$$\begin{aligned} \langle \text{FRE}_i^2 \rangle &= \sum_{k=3i-2}^{3i} \sum_{j, j'=7}^{3N} U_{kj} U_{kj'} \langle \tilde{e}_j \tilde{e}_{j'} \rangle \\ &= \sum_{k=3i-2}^{3i} \sum_{j, j'=7}^{3N} U_{kj} U_{kj'} (U^t W \Sigma W^t U)_{jj'}. \end{aligned} \quad (34)$$

The trace of Σ_{FRE} provides a general expression for $\langle \text{FRE}^2 \rangle$

$$\begin{aligned} \langle \text{FRE}^2 \rangle &= \sum_{k=1}^{3N} (\Sigma_{\text{FRE}})_{kk} \\ &= \sum_{k=1}^{3N} \sum_{j, j'=7}^{3N} U_{kj} U_{kj'} (U^t W \Sigma W^t U)_{jj'} \\ &= \sum_{j=7}^{3N} (U^t W \Sigma W^t U)_{jj}. \end{aligned} \quad (35)$$

The expressions in (33) through (35) give statistics for the weighted FRE. Statistics for unweighted FRE (but with weighting used in the registration step), which we designate by using lower case, fre , are given by replacing U_{jk} and $U_{kj'}$ by $(W^{-1}U)_{kj}$ and $(W^{-1}U)_{kj'}$ in each of these equations, resulting in

$$\langle \text{fre} \rangle_{kl} = \sum_{j,j'=7}^{3N} (W^{-1}U)_{kj} (W^{-1}U)_{lj'} (U^t W \sum W^t U)_{jj'} \quad (36)$$

$$\langle \text{fre}^2 \rangle = \sum_{k=3i-2}^{3i} \sum_{j,j'=7}^{3N} (W^{-1}U)_{kj} (W^{-1}U)_{kj'} (U^t W \sum W^t U)_{jj'} \quad (37)$$

and

$$\langle \text{fre}^2 \rangle = \sum_{j,j'=7}^{3N} (U^t W^{-2} U)_{jj'} (U^t W \sum W^t U)_{jj'}. \quad (38)$$

For isotropic weighting, i.e., $W_i = w_i I$, it can with some manipulation be seen that

$$\langle \text{FRE}_i^2 \rangle = w_i \langle \text{fre}_i^2 \rangle.$$

Our expression for Σ_{TRE} can be gotten by remembering that in $\text{TRE}(r)_j = \delta_j$ (31). Therefore,

$$\begin{aligned} \langle \Sigma_{\text{TRE}} \rangle_{kl} &= \langle (\delta_{\min})_k (\delta_{\min})_l \rangle \\ &= \sum_{j,j'=1}^6 (DV)_{kj} (DV)_{lj'} \frac{\langle \tilde{\epsilon}_j \tilde{\epsilon}_{j'} \rangle}{\Lambda_{jj} \Lambda_{j'j'}}. \end{aligned} \quad (39)$$

By applying (26) and (32) to $\langle \tilde{\epsilon}_j \tilde{\epsilon}_{j'} \rangle$, we achieve the following result:

$$\langle \Sigma_{\text{TRE}} \rangle_{kl} = \sum_{j,j'=1}^6 (DV)_{kj} (DV)_{lj'} \frac{(U^t W \sum W^t U)_{jj'}}{\Lambda_{jj} \Lambda_{j'j'}}. \quad (40)$$

The trace of Σ_{TRE} provides a general expression for $\langle \text{TRE}^2 \rangle$

$$\begin{aligned} \langle \text{TRE}^2 \rangle &= \sum_{k=1}^3 \langle \Sigma_{\text{TRE}} \rangle_{kk} \\ &= \sum_{k=1}^3 \sum_{j,j'=1}^6 (DV)_{kj} (DV)_{kj'} \times \frac{(U^t W \sum W^t U)_{jj'}}{\Lambda_{jj} \Lambda_{j'j'}}. \end{aligned} \quad (41)$$

It should be noted that (33) through (41) are completely general. Thus they are appropriate both for general fiducial localization error (i.e., possibly inhomogeneous and possibly anisotropic) and for general weighting (i.e., possibly not ideal). We have included in Fig. 4 in the Appendix an implementation of (33), (35), (40), and (41) written in Matlab (MathWorks, Inc., Natick, MA). This function is likewise completely general. While these equations have all been derived specifically for three-dimensional space, they all hold for the 2-D case as well with only minor changes to accommodate the reduced dimensionality. The details are given in Appendix C of the Appendix.

G. Statistical Independence of FRE and TRE for Ideal Weighting

In Fig. 1, we show examples for which FRE is zero and TRE is nonzero. It can easily be seen from (5) that the number of such examples is infinite. Indeed, whenever there exists a rigid transformation R, t for which $Rx_i + t = y_i + \xi_i$ for all i , then FRE is zero, but unless all the ξ_i are zero, TRE will be nonzero. The simplest example is for the localization errors in a given space to be equivalent to a global translation, i.e., all the x_i are equal to the same nonzero vector v_x , and all the y_i are equal to some other nonzero vector $v_y \neq v_x$. As a result, each $\xi_i = v_y - v_x$, the rigid transformation that minimizes FRE^2 is, $R = R^{(0)}$, $t = t^{(0)} + v_y - v_x$, and $\text{TRE}(r) = v_y - v_x \neq 0$. Substituting these values into (5) reveals that each of the terms in the summation is zero. Thus, each $\text{FRE}_i = 0$. There is likewise an infinity of examples in which TRE is zero and FRE is nonzero. These examples occur when no rigid transformation can reduce the effect of the localization errors, and the transformation that minimizes (5) is the identity. One such example is also shown in Fig. 1. These examples suggest that FRE may not be a good predictor of TRE, but they provide little insight into their statistical relationship. In this section we examine the statistical relationship between FRE and TRE for our model of independent, normally distributed localization errors when an important weighting scheme—“ideal weighting”—is employed, and we show that for this weighting scheme FRE and TRE are statistically independent.

Ideal weighting is important in the rigid point registration problem because it maximizes the probability that the resulting transformation is the true one. We derive ideal weighting in Appendix I-A of the Appendix and show that it has the form

$$W_i = w \times \sum_i^{-1/2} = w \times \left(R \sum_i^{(1)} R^t + \sum_i^{(2)} \right)^{-1/2}, \text{ where } w \text{ is a constant, which for}$$

normalized weighting is equal to $3^{1/2} \left(\sum_1^N \text{trace} \left(\sum_i^{-1} \right) \right)^{-1/2}$. Ideal weighting is not new. It was introduced into point registration as early as 1998 by Ohta and Kanatani [16], and it was employed by West *et al.* in 2001 [17] and Moghari *et al.* in 2009 [11], [19]. In both of these 2009 papers a first-order registration equation equivalent to (18) was derived for the special case in which ideal weighting is employed and $\sum_i^1 = 0^4$. As mentioned at the end of Section III-B, to first order R can be replaced by $R^{(0)}$ in the expression for \sum_i . This approximation is adequate for our purposes, and with it we can see that the R dependence of ideal weighting does not affect the first-order character of (13). Using this weighting in (32) gives us

$$\langle ee^t \rangle = W \sum W^t = w^2 \times \sum^{-1/2} \sum \sum^{-1/2} = w^2 I. \quad (42)$$

Since $\langle ee^t \rangle$ is diagonal, the elements e_i are uncorrelated; since they are uncorrelated and also normally distributed, they are statistically independent [2], [18]. Because the diagonal elements equal w^2 , the variance of each e_i is w^2 . As shown in Appendix I-B, since U is orthogonal, the elements of \tilde{e} are also independent and normally distributed with variance

⁴Equation (30) of [11] and (9) of [19].

equal to w^2 . In Section III-D, we found that each component of \mathbf{FRE}_i for each fiducial point, which is also an element of f_{\min} , is a function of only the last $3N - 6$ elements of $\tilde{\epsilon}$, while in Section III-E we found that each component of $\mathbf{TRE}(r)$, which is also a component of δ_{\min} , is a function of only the first six elements of $\tilde{\epsilon}$. Since the elements of $\tilde{\epsilon}$ are independent of each other, any functions of nonoverlapping subsets of them are independent each other as well. It is important to note that this conclusion depends crucially on the fact that all $3N$ of the elements of $\tilde{\epsilon}$ are mutually independent. That independence is a consequence of ideal weighting.

Therefore, for the case of ideal weighting, every component of \mathbf{FRE}_i of every one of the N fiducial points is statistically independent of any component of the vector $\mathbf{TRE}(r)$ at any point r . Thus fluctuations in $\mathbf{TRE}(r)$ are, in our first-order approximation, independent of fluctuations of any and all functions of the goodness of fit! We note that the overall fiducial registration error, FRE, from (3), is given in the linearized problem by (10) and can be expressed in terms of f_{\min} as $\text{FRE} = (f_{\min}^t f_{\min})^{1/2}$ and that the length of the vector $\mathbf{TRE}(r)$ can be expressed in terms of δ_{\min} as $|\mathbf{TRE}(r)| = (\delta_{\min}^t \delta_{\min})^{1/2}$. Since the elements of f_{\min} are statistically independent of those of δ_{\min} , any function of only f_{\min} is statistically independent of any function of only δ_{\min} . As primary examples, FRE is statistically independent of $|\mathbf{TRE}(r)|$. Since they are independent, they are also uncorrelated. We let $\text{CC}(\text{FRE}, \text{TRE})$ be the correlation coefficient of FRE and $|\mathbf{TRE}(r)|$

$$\text{CC}(\text{FRE}, \text{TRE}) = \frac{\langle (\text{FRE} - \mu_{\text{FRE}})(\text{TRE} - \mu_{\text{TRE}}) \rangle}{(\sigma_{\text{FRE}} \sigma_{\text{TRE}})} \quad (43)$$

where we have abbreviated $|\mathbf{TRE}(r)|$ as TRE, and are letting μ and σ denote means and standard deviations. It easy to see that, to first order in FLE, $\text{CC}(\text{FRE}, \text{TRE}) = 0$, meaning that there is to first order no correlation between FRE and TRE.

H. Two-Dimensional Space

The results of this section hold not only for 3-D space, but for 2-D space as well. The details are provided in Appendix I-C of the Appendix along with the 2-D forms of (33) through (41), but these equations can be transliterated to two dimensions merely by making the following changes to the summation limits: $3N \rightarrow 2N$, $7 \rightarrow 4$, $3i \rightarrow 2i$, $3i - 2 \rightarrow 2i - 1$, and $6 \rightarrow 3$.

I. Comparison With Previously Published Derivations

We compare our formulations with five previous derivations. These derivations, which are spread among many publications with differing notations, are, unlike our general derivation, limited in every case to uniform weighting (i.e., no weighting) or to ideal weighting. In our validation section below, we compare our formulas with these other authors' formulations on specific cases and find agreement in every case.

The first result is an expression for the probability density of FRE^2 when ideal weighting is employed. We note that $\text{FRE}^2 = f_{\min}^t f_{\min} = \sum_{k=1}^{3N} (f_{\min})_k^2$, which by (27) equals $\sum_{j=7}^{3N} \sum_{l=7}^{3N} \tilde{\epsilon}_j \tilde{\epsilon}_l \sum_{k=1}^{3N} U_{kl} U_{kj}$, and since U is orthogonal, we have $\text{FRE}^2 = \sum_{j=7}^{3N} \tilde{\epsilon}_j^2$. As noted in Section III-G, when ideal weighting is employed, the $\tilde{\epsilon}_j$ become statistically independent and normally distributed with equal variances ($= w^2$). Therefore, FRE^2 is chi-square distributed with $3N - 6$ degrees-of-freedom. The probability density of such a function is well known [22]

$$p(\text{FRE}^2) = \frac{\text{FRE}^{(3N-8)} e^{-(\text{FRE}^2)/2}}{w^2 2^{(3/2N-3)} \Gamma(\frac{3}{2}N-3)} \quad (44)$$

where Γ denotes the Gamma function⁵. This probability density is identical to that derived by Sibson for the case of isotropic, homogeneous FLE with no weighting [5]. Here we have not only re-derived a thirty-year-old result but also shown that it applies to a more general case, namely inhomogeneous, anisotropic FLE with ideal weighting. It differs from the distribution for ideal weighting published recently by Moghari [19] for fre (instead of FRE). Also, for ideal weighting, since the $\tilde{\epsilon}_j$ are independent and normally distributed with variances equal to w^2 , we have $\langle \text{FRE}^2 \rangle = (3N - 6)w^2$, and for normalized weighting

$$\langle \text{FRE}^2 \rangle = \frac{3(3N-6)}{\sum_1^N \text{trace}(\Sigma_i^{-1})} \quad (45)$$

For the special case of homogeneous FLE (45) reduces to $\langle \text{FRE}^2 \rangle = 3(3N-6) / \sum_k^3 \sigma_k^2$, where σ_k^2 is the variance of FLE along one of its principal axes, and for the case of homogenous, isotropic FLE, which is the case originally treated by Sibson, it reduces exactly to (1).

The second result is an expression for $\langle \text{FRE}^2 \rangle$ for ideal weighting, first derived for the case of isotropic, homogeneous FLE in 1998 [6] and given above by (2) and derived for general FLE in 2009 [11]. With ideal weighting we find from (32) and (42) that $(U^t W \Sigma W^t U)_{jj'} = \delta_{jj'}$ (i.e., the Kronecker delta). Using this result in (41) yields

$$\langle \text{TRE}^2 \rangle = \sum_{k,j=1}^{3,6} \frac{(DV)_{kj}^2}{(\Lambda_{jj}^2)} \quad (46)$$

Despite the considerable detail in our derivations of this result, its calculation is quite simple, requiring evaluation only of (16), (15), (19), (29), and (46) in that order.

The third result is an expression for $\langle \text{TRE}^2 \rangle$ when FLE is inhomogeneous and anisotropic but no weighting is employed, first derived in 2009 by Wiles first for the homogeneous case

⁵If the distribution for $|\text{FRE}|$ is desired, we note that $\rho(|\text{FRE}|) = 2|\text{FRE}|\rho(\text{FRE}^2)$.

[9] and then for the general case [23]. Our expression for this case is gotten simply by removing W from the expression in (41) and noting that, in the construction of the block-diagonal covariance matrix Σ , the nonzero blocks Σ , are all equal.

The fourth result is the decomposition of TRE into three statistically independent orthogonal normally distributed components when FLE is homogeneous and isotropic and no weighting is employed, which was first accomplished in 2001 [7]. We begin by noting that, for this case, having no weighting is equivalent to employing ideal weighting with $w = 1$. We then have $(U'W\Sigma W'U)_{jj'} = \delta_{jj'}$, as in the second result above. Using this result in (39), we find that

$(\Sigma_{\text{TRE}})_{kl} = \sum_{j=1}^6 (DV)_{kj} (DV)_{lj} / \Lambda_{jj}^2$. Performing an eigen decomposition yields $\Sigma_{\text{TRE}} = V_{\text{TRE}} \Lambda^{(\text{TRE})} V_{\text{TRE}}'$. Since Σ_{TRE} is symmetric, V_{TRE} is orthogonal. If we define $\eta = V_{\delta}' \delta_{\text{min}}$, then, since V_{TRE} is orthogonal and the elements of $\tilde{\epsilon}$ are independent and normally distributed, the three components, η_k , are uncorrelated, normally distributed, orthogonal components of TRE with variances equal to η_k .

The fifth result is the set of expressions in (36) through (38) for the statistics of the unweighted individual fiducial registration error with ideal weighting used the registration step when FLE is homogeneous or isotropic. Analytical expressions for these two cases were first derived in 2009 [19]. They reduce in the homogeneous *and* isotropic case to expressions first derived in 2008 [10].

IV. Validation

Our theory has three major results, which can be summarized as follows: For arbitrary weighting and arbitrary FLE, (a) (33) gives the covariance of FRE, (b) (40) gives the covariance of TRE, (c) FRE and TRE are statistically independent. Our other results are derived from (a) and (b). We tested our results as follows (1) We compared (33), (37), (41) and (46) to previously published analytic expressions for special cases. (2) We compared (33) and (39) to true values obtained by means of simulations, and we measured their cross correlation and dependence. All calculations were performed using Matlab 2010b (Mathworks, Natick, MA) on a Dell Latitude D830 with an Intel Centrino Duocore 2.2 GHz CPU with 2 GB RAM.

A. Comparisons to True Values via Simulations

For each simulation experiment, the following was performed: Randomly select a set $S(N)$ of N unperturbed fiducial positions x_i inside a 200 mm cube and one random target position r inside a 400 mm cube (the cubes share a common corner). Then repeat the following 15 times.

1. For each x_i , define a covariance FLE matrix $\Sigma_i^{(1)} = V_i^{(1)} \Lambda_i^{(1)} (V_i^{(1)})^t$, where the columns of $V_i^{(1)}$ are the principal axes of FLE and the diagonal elements $(\Lambda_i^{(1)})_{jj} = (\sigma_i^{(1)})_j^2$ the variances of the independent components of FLE.

2. Construct a random rotation $R^{(0)}$ and a random translation $t^{(0)}$ and apply them to the x_i to produce N corresponding unperturbed positions y_i [see (4)].
3. Generate a covariance FLE matrix $\sum_i^{(2)}$ for each y_i in a similar manner to that used to define $\sum_i^{(1)}$.
4. Generate a weighting matrix W_i for each point.
5. Repeat the following steps 10 000 times.
 - a. Perturb each fiducial position in each space to produce $x_i + \delta x$ and $y_i + \delta y$ according to $\sum_i^{(1)}$ and $\sum_i^{(2)}$, respectively, with random, normally distributed, zero-mean components.
 - b. Employ a registration algorithm to find R and t that minimize the exact registration formula of (3).
 - c. Calculate TRE(r) and FRE.
6. Calculate CC(FRE, TRE) for the 10 000 registrations of Step 5.
7. Perform the chi-square test described in Appendix I-D of the Appendix to detect dependence between FRE and TRE for the 10 000 registrations of Step 5.
8. Compute rms(TRE) and rms(FRE) for the 10 000 registrations of Step 5.
9. Calculate $\text{rms}(\text{TRE}) = \langle \text{TRE}^2 \rangle^{1/2}$ and $\text{rms}(\text{FRE}) = \langle \text{FRE}^2 \rangle^{1/2}$ our formulas.
10. Calculate CC between rms(TRE) from Steps 8 and 9; calculate CC between rms(FRE) from Steps 8 and 9.

The registration algorithm employed to accomplish Step 5b, depends on the weighting being used. For uniform weighting, the method of singular-value decomposition was employed (e.g., [20, Algorithm 8.1]). For ideal weighting, the algorithm of [21] was employed, except

that in Step 1 of that algorithm, $W_i = w \times \left(R \sum_i^{(1)} R^t + \sum_i^{(2)} \right)^{-1/2} = \langle \text{fre}_i^2 \rangle$, as given in Section III-G. This amendment to the algorithm is necessary for ideal weighting when FLE is anisotropic and/or inhomogeneous.

We employed these simulation experiments to determine the accuracy of 1) our expressions for TRE and FRE and 2) our claims that TRE and FRE are statistically independent and hence uncorrelated. As pointed out in [2], a failure to detect dependence does not guarantee that correlation is zero. The more specific correlation test may succeed where the dependence test fails, so we apply separate tests for correlation and dependence—the t-test to detect correlation and the chi-square test based on contingency tables to detect dependence [22]. The latter test is described Appendix I-D of the Appendix. For all tests we chose $p = 0.05$ with Bonferroni correction applied where appropriate. These tests were also used in [2].

V. Results

A. Comparison to Previously Published Expressions

We compared (33), (37), (41), and (46) with previously published and validated analytic expressions for the first three cases and the fifth case in Section III-A. Regarding (33), as pointed out in Section III-A, the simplification of that equation produced by the assumption of isotropy and homogeneity for FLE yields (45), which agrees exactly with the formula for FRE^2 derived by Sibson that case. We compared the outputs of (41) to the formulas for $V_i = \langle \text{TRE}^2 \rangle$ published by Moghari [11] and Wiles [9] for a large set of inputs. Similarly we compared (37) to the formula for $V_i = \langle \text{fre}_i^2 \rangle$ published by Moghari [19]. For each comparison, we repeated the following 1000 times: We generated sets $S(N)$ of random fiducials and a random target and in the manner described above for the simulations, with $N = 3, 4, 5, \dots, 10, 20, 30$ and 40 fiducials. We generated a two-space covariance Σ_i in the same manner as that used in the simulations to generate $\sum_i^{(1)}$, based on V_i and $(\sigma_i)_j$. We chose random orientations for each set of principle axes and chose random values of $(\sigma_i)_j$. We compared (46), with the formulation of Moghari, who provided an analytic formula for the case of isotropic, inhomogeneous FLE with ideal weighting [11], and we compared (41) with W set to the identity (i.e., no weighting) to the formula given by Wiles's for the case of anisotropic, homogeneous FLE with no weighting [9]. We compared (37) to Moghari's formula for the case of isotropic inhomogeneous FLE with ideal waiting [17]. In every case, our outputs agreed with those of these previously published formulas (maximum difference = 1.2×10^{-10}). In summary, our new general formulas agree with existing formulas, each of which is limited to a special case.

B. Comparison With Simulations

The important new feature that our theory supports is arbitrary weighting for arbitrary FLE. The sets of possible weightings and FLEs are each infinite, but we focused on two types of weighting—ideal weighting and uniform weighting, which is equivalent to no weighting—and on two sizes of FLE—small, which we defined as $\text{rms}(\text{FLE}) < 1$ mm, and large. For the latter case we used $\text{rms}(\text{FLE}) = 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60,$ and 70 mm. Our motivation for the small-versus-large-FLE dichotomy is to explore the limitations of our linear approximation, which is strictly correct only in the limit of small FLE. For small FLE, we tested with $N = 3, 4, 5, 6, 7, 8, 9, 10, 20, 30,$ and 40 fiducials. For the large FLE values, we limited our testing to four fiducials in order to control the requisite simulation time, which for each value of N requires 24 h for 15 values of $\text{rms}(\text{FLE})$. For the case of no weighting, we investigated only the case of small FLE, because we found correlation without moving above 1 mm.

Utilizing the simulation scheme above, in every case, we found the values calculated in Steps (8) and (9) to be statistically equivalent (no significant difference via paired t-test at $p < 0.05$). The maximum absolute percentage differences were less than 1.5% for FLE in the range from 1 to 10 mm ($n = 555$) and were less than 4.1% for FLE in the range from 1 to 50 mm ($n = 120$). For Step 10, in which we calculated the CC between the values of $\text{rms}(\text{TRE})$ of Steps 8 and 9, and the CC between the values of $\text{rms}(\text{FRE})$ calculated in Steps 8 and 9,

we found in both cases that CC was well above 99.9%. The minimal lower bound of 95% confidence was 99.99%. Thus, our formulas for both FRE and TRE agree with simulation and show only a slight degradation in accuracy even when FLE is far larger than any value that would be tolerated in surgical guidance (50 mm is equal to one-fourth of the side of the 200 mm cube within which the fiducials were randomly placed for this simulation).

Figs. 2 and 3 summarize the results of our investigation into the correlation and dependence of FRE and TRE. As predicted by our theory, neither correlation nor dependence is detectable when ideal weighting is employed, and FLE is less than 1 mm [Fig. 2(a)]. For $FLE > 20$ mm, correlation is statistically significant but remains negligible ($CC < 0.1$) even for $FLE = 70$ mm [Fig. 2(b)]. For uniform weighting (i.e., no weighting) a very small correlation is observed, $CC \approx 0.1$ that decreases with increasing N [Fig. 3(a)]. This correlation is statistically significant for all values of N , showing that, while FRE has some predictive power when higher order terms are included, that power is negligible. The dependence that this tiny correlation implies was undetectable for all but a very few cases. Detection is achieved when the chi-square statistic q (Appendix I-D of the Appendix) rises above a critical value shown by the horizontal line in Fig. 3(b). Such values were observed very rarely, as can be seen by the sub-critical means and, for $N = 20$, the sub-critical values of mean + standard deviation.

Finally, we repeated our experiments with targets positions chosen to be equal to fiducial positions in order to determine whether some dependence between TRE and FRE might appear in the special case in which a target is close to a fiducial. No such correlation is possible according to our theoretical development, but we explored this case to eliminate the possibility that appreciable second-order dependence might appear. The results were the same as for the randomly chosen target positions—no detectable dependence and no significant correlation.

VI. Discussion

We have generalized a method of first-order analysis of the point-registration problem introduced in 2009 [2] so that it accommodates inhomogeneous and anisotropic FLE with arbitrary weighting. With this generalization we are able to provide first-order expressions for the covariance matrices of FRE and TRE—(33) and (40)—an expression for individual FRE—(34)—and expressions for $\langle FRE^2 \rangle$ and $\langle TRE^2 \rangle$ —(35) and (41). As a result, we have unified the theory of first-order FRE and TRE statistics, which has heretofore been solved piecemeal only for special cases, each of which involved a separate formulation [6], [7], [9], [11], [19]. We have included in Appendix I-E of the Appendix an implementation of (33), (35), (40), and (41) in Matlab. With this implementation, we have reduced to 25 lines of Matlab not only all the many special cases of weightings and FLE anisotropies heretofore spread over many pages of many papers with different notations but also all other possible combinations of weightings, inhomogeneities, and anisotropies.

In addition, by means of our new method we are able to generalize the major result of [2]: namely, that, to first order, FRE and TRE are independent when FLE is homogeneous and isotropic. Our generalization reveals that, in addition, if ideal weighting is employed, FRE

and TRE are independent to first order when FLE is inhomogeneous and/or anisotropic. Ideal weighting, which we define in Section III-G, is not a new concept. It was introduced to point registration by Ohta and Kanatani in 1998 [16] and used by West *et al.* in 2001 [17] and by Moghari *et al.* in 2009 [11], [19] as well, but the proof that it produces independence between FRE and TRE is new.

We note that this independence in no way detracts from the importance of (1) and (2), from the generalizations of these expressions provided in the excellent work by Wiles [9], [23], Balachandran [10], and Moghari [11], [19], and further generalizations given in (33) through (38), (40) and (41), each of which relates one expectation value to another. Equation (1) and Wiles's generalization of that relationship [23] can be used to estimate the FLE covariance from measurements of FRE obtained from repeated registrations based on independent fiducial localizations. Based on this estimate, (2) or one of its generalizations can then be used to estimate an expectation value for TRE for a given fiducial configuration and target position. What we have shown is that the value of FRE observed for a given registration is unreliable as means to estimate the deviation of TRE from its expected value for that same registration.

We have compared our results to previously published formulas for FRE and TRE and to our own simulations to validate our new formulation. While there is an infinite variety of FLE patterns and weighting schemes available, we focused on inhomogeneous and anisotropic FLE with both ideal weighting and uniform weighting. We generated random fiducial and target positions and used both small FLE— $\text{rms}(\text{FLE}) < 1$ mm, and large FLE—up to 70 mm. We found excellent agreement in all cases. We also used simulations to evaluate correlation and dependence between FRE and TRE when the first-order approximation is not made. We found, in agreement with our theoretical results, that correlation and dependence approach zero for both weighting schemes as FLE approaches zero, for every case tested. We found for very large FLE that, while there is clearly a statistically significant correlation at higher order, that correlation is so small as to have negligible value as a predictor of TRE.

Our experiments on uniform weighting explain why neither Steinmeier nor Woerdeman, both of whom experimented with systems that apply this weighting scheme, were able to detect appreciable correlation between FRE and TRE, and it provides further support for the argument advanced in [2] that estimates of fluctuations of registration accuracy, i.e., TRE, based on the goodness of fit of the fiducials, e.g., FRE, are unreliable.

VII. Conclusion

By means of a new unified approach to the first-order analysis of the point registration problem, we are able to provide general expressions for the covariances of fiducial registration error and of target registration error. These two expressions and all the other expressions that we derived from them are appropriate for any set of fiducial localization error patterns, whether inhomogeneous or not and whether anisotropic or not. Furthermore, they are appropriate for any arbitrary set of weightings in the point registration problem, including ideal weighting and uniform weighting. Furthermore, this approach allows us to

show that, for ideal weighting, FRE and TRE are to first order neither correlated nor dependent. We have verified these results by means of comparison to previous derivations and to simulations. Finally, by means of simulation we have shown that, for all cases examined, correlation between FRE and TRE is also negligible for uniform weighting even for the exact case. These results reinforce the message delivered in [2] that for ideal weighting and uniform weighting, the latter of which is used in commercial guidance systems, fluctuations of measures of the goodness of fit of the fiducials, e.g., fiducial registration error, have no statistical relationship, or at most a negligible relationship, to fluctuations of individual measures of registration accuracy, i.e., target registration error. Therefore, as estimators of registration accuracy, they are unreliable and should approach with extreme caution both by the purveyors of guidance systems and by the practitioners who use them.

Acknowledgments

This work was supported in part by the National Institute on Deafness and Other Communication Disorders under Grant RO1DC008408 and Grant R01DC010184, in part by the National Institute of Biomedical Imaging and Bioengineering under Grant R01EB006136, and in part by Vanderbilt University's International Office.

The authors would like to thank the anonymous reviewers whose suggestions greatly improved the form and clarity of this paper.

References

1. Danilchenko, A.; Fitzpatrick, JM. General approach to error prediction in point registration. presented at the SPIE Med. Imag; San Diego, CA. Feb. 2010.;
2. Fitzpatrick, JM. Fiducial registration error and target registration error are uncorrelated. presented at the SPIE Med. Imag; Lake Buena Vista, FL. Feb. 2009.;
3. Steinmeier R, Rachinger J, Kaus M, Ganslandt O, Huk W, Fahlbusch R. Factors influencing the application accuracy of neuron-avigation systems. *Stereotactic Functional Neurosurg.* 2000; 75:188–202.
4. Woerdeman PA, Willems PW, Noordmans HJ, Tulleken CA, van der Sprenkel JW. Application accuracy in frameless image-guided neurosurgery: A comparison study of three patient-to-image registration methods. *J Neurosurg.* Jun; 2007 106(6):1012–6. [PubMed: 17564173]
5. Sibson R. Studies in the robustness of multidimensional scaling: Perturbational analysis of classical scaling. *Journal of the Royal Statistical Society: Series B.* 1979; 41:217–229.
6. Fitzpatrick JM, West JB, Maurer CR Jr. Predicting error in rigid-body, point-based image registration. *IEEE Trans Med Imag.* Oct; 1998 17(5):694–702.
7. Fitzpatrick JM, West JB. The distribution of target error in rigid-body, point-based registration. *IEEE Trans Med Imag.* Sep; 2001 20(9):917–927.
8. West JB, Maurer CR Jr. Designing optically tracked instruments for image-guided surgery. *IEEE Trans Med Imag.* May; 2004 23(5):533–545.
9. Wiles AD, Likholyot A, Frantz DD, Peters TM. A statistical model for point-based target registration error with anisotropic fiducial localizer error. *IEEE Trans Med Imag.* Mar; 2007 27(3): 378–390.
10. Balachandran, R.; Fitzpatrick, JM. The distribution of registration error of a fiducial marker in rigid-body point-based registration. presented at the SPIE; San Diego, CA. 2008.
11. Moghari M, Abolmaesumi P. Distribution of target registration error for anisotropic and inhomogeneous fiducial localization error. *IEEE Trans Med Imag.* Jun; 2009 28(6):799–813.
12. Suess O, Picht T, Kuehn B, Mularski S, Brock M, Kombos T. Neurosurgery without rigid pin fixation of the head in left frontotemporal tumor surgery with intraoperative speech mapping. *Operative Neurosurg Suppl.* 2007; 2(60):330–338.

13. Fitzpatrick JM, Galloway RL. The use of fiducials in the registration of 3D images. *Automedica Special Issue: Eng Approaches Neurological Surg.* 2001; 20
14. Fitzpatrick, JM. Detecting Failure, Assessing Success Medical Image Registration. Hajnal, JV.; Hill, DLG.; Hawkes, DJ., editors. Boca Raton, FL: CRC; 2001. p. 117-139.
15. Press, WH.; Teukolsky, SA.; Vetterling, WT.; Flannery, BP. *Numerical Recipes in C: The Art of Scientific Computing.* 2. New York: Cambridge Univ. Press; 1992.
16. Ohta, N.; Kanatani, K. *Lecture Notes in Computer Science.* Berlin, Germany: Springer; 1998. Optimal estimation of 3-D rotation and reliability evaluation; p. 175-187.
17. West, JB.; Fitzpatrick, JM.; Batchelor, PG. Point-based registration under a similarity transform. *Proc. SPIE Med. Imag.*; San Diego, CA. Feb. 2001; p. 611-622.
18. Papoulis, A.; U, Pillai S. *Probability, Random Variables, and Stochastic Processes.* 4. Boston, MA: McGraw Hill; 2002. p. 266-268.
19. Moghari M, Abolmaesumi P. Distribution of fiducial registration error in rigid-body point-based registration. *IEEE Trans Med Imag.* Nov; 2009 28(11):1791–1801.
20. Fitzpatrick, JM.; Hill, DLG.; Maurer, CR, Jr. Registration. In: Sonka, M.; Fitzpatrick, JM., editors. *Medical Image Processing: Volume II of the Handbook of Medical Imaging.* Bellingham, WA: SPIE; 2000. p. 447-513.
21. Balachandran, R.; Fitzpatrick, JM. Iterative solution for rigid-body point-based registration with anisotropic weighting. presented at the SPIE Med. Imag; Lake Buena Vista, FL. Feb. 2009.;
22. Spiegel, MR. *Probability and Statistics, Schaum's Outline Series.* New York: McGraw-Hill; 1975. p. 116-116.
23. Wiles, AD. Navigation accuracy of image-guided interventional systems. Dept. Med. Biophys., Univ. Western Ontario; London, ON, Canada: 2009.
24. Roche A, Malandain G, Ayache N. Unifying maximum likelihood approaches in medical image registration. *Int J Imag Syst Technol: Special Issue 3D Imag.* 2000; 11:71–80.

Appendix

Derivation of Ideal Weighting

The standard approach to finding the rigid transformation by minimizing (3) is based on maximum likelihood. Maximum likelihood was introduced into the point registration problem by Ohta and Kanatani in 1998 [16] and used by West *et al.* in 2001 [17] and by Moghari in 2009 [11], [19]. It has also been invoked for intensity-based registration methods [23]. The goal in point registration is to find the transformation that is most likely to be correct given the measured positions of the fiducial points in the two spaces. That goal can be reached only if the matrices W_i in that equation provide “ideal” weighting. In this section we derive the ideal weighting matrices and show that they provide the maximum-likelihood solution. We begin with the given sets $\{x'_i\}$ and $\{y'_i\}$, $i = 1, \dots, N$ of measured vector positions in two spaces, where $x'_i = x_i + \Delta x_i$ and $y'_i = y_i + \Delta y_i$, x_i and y_i are the true but unknown positions of corresponding points in the respective spaces, and Δx_i and Δy_i are their unknown localization errors. We wish to find the most likely rigid transformation, but, as we do, we will also find the most likely true positions. Specifically, we wish to find the rotation matrix and translation vector R, t and the $\{x_i\}$ and $\{y_i\}$ that together maximize the conditional probability density $p(R, t, \{y_i\} | \{x'_i\}, \{y'_i\})$. The set of true points $\{x_i\}$ is missing from the variables in this density because they are completely determined by R, t , and the $\{y_i\}$, since by definition $x_i = R^t y_i - R^t t$.

Thus, we wish to maximize $p(R, t, \{y_i\} | \{x'_i\}, \{y'_i\})$ with respect to R, t , and $\{y_i\}$ for fixed $\{x'_i\}$ and $\{y'_i\}$. From Bayes' law, we have that

$$p(R, t, \{y_i\} | \{x'_i\}, \{y'_i\}) = p(\{x'_i\}, \{y'_i\} | R, t, \{y_i\}) \frac{p(R, t, \{y_i\})}{p(\{x'_i\}, \{y'_i\})} \quad (47)$$

and, absent prior knowledge of $p(\{x'_i\}, \{y'_i\})$, we neglect its dependence on R, t , and $\{y_i\}$. Furthermore, $p(\{x'_i\}, \{y'_i\})$ is fixed because both $\{x'_i\}$ and $\{y'_i\}$ are fixed. Therefore, we may achieve our goal by maximizing the conditional probability $p(\{x'_i\}, \{y'_i\} | R, t, \{y_i\})$ with respect to R, t , and $\{y_i\}$.

We now assume that all localization errors are independent and that their components have zero mean and are normally distributed about their principal axes. With this assumption we have

$$p(\{x'_i\}, \{y'_i\} | R, t, \{y_i\}) = c \exp \left(- \sum_{i=1}^N (\Delta x_i^t \sum_{x_i}^{-1} \Delta x_i + \Delta y_i^t \sum_{y_i}^{-1} \Delta y_i) \right) \quad (48)$$

where \sum_{x_i} and \sum_{y_i} are respectively the covariances of x_i and y_i . It can be seen from our

definitions above and from minor manipulation that $\Delta x_i^t \sum_{x_i}^{-1} \Delta x_i$ equals

$(Rx'_i + t - y_i)^t S_{x_i} (Rx'_i + t - y_i)$, where $S_{x_i} = R \sum_{x_i}^{-1} R^t$, and that $\Delta y_i^t \sum_{y_i}^{-1} \Delta y_i$ equals

$(y'_i - y_i)^t S_{y_i} (y'_i - y_i)$, where $S_{y_i} = \sum_{y_i}^{-1}$. After these substitutions, we have

$p(R, t, \{x_i\}, \{y_i\} | \{x'_i\}, \{y'_i\}) = c \exp(-G(R, t, \{x'_i\}, \{y'_i\}, \{y_i\}))$, where

$$G(R, t, \{x'_i\}, \{y'_i\}, \{y_i\}) = \sum_{i=1}^N (Rx'_i + t - y_i)^t \times R \sum_{x_i}^{-1} R^t (Rx'_i + t - y_i) + (y'_i - y_i)^t S_{y_i} (y'_i - y_i). \quad (49)$$

Thus, in order to maximize the likelihood, we need to minimize $G(R, t, \{x'_i\}, \{y'_i\}, \{y_i\})$ with respect to R, t , and $\{y_i\}$. We minimize with respect to the $\{y_i\}$ by taking the derivative of G with respect to each of the components of each of the and setting all $3N$ derivatives to zero. The result is

$$y_i = (S_{x_i} + S_{y_i})^{-1} (S_{x_i} (Rx'_i + t) + S_{y_i} y'_i). \quad (50)$$

We note that the expression for y_i given in (50), which is a weighted mean of the i th transformed measured point in x space and its corresponding measured point in y space, provides the best available approximation for the true positions needed in all the expressions for TRE and FRE statistics derived in the body of this paper. In the case of isotropic,

homogeneous FLE, as, for example, in (2), this expression reduces to $y_i = ((Rx'_i + t) + y'_i) / 2$, i.e., a simple unweighted average.

Finally, we substitute our expression for y_i into G to obtain (after considerable manipulation)

$$G(R, t, \{x'_i\}, \{y'_i\}, \{y_i\}) = \sum_{i=1}^N (Rx'_i + t - y'_i)^t W_i^2 (Rx'_i + t - y'_i) \quad (51)$$

where $W_i = (R \sum_{xi} R^t + \sum_{yi})^{-1/2}$. In order to reach our goal, the remaining task is to find the transformation R, t that minimizes G . In this minimization process, we can use $W_i = w(R \sum_{xi} R^t + \sum_{yi})^{-1/2}$, where w is any constant. For this reason we refer to $W_i = w(R \sum_{xi} R^t + \sum_{yi})^{-1/2}$ as “ideal” weighting. The right side of (51) has the same form as the right side of (3), but the weighting is unspecified in (3), whereas the derivation above shows that only ideal weighting can achieve maximum likelihood.

Statistical Independence of the Elements of \tilde{e}

In this appendix we show that, if the M elements of the vector e are statistically independent, normally distributed variables with zero mean, then, if the $M \times M$ matrix V is orthogonal, the following vector has these same three characteristics:

$$\tilde{e}_i = \sum_{k=1}^M V_{ki} e_k. \quad (52)$$

1. *If the mean of e is zero, then the mean of \tilde{e} is zero.*

That the mean is zero is easily seen by computing the expectation value, which we denote by $\langle \cdot \rangle$, of both sides of (52)

$$\langle \tilde{e}_i \rangle = \left\langle \sum_{k=1}^M V_{ki} e_k \right\rangle = \sum_{k=1}^M V_{ki} \langle e_k \rangle = \sum_{k=1}^M V_{ki} \cdot 0 = 0. \quad (53)$$

2. *If the elements of e are uncorrelated with equal variances and V is orthogonal, then the elements \tilde{e} are uncorrelated.*

Because of (53), the correlation of \tilde{e}_i and \tilde{e}_j is equal to $\langle (\tilde{e}_i - \langle \tilde{e}_i \rangle) (\tilde{e}_j - \langle \tilde{e}_j \rangle) \rangle$

$$\begin{aligned} \langle (\tilde{e}_i - \langle \tilde{e}_i \rangle) (\tilde{e}_j - \langle \tilde{e}_j \rangle) \rangle &= \left\langle \sum_k V_{ki} (e_k - \langle e_k \rangle) \sum_l V_{lj} (e_l - \langle e_l \rangle) \right\rangle \\ &= \sum_k \sum_l V_{ki} V_{lj} \langle (e_k - \langle e_k \rangle) (e_l - \langle e_l \rangle) \rangle \\ &= \sigma^2 \sum_k \sum_l V_{ki} V_{lj} \delta_{kl} \\ &= \sigma^2 \sum_k V_{ki} V_{kj} = \sigma^2 \delta_{ij} \end{aligned} \quad (54)$$

where we used the fact that the elements of e are uncorrelated and equal in going from the first line of (54) to the second, defined $\sigma^2 = \langle e_i^2 \rangle$ to be the variance of an element of e , and used the orthogonality of V in the last step.

3. *If the elements of e are independent and normally distributed with equal variances and V is orthogonal, then the elements of \tilde{e} are independent and are normally distributed with the same variance.*

We note that, since the elements e_i are independent and normally distributed with variance σ^2 , their joint probability density has the form

$$\begin{aligned} p_{1,2,\dots,M}(e_1, e_2, \dots, e_N) &= \prod_{i=1}^M (\sigma \sqrt{2\pi})^{-1} \exp\left(\frac{-e_i^2}{2\sigma^2}\right) \\ &= (\sigma \sqrt{2\pi})^{-M} \exp\left(\frac{-1}{2\sigma^2} \sum_{i=1}^M e_i^2\right) \end{aligned} \quad (55)$$

where we have used the fact that the means of the e_i are zero⁶. Because each \tilde{e}_i is a linear combination of the e_i , their joint density has the form (by generalizing the derivations in [18, p. 199–201] from two variables to M variables)

$$q_{1,2,\dots,M}(\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_M) = |J|^{-1} p_{1,2,\dots,M}(z_1, z_2, \dots, z_M) \quad (56)$$

where $J = \det(V^t)$ and $z = (V^t)^{-1} \tilde{e}$. Since V is orthogonal, $|J| = 1$ and $(V^t)^{-1} = V$. Using these properties and (55) in (56) gives us that

$$\begin{aligned} q_{1,2,\dots,M}(\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_M) &= (\sigma \sqrt{2\pi})^{-M} \exp\left(-\sum_{i=1}^M \frac{(V\tilde{e})_i^2}{2\sigma^2}\right) \\ &= (\sigma \sqrt{2\pi})^{-M} \exp\left(-\frac{(V\tilde{e})^t V \tilde{e}}{2\sigma^2}\right) \\ &= (\sigma \sqrt{2\pi})^{-M} \exp\left(-\frac{\tilde{e}^t V^t V \tilde{e}}{2\sigma^2}\right) \\ &= (\sigma \sqrt{2\pi})^{-M} \exp\left(-\frac{\tilde{e}^t \tilde{e}}{2\sigma^2}\right) \\ &= (\sigma \sqrt{2\pi})^{-M} \exp\left(-\sum_{i=1}^M \frac{\tilde{e}_i^2}{2\sigma^2}\right) \end{aligned} \quad (57)$$

where we have used the orthogonality of V in going from the second line of (57) to the third line. Equation (57) can be written as a product of functions of the

individual \tilde{e}_i as follows: $q_{1,2,\dots,M}(\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_M) = \prod_{i=1}^M q_i(\tilde{e}_i)$, where

$q_i(\tilde{e}_i) = (\sigma \sqrt{2\pi})^{-1} \exp\left(-\tilde{e}_i^2 / (2\sigma^2)\right)$, which is of normal form. Since the joint probability density of the \tilde{e}_i is equal to the product of individual functions, the \tilde{e}_i are mutually statistically independent. Since the individual functions are of normal form, they are normally distributed with variance σ^2 .

⁶Zero means are not necessary for this derivation. They simply reduce its complexity. Since the means are in fact zero, we take advantage of that fact to simplify the equations.

Two-Dimensional Case

The formulas and the proof of independence are derived in Section III only for 3-D transformations, but they all carry over to the 2-D case as well. We highlight here the salient changes that result when the points and the transformations are 2-D. *First*, x_i, y_i, ξ_i , and $\Delta\xi_i^{(0)}$ are 2×1 . *Second*, $W_i, R, R^{(0)}, \sum_i^{(1)}, \sum_i^{(2)}, \Sigma_i, I$, and R , are now 2×2 matrices, t , and

$$\Delta R = \begin{bmatrix} 0 & -\Delta\theta \\ \Delta\theta & 0 \end{bmatrix} \quad (58)$$

Third, C is now a $2N \times 3$ matrix, and q is a three-element vector whose elements are θ, t_1, t_2 . *Fourth*,

$$X_i^{(0)} = \begin{bmatrix} -x_{i2}^{(0)} \\ x_{i1}^{(0)} \end{bmatrix} \quad (59)$$

and

$$e_{2(i-1)+j} = (W_i \Delta\xi_i^{(0)})_j, \quad i=1, \dots, N, j=1, 2 \quad (60)$$

Fifth, Λ is $2N \times 3$, V is 3×3 , Λ^+ is $3 \times 2N$, and the proof that all singular values (there are now only three of them) are nonzero involves only the third of the three equations in (20). It is now the first three (instead of first six) diagonal elements of $\Lambda\Lambda^+$ that are equal to 1, and the last $2N-3$ (instead of the last $2N-6$) are equal to 0. As a result, each component of \mathbf{FRE}_i for each fiducial point is a function of only the last $2N-3$ elements of \tilde{e}_i , and each component of $\mathbf{TRE}(r)$ is a function of only the first three elements of \tilde{e} . The proof that the $2N$ elements of \tilde{e} are independent makes no reference to the dimension of the space (see Appendix I-B above). Therefore, \mathbf{FRE}_i and $\mathbf{TRE}(r)$ are statistically independent for the 2-D case, just as they are for the 3-D case.

Sixth, the block-diagonal matrices have slightly different forms to accommodate the reduced spatial dimensionality: $\Sigma_{2(i-1)+j, 2(i-1)+j'} = (\Sigma_i)_{jj'}$ and $W_{2(i-1)+j, 2(i-1)+j'} = (W_i)_{jj'}$, for $i=1, \dots, N, j=1, 2, j'=1, 2$ with all other elements equaling zero.

Seventh, (33) through (41) must be altered to accommodate the reduced dimensionality, as follows:

$$\begin{aligned} (\Sigma_{\mathbf{FRE}})_{kl} &= \sum_{j, j'=4}^{2N} U_{kj} U_{lj'} \langle \tilde{e}_j \tilde{e}_{j'} \rangle \\ &= \sum_{j, j'=4}^{2N} U_{kj} U_{lj'} (U^t W \Sigma W^t U)_{jj'} \end{aligned} \quad (61)$$

$$\begin{aligned} \text{FRE}_i^2 &= \sum_{k=2i-1}^{2i} \left(\sum_{j,j'=4}^{2N} U_{kj} U_{kj'} \langle \tilde{e}_j \tilde{e}_{j'} \rangle \right)_{kk} \\ &= \sum_{k=2i-1}^{2i} \left(\sum_{j,j'=4}^{2N} U_{kj} U_{kj'} (U^t W \Sigma W^t U)_{jj'} \right)_{kk} \end{aligned} \quad (62)$$

$$\begin{aligned} \langle \text{FRE}^2 \rangle &= \sum_{k=1}^{2N} (\sum_{\text{FRE}})_{kk} \\ &= \sum_{k=1}^{2N} \sum_{j,j'=4}^{2N} U_{kj} U_{kj'} (U^t W \Sigma W^t U)_{jj'} \\ &= \sum_{j=4}^{2N} (U^t W \Sigma W^t U)_{jj} \end{aligned} \quad (63)$$

$$(\sum_{\text{fre}})_{kl} = \sum_{j,j'=4}^{2N} (W^{-1}U)_{kj} (W^{-1}U)_{lj'} (U^t W \Sigma W^t U)_{jj'} \quad (64)$$

$$\langle \text{fre}_i^2 \rangle = \sum_{k=2i-1}^{2i} \sum_{j,j'=4}^{2N} (W^{-1}U)_{kj} (W^{-1}U)_{kj'} (U^t W \Sigma W^t U)_{jj'} \quad (65)$$

$$\langle \text{fre}^2 \rangle = \sum_{j,j'=4}^{2N} (U^t W^{-2}U)_{jj'} (U^t W \Sigma W^t U)_{jj'} \quad (66)$$

$$(\sum_{\text{TRE}})_{kl} = \langle (\delta_{\min})_k (\delta_{\min})_l \rangle = \sum_{j,j'=1}^3 (DV)_{kj} (DV)_{lj'} \frac{\langle \tilde{e}_j \tilde{e}_{j'} \rangle}{\Lambda_{jj} \Lambda_{j'j'}} \quad (67)$$

$$(\sum_{\text{TRE}})_{kl} = \sum_{j,j'=1}^3 (DV)_{kj} (DV)_{lj'} \frac{(U^t W \Sigma W^t U)_{jj'}}{\Lambda_{jj} \Lambda_{j'j'}} \quad (68)$$

$$\langle \text{TRE}^2 \rangle = \sum_{k=1}^2 (\sum_{\text{TRE}})_{kk} = \sum_{k=1}^2 \sum_{j,j'=1}^3 (DV)_{kj} (DV)_{kj'} \frac{(U^t W \Sigma W^t U)_{jj'}}{\Lambda_{jj} \Lambda_{j'j'}}. \quad (69)$$

Chi-Square Test for Dependence

To detect dependence, we employ the standard chi-square test based on the contingency tables [22]. Each element, $C(i,j)$, $i = 1, 2, \dots, N_c, j = 1, 2, \dots, N_c$, of an $N_c \times N_c$ contingency table is a count of the number of registrations for which $\text{FRE}_i < \text{FRE} < \text{FRE}_{i+1}$ and $\text{TRE}_j < \text{TRE} < \text{TRE}_{j+1}$, where the thresholds, FRE_i and TRE_j can be chosen arbitrarily. The resulting table is then used to calculate the following test statistic:

$q = \sum_{i,j=1}^{N_C} (C(i,j) - C_0(i,j))^2 / C_0(i,j)$, where $C_0(i,j)$ is the count that is to be expected when FRE and TRE are independent, namely, $C_0(i,j) = C_1(i)C_2(j)/M$, where M is the total of all counts, $C_1(i) = \sum_{j'=1}^{N_C} C(i,j')$, and $C_2(j) = \sum_{i'=1}^{N_C} C(i',j)$. The variance in q over repeated experiments (i.e., with different random localization errors) can be reduced by choosing the thresholds FRE_i and TRE_j such that $\sum_{j'=1}^{N_C} C(i_1,j') \approx \sum_{j'=1}^{N_C} C(i_2,j')$ and $\sum_{i'=1}^{N_C} C(i',j_1) \approx \sum_{i'=1}^{N_C} C(i',j_2)$. If q is sufficiently large, then it can be concluded that FRE and TRE are dependent. A “sufficiently large” value is that of the chi-square distribution (i.e., cumulative distribution) $\chi_\nu^2(1-p_0)^7$, where $1-p_0$ is the desired confidence in the truth of dependence. The number, ν , of degrees of freedom for this problem equals $(N_C - 1)^2$. Thus, a value of q that is larger than $\chi_\nu^2(1-p_0)$ indicates that dependence between FRE and TRE is detected at $p = p_0$. Failure to detect q exceeding $\chi_\nu^2(1-p_0)$ suggests that our theory of independence can be extended from the linear approximation to the exact problem encountered in surgical guidance.

Computer Code to Implement the Derived Formulas

We provide in Fig. 4 a Matlab function to implement our new formulas for FRE and TRE. The values that it returns correspond to our equations as follows: $(\text{RMS_TRE})^2 = \langle \text{TRE}^2 \rangle$ from (41), $(\text{RMS_FRE})^2 = \langle \text{FRE}^2 \rangle$ from (35), $\text{Cov_TRE} = \Sigma_{\text{TRE}}$ from (40), and $\text{Cov_FRE} = \Sigma_{\text{FRE}}$ from (33). The inputs correspond to our definitions above as follows:

$$X_0(:,i) = x_i^{(0)} = R^{(0)} x_i, \quad W(:,i) = W_i, \quad \text{Cov_FLE}(:,i) = \sum_i R^{(0)} \sum_i R^{(0)t} + \sum_i^{(2)}, \quad \text{and } r_0 = r^{(0)} = R^{(0)} r.$$

Like the formulas that it implements, this code is general. This relatively simple function handles the special cases of uniform weighting for homogeneous, isotropic FLE, first solved by Fitzpatrick *et al.* [6], uniform weighting for homogeneous, anisotropic FLE first solved by Wiles *et al.* [9], and ideal weighting for inhomogeneous, anisotropic FLE first solved by Moghari *et al.* [11], and it handles all other cases as well.

⁷Note that χ_ν^2 denotes a variable, while $\chi_\nu^2(1-p_0)$ denotes a distribution.

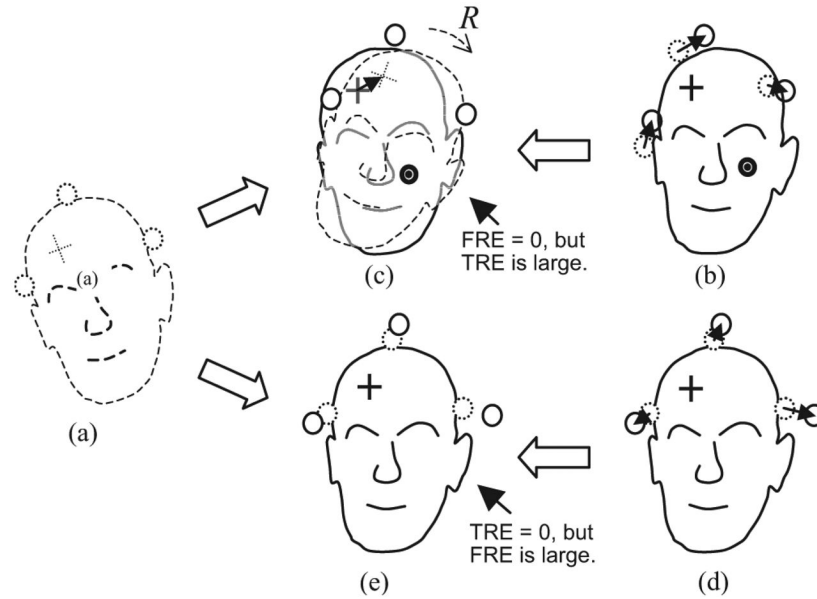


Fig. 1.

FRE and TRE for two example registrations (a) Image space showing patient (dashed outline), three fiducials (dotted circles), and an anatomical target (dotted cross). For simplicity localization errors are zero (b) Physical space showing a set of fiducial localization errors. Arrows show the displacements from true positions (dotted outlines) to localized positions (solid outlines). The same anatomical target is shown (solid cross). This set of localization errors can be duplicated exactly by a rigid transformation that comprises a clockwise rotation R about the “bull’s eye” that is located just to the right of the nose (c) Point registration has been applied to the image to register the localized positions in (a) with those in (b) The transformation is incorrect by the same rotational error R (dashed arrow) about the bull’s eye, but it achieves an FRE of zero. TRE (arrow) is large, however, because R is large (d) Physical space showing a second set of fiducial localization errors. This set of localization errors can be duplicated by expansion from the target point but cannot be approximated by any rigid transformation (e) Point registration has been applied to the image in (a), and the resulting rigid transformation is perfect. Since the transformation is perfect, TRE is zero, but since no rigid transformation can approximate the localization errors, the fiducial fit is poor, as can be seen from the relatively large size of the individual FREs (distance between circles with dotted and solid outlines, so the root-mean square FRE is large).

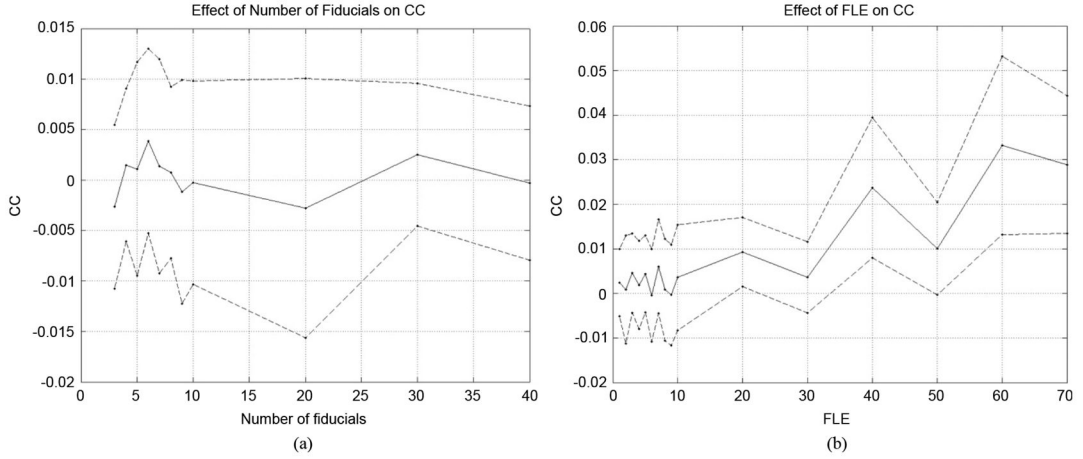


Fig. 2. FRE-TRE correlation coefficient CC for ideal weighting (a) FLE < 1 mm with varying number of fiducials (b) Varying FLE with four fiducials. In each case, the solid line is the mean of 15 sets of 10,000 registrations for randomly selected fiducial and target positions (see text). The dashed lines are mean \pm one standard deviation. Correlation is insignificant at $p < 0.05$ for all of (a) and for FLE < 10 mm in (b). Statistically significant but negligible (<0.1) correlation occurs for FLE ≥ 20 mm, showing that, while FRE has some predictive power when higher order terms are included, the power is negligible.

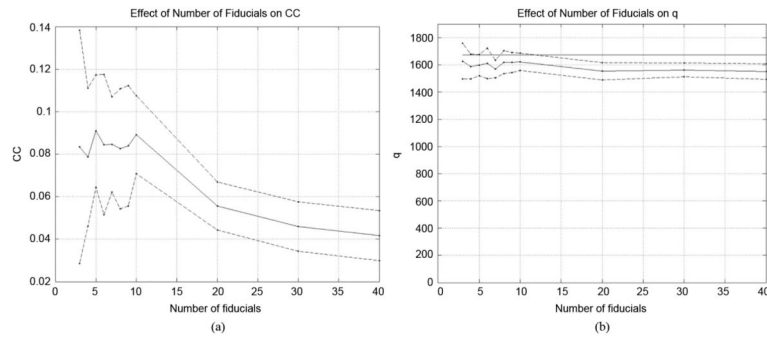


Fig. 3. FRE-TRE statistics for $\text{FRE} < 1$ mm when uniform weighting (i.e., no weighting) is employed (a) Correlation coefficient CC versus number of fiducials. A statistically significant but negligible (< 0.1), correlation is apparent that decreases with increasing number of fiducials (b) The chi-square statistic q for dependence versus number of fiducials. The dashed lines are mean \pm one standard deviation. The horizontal solid line in (b) is the critical level above which values of q indicate dependence. The slight dependence implied by the small correlation in (a) is not detected by q , whose mean remains below the critical level.


```

function [RMS_TRE,RMS_FRE,Cov_TRE,Cov_FRE]...
    = TRE_FRE_approx(X0,W,Cov_FLE,r0)
% [RMS_TRE,RMS_FRE,cov_TRE,cov_FRE] =
%     TRE_FRE_approx(X0,W,Cov_FLE,r0)
%
% Calculates root-mean-squares of TRE and FRE
% and covariance matrices of TRE and FRE for
% fiducials X0, weightings W, FLE covariance
% matrices Cov_FLE and target r0. X0 is a
% 3-by-N matrix with a fiducial point in each
% column, W is 3-by-3-by-N with a weighting
% matrix for one fiducial on each page.
% Cov_FLE 3-by-3-by-N with an FLE covariance
% matrix for one fiducial on each page. r0 is
% the 3-by-1 target. Note that FRE is the
% weighted fiducial registration error.
%
% Authors: Andrei Danilchenko and J Michael
% Fitzpatrick Created: August 2010

N = size(X0,2);

% Create matrix C and decompose it
W1 = W(:,1,:); W2 = W(:,2,:); W3 = W(:,3,:);
X1 = X0(1,:); X2 = X0(2,:); X3 = X0(3,:);
Xreshaped = reshape([ repmat(X1,3,1);...
    repmat(X2,3,1); repmat(X3,3,1) ], size(W));
X1 = Xreshaped(:,1,:); X2 = Xreshaped(:,2,:);
X3 = Xreshaped(:,3,:);
C = [-W2.*X3+W3.*X2, +W1.*X3-W3.*X1, ...
    -W1.*X2+W2.*X1, W1, W2, W3];
C = reshape(permute(C,[1,3,2]),[],6);
[U,S,V] = svd(C);

% Create 3N-by-3N block-diagonal matrices
% and matrix M:
temp = num2cell(W,[1 2]);
Wlarge = blkdiag(temp{:});
temp = num2cell(Cov_FLE,[1 2]);
Cov_FLElarge = blkdiag(temp{:});
M = U' * Wlarge * Cov_FLElarge * Wlarge' * U;

% FRE
U(:,1:6) = 0;
Cov_FRE = U*M*U';
RMS_FRE = sqrt(trace(Cov_FRE));

% TRE
Dleft = [0,r0(3),-r0(2);-r0(3),0,r0(1);...
    r0(2),-r0(1),0];
D = [Dleft,eye(3)];
DV = D * V;
S1 = [inv(S(1:6,1:6)), zeros(6,3*N-6)];
Cov_TRE = DV*S1*M*S1'*DV';
RMS_TRE = sqrt(trace(Cov_TRE));

```

Fig. 4.

Computer code to implement the derived formulas. The code is a function written in Matlab (The Mathworks, Natick, MA). Its input and output parameters are described in comments (lines beginning with a percent sign) and in the text.