FIDUCIAL-BASED REGISTRATION

WITH ANISOTROPIC LOCALIZATION ERROR

By

Andrei Danilchenko

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Computer Science

May, 2011

Nashville, Tennessee

Approved:

J. Michael Fitzpatrick

Benoit M. Dawant

Robert F. Labadie

Robert J. Webster III

ACKNOWLEDGEMENTS

I would like to thank all my committee members for their support throughout my dissertation research. First of all, I would like to thank Dr. Fitzpatrick for being the best possible advisor. For the past four and a half years he was the one who made me work a lot to be able to complete a PhD degree and to learn everything I know. He was a very understanding advisor who agreed to put up with all my habits and desires. He helped a lot during these four and a half years. In all good and bad situations I always felt his support. This dissertation would not have been possible without his guidance. Beside the scientific research, he taught me about writing papers and making presentation. His advises about dealing with people could never be overrated. Without him I would not be where I am right now.

I would like to thank Dr. Labadie for his great support over all these years. I was very fortunate to become a part of Computer-Assisted Otologic Lab as soon as I came to Vanderbilt. Working under his guidance in the lab was a pure pleasure. Even that he is an extremely busy person involved in several projects, he was always aware of each single detail of my work and ready to give his advice any minute. Letting me work on projects involving medical devices was something new for me, and that became something I really liked. In future I hope to continue working in this area.

I would like to thank Dr. Dawant, Dr. Webster, and Dr. Peters for being a part of my committee and their efforts and time. The classes of Dr. Dawant and Dr. Peters during my first year at Vanderbilt gave me the first real knowledge about medical imaging. Taking those classes made the whole future research much easier. The discussions about

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Registration is the process of determining a geometric transformation aligning the objects in two views so that corresponding features can be related [1]. In other words, registration is the process of mapping specific points of the object in one view to corresponding points in another view [2]. The two views are typically acquired from the same object. In most applications those views are two-dimensional images, such as x-ray image, or three-dimensional images, such as computed tomography (CT), magnetic resonance image (MRI), positron emission tomography (PET), or single photon emission computed tomography (SPECT). Additionally, in computer-assisted surgery (CAS) a model created by a computer may be aligned to a three-dimensional image of a patient. In image-guided surgery (IGS) the physical arrangement of the object in space is aligned to a three-dimensional, or in some cases two-dimensional, image of the same object. Because our research is highly concerned with IGS and CAS, we restrict our work to the cases for which registration is rigid, which means that registration involves finding a rigid transformation. A geometrical transformation is called rigid if it preserves all distances. It can be shown that it also preserves the straightness of lines, the planarity of surfaces, and all angles between straight lines. Furthermore, any rigid transformation can be accomplished by means of rotation followed by translation (the other order works as well). From now, with the term "registration" we will refer to rigid registration. Though our proposed algorithms work with images of two or three dimensions, specific examples and equations in this work will be shown for three-dimensional cases only.

1. History

Medical imaging as a new science was initiated with the discovery of x-rays in December 1895. Two weeks after the discovery, a first surgical procedure was undertaken with guidance by the pre-operative images: in order to remove an inserted broken needle, a radiograph of a patient's hand was aligned with the hand [1]. Since then, many methods have been created to increase the quality of image registration. A very important advance in medical imaging was made in 1947 by Spiegel et al. [3]. To guide surgery in humans, they registered plain images to the anatomy of patients. They used a stereotactic frame, which was rigidly attached to a patient's head and could stay in the same position from the beginning of acquiring images to the end of the surgery.

In the past 35 years medical imaging in general and image registration in particular have made a remarkable progress. The progress is mainly caused by the invention of three-dimensional imaging techniques, such as CT, MRI, SPECT, and PET. A wide range of imaging modalities, which mean methods of acquiring images, exists nowadays. It all started in 1973, when Hounsfield demonstrated CT technology [4] and Lauterbur introduced a primitive version of MRI [5]. In about five years both CT and MRI were clinically available for the surgeons. The major principle of CT is as follows: CT image is obtained by constructing the three-dimensional image from a series of lower dimensional x-ray projections taken at known directions. The principle of MRI, which uses a powerful magnetic field to align the nuclear magnetization of hydrogen atoms in water in the body of the patient [6], is more complicated, but it also constructs a three-dimensional image from a series of signals derived from lower dimensional projections.

All imaging modalities are good for showing the surgeon specific information about the patient. For example, PET and SPECT provide functional information, such as cellular activity. CT is good for highlighting bones, but not good for soft tissues. MRI, on the contrary, performs well on soft tissues and not so well on bones. Ultrasound provides information about discrimination of different tissue types. Doppler ultrasound shows flowing blood. Nowadays, in patients' diagnosis and treatment, surgeons often use more than one modality. For many surgeries they are required to align images acquired via different modalities, for example, CT and PET images. Besides that, using pre-operative images and post-operative images is a common practice for several surgeries now. Because of this growing use of multiple images for diagnosis and treatment, image registration has evolved from being a minor procedure to a significant subdiscipline in itself.

In the first half of the 1990s, a new big step in image registration was made when fully automatic retrospective algorithms were developed for aligning the images of the same modality [7], [8] and different modalities [8]-[12]. Currently, the usefulness of image registration methods is rapidly advancing because of the highly increasing power of computation systems. Some algorithms, which looked impossible 20 years ago because of complicated computations, now can be completed in a couple of seconds.

2. Registration methods

Since the introduction of multiple imaging modalities, a large number of registration approaches have been developed [13]-[15]. They can be classified in many ways, but in general all registration methods minimize some kind of cost function. For

several methods the cost function is a sum of distances between known corresponding points (point-based methods). For other methods it is the distance between the surfaces of the object (surface-based methods). For some methods it is a function of image intensities (intensity-based methods).

Point-based methods [2], [16]-[21] are the most intuitive registration procedures. If some points of a rigid object can be identified a priori on both views, then a point-based method can be used. The transformation that minimizes the sum of the squared distances between each pair of corresponding points needs to be found. (See "Derivation of ideal weighting" in Section 8 of Chapter II (page 70) for an explanation of the use of the squared distance.) Anatomical landmarks can be chosen as identified points for this type of registration. Often because of difficulties in accurate localization of anatomical points, special markers, which are designed in such a way that they can be accurately localized in both views, are used. The points used for point-based registration are called "fiducial points", or "fiducials". The markers from which the fiducials are derived are called "fiducial markers". In the standard algorithm on which most point-based methods are based, the translation is calculated as the difference between the means of the two point sets. The "de-meaned" points in each set are then calculated by subtracting the respective mean from each point, and one of the de-meaned sets is then rotated to minimize the total squared sum of the de-meaned point differences. Closed-form solutions exist for determining the transformations [22]-[27].

Surface-based methods [28]-[33] are based on aligning corresponding surfaces of the same object or two different objects in two views. Usually these methods are used when attachment of the fiducial is not feasible. Sometimes it is also hard to find good

natural landmarks for point-based registration. For example, the tip of the nose can be localized in physical space, but it is hard to localize it in CT or MRI. Even if it can be localized, there is a chance that skin-based landmarks will move relative to each other, which will cause additional error. In this situation, it is better to align the whole surface instead of a small number of landmarks. The most easily accessible surface on a patient's body is the surface of the patient's skin. When the surface is localized, only the coordinates in the direction perpendicular to the skin are important for determining the transformation. Thus, if the whole region slides along the surface, the registration will still appear be good.

The most popular surface-based algorithm is "Iterative Closest Point" algorithm [29]. It aligns two surfaces by iteratively performing three major steps: the closest point on the surface is found for each localized point, then a point-based algorithm is used to align these two sets of points, and finally surface points are replaced with the found ones. This algorithm does not perform well if the initial two sets of points are not approximately aligned. Recent patient studies have consistently measured accuracy of this method of two to five millimeters [34]-[36].

Volume-based methods recently have became the most widely used registration basis for applications in which two images of different modalities need to be registered [7], [37]-[41]. This approach uses intensities of image voxels to calculate a similarity measure which then is optimized – maximized or minimized. The advantage of volume-based methods is that they require less user-interaction than surface-based or point-based methods. The disadvantage is that there is no uniform algorithm that works for all images. Some similarity measures work well for one set of modalities and anatomical regions, but

fail for others. The most well known similarity measures are absolute difference, squared difference, correlation coefficient, ratio-image uniformity, partitioned intensity uniformity, joint histogram and joint probability distribution, joint entropy, mutual information, and normalized mutual information [2].

3. Image-guided surgery

In image-guided surgery, registration is used to provide the mapping between the arrangement of the patient in physical space and a three-dimensional image. This procedure is called image-to-physical registration. Image-to-physical registration is one of the key procedures in neurosurgery.

In 1906 Horsley and Clark performed the first image-guided surgery experiments on animals [42], [43]. They fixed a stereotactic frame on the head of an animal. About 40 years later, the first surgery on a human was performed using the same idea of rigidly-attached frame to allow guidance of the operation by the images. After the introduction of three-dimensional images, such as CT and MRI, image-guided surgery changed neurosurgery. In 1979 a surgery guided by CT image was performed [44]. It still used the stereotactic frame, though it was enhanced for three-dimensional nature of the images. In early 1980s, a new approach to image guidance became possible with the invention of methods for real-time updating of the images displayed on screens. Kelly with colleagues displayed tumor during surgery in the OR, updating the images and a cursor position based on the movement of the surgical microscope [45], [46]. Roberts with colleagues displayed similar image based on the tracking of a microscope [47]. An innovation in image-guided surgery was suggested by Allen in 1987 [48]. He suggested to use bone

implanted markers, or fiducials, instead of the stereotactic frame. Ten years later it was shown that with newly developed fiducial markers, it is possible to achieve the accuracy at the level of one millimeter in neurological guidance systems [20].

The registration problem in neurosurgery can be defined as follows. One of the views is the patient's three-dimensional image, usually CT or MRI. This image is typically obtained before the surgery. The other view is the physical arrangement of the patient in the space. During the surgery, the surgical tool is tracked, so that the position of the surgical tool's tip can be displayed on the screen in image space after performing registration of the patient in the two views. The registration should be performed with high accuracy, otherwise the surgeon may be misled, which is potentially dangerous for the patient.

The mapping between physical space and the image can be performed using any method from two categories – point-based registration and surface-based registration. Volume-based registration cannot be applied here because there is no good measure for intensities in physical space (ultrasound is blocked by the skull).

Because of the rigidity of a human head provided by the skull, IGS is now an important part of many neurological surgeries, but in the field of otologic surgery, where anatomy is also kept rigid by bones, it is relatively new. The temporal bone, which in human anatomy protects the inner ear, can be assumed to be rigid during surgeries. So can the region surrounded by the temporal bone, which makes it possible to use IGS systems for otologic surgeries. Thus, otology represents an excellent opportunity for new IGS advances. Before operating on the inner ear, a surgeon performs mastoidectomy – drilling out a part of the temporal bone. Some critical structures are hidden behind this

part of the bone. They are the facial nerve (an injury may result in paralysis of the face), the inner ear (an injury may result in permanent hearing loss and vertigo), the floor of the cranial vault (an injury may result on leakage of cerebrospinal fluid), and the internal jugular vein and carotid artery (an injury may result in blood loss which may be life threatening) [49]. With the help of IGS systems, an otologic surgeon can be more confident in his or her actions because by having pre-operative tomographic images mapped to the patient in the OR during surgery, the surgeon will have an option to see on a display how far the surgical tool is from the critical structures.

In 1999 the question of whether IGS is required in otologic surgeries was discussed [50]. In 2002 it was reported that using IGS might reduce the complications associated with procedures in sinus surgery [51]. Nowadays IGS is widely used for performing endoscopic sinus surgery in cases when landmarks are ambiguous or distorted. The use of IGS gives greater confidence and allows more complete exploration of the paranasal sinuses. The operation time is reduced, the accuracy of the system is within 2 mm [52]. This success in sinus surgery gives strong evidence that IGS can be an important advantage in ear surgery as well. In 2008 a minimally invasive, image-guided, facial-recess approach to the middle ear using a single drill hole was performed [53]. It was shown that accessing the cochlea using customized drill guides based on pre-operative CT scans and image-guided surgery technology can be safely accomplished with less destruction of the bone. This finding further predicts a bright future for IGS systems in otology.

The specification of ear surgeries does not always allow invasive fiducial systems. They are usually used only in the most malignant cases. The systems with skin-affixed

markers have errors of at least 1.5 mm because of possible skin movement relative to the skull [54]. The other possibility is using a surface-based registration method by using laser skin contouring. The experiments show that this method is also not suitable for otologic surgeries because of the relatively large errors (two to five millimeters). Recently, new fiducial systems were suggested specifically for the otologic surgeries [57], [58]. The major principle of these fiducial systems is taken from the stereotactic frames. However, the new fiducial systems are much smaller in size, and they provide submillimetric accuracy.

4. Computer-assisted surgery and robotic surgery

Any computer-based procedure using technologies such as imaging and real-time planning can be called a computer-assisted surgery (CAS). The major purpose and advantage of CAS is allowing much better visualization of the operated region and more accurate pre-operative surgical planning. That means that a surgeon can plan the operation in a pre-operative virtual environment and make sure that no critical structures will be hit during procedure and no damage will be done to the patient. It also has the potential to help during the surgery by decreasing the chance of surgical error, reducing the operating time, and improving the accuracy of the surgeon's gestures.

Robotic surgery implies the use of a robot during an operation on a patient. There are different robotic systems. The difference between them can be described in terms of the level of the surgeon's interactions during the surgery. There are three major categories: supervisory-controlled systems, telesurgical systems, and shared-control systems [79]. In supervisory-controlled system, the whole procedure is executed by a

robot which follows a pre-defined program that implements planning performed by the surgeon. The best known representative of this category is ROBODOC® (Curexo Technology Corp., Fremont, CA), which performs minimally invasive hip and knee replacement. In telesurgical systems, the robot does not follow a pre-defined program as in supervisory-controlled systems. Instead, with telesurgical systems the surgeon manipulates robotic arms during the surgery. The Da Vinci® surgical system (Intuitive Surgical Inc., Sunnyvale, CA) allows performing complicated minimally invasive surgeries with a low chance of surgical error and with reduced operating time. It provides the surgeon with a three-dimensional visualization of organs, such as heart, that are otherwise hard to see. It also gives the surgeon the opportunity to use tiny instruments. In shared-control systems, both the surgeon and the robot perform the operation. The robot is usually programmed in such a way that it recognizes safe territories and boundaries. If the surgeon is moving a surgical tool beyond the safety bounds, the robot warns the surgeon or limits the movement of the tool.

The workflow of a robotic surgery is as follows [80]. First, pre-operative images of the patient are acquired. Second, a three-dimensional model of the region of interest is constructed. Third, computer-assisted planning of the surgical procedure is performed. Finally, after performing the registration of the computer model to the patient in operating room, the surgery plan is executed.

Though robotic systems are widely used in urology, radiosurgery, orthopedics, neurosurgery, and other medical areas, there has not been such a wide usage of robots in otologic surgeries. The first method of computer-assisted surgery in head and neck surgery was developed in 1989 by Schloendorff [59]. During the next 15 years, there

were no reported ENT applications; only cadaveric and animal studies were made [60]. In 2005 the first robotic otologic surgery was performed by McLeod and Melder [61].

5. Definition of errors

Regardless of the type of fiducial markers used for point-based registration and methods of fiducial localization, in real-life applications it is impossible to align the markers perfectly. The reason for that is that it is highly unlikely to get an absolutely accurate location of a point in any view. Fortunately, registration does not necessarily need to be perfect; it just needs to be adequate for the problem. To understand how good a registration is, the following error measures are used.

Fiducial Localization Error: Fiducial localization is the process of estimating the geometrical positions of a fiducial point, such as the centroid of a fiducial marker. In all applications fiducial localization must be performed in both views. Regardless of the method used for localization, whether it is interactive visual identification of anatomical landmarks or an automatic algorithm for localizing a fiducial marker, the point will inevitably be erroneously displaced somewhat from its correct location [2]. This displacement is called "fiducial localization error" (FLE). FLE of each fiducial is present in both views, but its value is unknown. This error is the root cause of further errors in the registration process. In most cases, the mathematical definition of FLE is the Euclidean distance between the true point location and estimated fiducial location (if direction is considered to be important, it may be measured as a vector displacement).

Systems for image-guided surgeries that are based on fiducials are implemented with fiducial registration algorithms that assume statistical independence of FLE of each

fiducial in the system (see "Derivation of ideal weighting" in Section 8 of Chapter II (page 70)). FLE of bone-implanted markers is modeled as random error produced by imaging artifacts, fiducial fixation defects, and intra-operative tracking in accuracy. The assumption about the independence of such FLE is reasonable for bone-implanted fiducials, but for skin-affixed markers FLE independence is highly questionable. The reason is that FLE of such markers will include, in addition to the errors listed above for bone-implanted markers, errors caused by fiducials movement as a result of skin movement. Because skin movement may span more than one marker, marker movements may be correlated, resulting in correlated FLE among them. Nevertheless, FLE among the markers is usually treated as approximately independent by the registration algorithms, regardless of whether they are bone-implanted or skin-attached. It is also assumed that FLE is normally distributed with zero-mean.

Fiducial Registration Error: As mentioned earlier, because of the presence of FLE, the corresponding fiducial points in two views are not aligned perfectly after the registration. The Euclidean distance between the corresponding fiducial points after registration is called "fiducial registration error" (FRE) (if direction is considered to be important, FRE may be measured as a vector displacement). Though FRE can be calculated for each fiducial in the system and thus be used as a measure of registration accuracy, it is more common to use the root-mean-square (RMS) of the set of Euclidean distances between corresponding points in a given registration as a registration accuracy measure. In this work we will refer to the latter definition when we talk about FRE. In the general case, FRE is weighted differently for different markers and/or different directions. Thus, FRE for different markers (and possibly different directional

components) may be multiplied by differing weights in the calculation of the RMS. Usually FRE is lacking in any intrinsic clinical meaning, because it only shows how well the fiducials are aligned. Moreover, as we discuss later, FRE is not a reliable measure of registration accuracy.

Target Registration Error: A displacement between two corresponding points not used to determine the registration transformation but after applying the transformation is called "target registration error" (TRE). The term "target" is used because of the implication that the error is being measured at some important anatomical position within a region of interest. Because of the presence of FLE, targets are rarely aligned perfectly. If the direction of the error is important for an application of registration, the displacement vector of TRE is reported. Otherwise, typically the Euclidean distance between corresponding targets after registration is referred to as TRE.

Figure 1 shows a schematic to explain the errors defined above. FLE, FRE, and TRE are here depicted as vector quantities, though they are usually reported as scalar RMS values. As said above, FRE does not have a practical use for the surgeon. TRE is the true measure of system accuracy. Only TRE gives the surgeon the understanding of how large the error is at certain position while using the system.
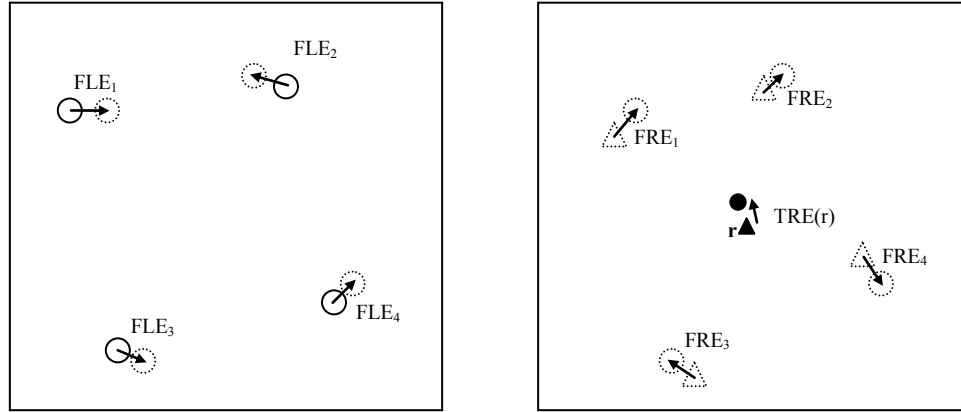
Figure 1. Schematic showing various errors in registration. Solid circles represent true position of the fiducials. Dashed circles and dashed triangles represent the localized position of the fiducials. Dashed triangles represent the transformed localized positions produced by registration. The filled circle represents the true target point in the right-hand space. The filled triangle represents the target point transformed from the left-hand space. Arrows depict errors: Arrows between solid and dashed circles are fiducial localization errors (FLE). Arrows between dashed triangles and circles represent individual fiducial registration errors (FRE). The arrow between the solid triangle and circle represent target registration error TRE(**r**) at position **r**.

Target Localization Error: In systems using a probe to localize fiducials, there is one more important source of inherent error. Because the probe tip position is usually calculated using fiducials on the same probe, the localization errors of those fiducials cause an erroneous calculation of the probe tip location. The distance from the true tip position to the expected one is called "target localization error" (TLE).

Total Target Error: If there is TLE present in the system, since it can be expected to be uncorrelated with TRE, the RMS value of the "total target error" (TTE) at target **p** can be calculated using this formula: $\left(\mathrm{rms\,TTE}(\mathbf{p})\right)^2 = \left(\mathrm{rms\,TRE}(\mathbf{p})\right)^2 + \left(\mathrm{rms\,TLE}(\mathbf{p})\right)^2$. In presence of a probe used to localize fiducials in the system, TTE is the proper measure to be reported to the surgeon to analyze the accuracy of the system.

14

6. Point-based registration

Fiducial localization in two views produces two sets of corresponding fiducial points, $\mathbf{p}_i$ and $\mathbf{q}_i$ for $i=1...N$. In point-based registration, a rigid transformation is sought of the form $\mathbf{p}'_i = R\mathbf{p}_i + \mathbf{t}$, where for the three-dimensional case R is a 3-by-3 rotation matrix, which by definition means that $R^t R$ equals the identity matrix, and $\mathbf{t}$ is a 3-by-1 vector, whose elements are the components of translation along the x, y, and z axes. This optimum R and $\mathbf{t}$ are found, meaning that they maximize the goodness of fit of the fiducials, by minimizing a weighted sum of the squares of the distances, $\left|\mathbf{p}'_i - \mathbf{q}_i\right|$, between corresponding fiducials after the transformation has been applied. In particular, the quantity,

$$\text{FRE}^2 = \sum_{i=1}^{N} \left| W_i \left(\mathbf{p}'_i - \mathbf{q}_i \right) \right|^2, \tag{1}$$

is minimized, where W_i represents 3-by-3 weighting matrix for the i-th marker. W_i may differ from fiducial to fiducial and may also depend on spatial direction as well. Directional dependence means that FLE is anisotropic. Up to today there is no closed-form solution for this problem. Iterative approaches [62]-[65] are used to find registration parameters. However, for many applications the anisotropy is typically small and is ignored. Making this assumption, we get the isotropic problem, for which W_i is a scalar. In the statistics literature, the minimization of Eq. (1) is known as the "Orthogonal Procrustes Problem". For this problem, closed-form solutions for rigid transformations do

exist. The first solution was given by Green [22], followed by other published solutions [23]-[27]. All of those solutions use singular-value decomposition, first introduced to this problem by Schönemann [23] for computation of rotation matrices or finding unit quaternions to represent rotations [26]. In our work, we will follow a simple solution described in [2] that uses singular-value decomposition of the cross-covariance matrix of the fiducial positions.

7. Error Measures

As mentioned above in Section 5 (page 11), during fiducial point-based registration there is one major source of errors – FLE. As was stated earlier, FLE cannot be measured directly. On the other hand, FLE can be characterized statistically. Typically, FLE for a particular marker is different in the two spaces. Let $\left\langle FLE_1^2 \right\rangle$ and $\left\langle FLE_2^2 \right\rangle$ be the expected squared FLE in the two spaces, where $\langle . \rangle$ represents expected value. Then making the usual and reasonable assumption that the errors are uncorrelated, the total FLE is related to FLE in the two spaces as follows [66]:

$$\left\langle FLE^2 \right\rangle = \left\langle FLE_1^2 \right\rangle + \left\langle FLE_2^2 \right\rangle. \tag{2}$$

Because of the lack of a closed-form solution of anisotropic registration problem and the infeasibility of running iterative methods due to large computation time and resources, for most applications it was assumed that FLE is small relative to the spread of the fiducials. The following formulas are derived under this assumption.

In 1979 Sibson found a relationship between FRE and FLE that holds for small values of FLE and can be used to compute the expected value of the total FLE [67]. The relationship is

$$\left\langle FRE^2 \right\rangle = \left(1 - 2/N\right)\left\langle FLE^2 \right\rangle, \tag{3}$$

where N is the number of fiducials used for the registration. From this equation it is clear that FRE is independent of the fiducial configuration.

The first formula for calculating the expected squared value of TRE was given by Fitzpatrick et al. in 1998 [68]. According to their theory, TRE at any target position r can be estimated using following formula:

$$\left\langle TRE\left(r\right)^2 \right\rangle = \frac{1}{N}\left(1 + \frac{1}{3}\sum_{k=1}^{3}\frac{d_k^2}{f_k^2}\right)\left\langle FLE^2 \right\rangle, \tag{4}$$

where d_k is the distance of the target r from the principal axis k of the fiducial set, and f_k is the RMS distance of the fiducials from the same axis. This formula quantifies the effect on registration accuracy of the target's position relative to the fiducial configuration. The following characteristics are derived from the above formula (omitting the expectation-value symbol): a) TRE is proportional to FLE, b) TRE depends on the configuration and the number of the fiducials used for the registration, c) TRE depends on the position of the target, d) the minimum value of TRE is equal to $FLE/\sqrt{N}$. This minimum is reached at the fiducial configuration centroid and is purely the translational

component of the registration error. Its value decreases as the number N of fiducials increases, e) TRE increases as the target point moves away from the centroid, and it exhibits ellipsoidal isocontours.

A later result by Fitzpatrick and West [69] showed that TRE is anisotropic in nature. TRE's component along the direction parallel to the line joining the origin of the principal axes to the target point has a lower value relative to the components along the other two perpendicular directions.

In the last 10 years attempts have been made to find the formula for estimating TRE and FRE for the case without neglecting anisotropy. The reason is widely-used in real-life applications optical tracking systems. In those systems, the error along the direction parallel to the line from the tracking device to the object is about three times larger than errors in the two perpendicular directions. Wiles et al. [70], [73] and Moghari et al. [71], [72] recently derived formulas for estimating TRE and FRE. However, their derivations and equations work only for specific cases of FLE distribution.

8.   Overview of dissertation

Here we briefly describe the main problems we set out to solve in our work and give an overview of the remainder of the dissertation.

Finding a new general method of estimating the point-based registration accuracy (Chapter II)

As was described above, for a long time in calculations of the registration accuracy measures it was assumed that the anisotropy is very small and thus can be

neglected. The formulas for estimating TRE and FRE were derived on this basis. But with the increase of usage of optically tracked systems in medical applications and the importance of accurate image alignment, the fact that FLE in one direction is much larger than FLE in other directions cannot be neglected anymore. A couple of research groups found formulas for estimating registration accuracy measurements in specific cases. Yet, neither of the developed methods was able to produce general formulas for TRE and FRE. In Chapter II we derive the general form of formulas estimating TRE and FRE. Our formulas work for any fiducial configuration, for any FLE distribution, and for any weighting during registration. Any FLE distribution means that FLE in either of two spaces can be homogeneous or inhomogeneous – same or different FLE for all fiducials – and isotropic or anisotropic – same or different FLE in all directions for each fiducials. Our formulas are a unification of first-order TRE and FRE statistics, which was previously developed piecewise for special cases. We present formulas for estimating TRE and FRE calculating covariance matrices for TRE and FRE and estimating individual FRE.

In addition, for the case of ideal weighting we show that FRE and TRE are independent to first order for any FLE distribution. This result disproves a popular belief that small FRE corresponds to small TRE, previously disproved only for one special case – isotropic FLE with uniform weighting. Both analytically and experimentally we show that small FRE does not always lead to small TRE. Thus, even though there is dependence between the covariances of these two errors, for each given case the value of FRE does not give any good estimation of TRE.

To prove the correctness of our formulas and independence theory, we present the results of a set of experiments. In those experiments, we used computer simulations to find true values and compared them to values calculated using our new formulas.

At the end of Chapter II we provide a computer program for calculating covariance matrices and estimates of TRE and FRE.

Developing and evaluating a feasible method of solving the problem of anisotropic point-based registration (Chapter III)

As mentioned above, modern tracking systems cause high anisotropy in localizing fiducials in physical space. Using registration methods disregarding the anisotropy and instead looking for isotropic solution increases overall registration error. Because tracking systems are widely used for medical applications on real patients, the accuracy of such systems becomes an important issue. Thus, solving anisotropic point-based registration is of big value. To our knowledge, no closed-form solution of anisotropic point-based registration problem exists. Recently some iterative algorithms solving the anisotropic registration problem were developed. Most of them require multiple parameters adjustment before registration. In Chapter III we present and evaluate a new iterative algorithm solving the problem of anisotropic point-based registration that requires just one parameter adjustment.

Like the formulas for estimating errors presented in Chapter II, the registration algorithm presented in Chapter III is general. Specifically, it works for any non-linear fiducial configuration (no solution exists for a linear configuration), any type of FLE – homogeneous or inhomogeneous, isotropic or anisotropic, and any set of weights.

To evaluate the new algorithm, we selected for comparison two other iterative algorithms that solve anisotropic point-based registration. The criterion for selecting the algorithms was the number of parameters which need to be set. We selected two algorithms which similar to our new algorithm require just one parameter – the stopping criterion. For those selected algorithms, we ran computer simulations as well as an experiment with real data. All the experiments showed that our new algorithm produces a smaller TRE than either of the selected algorithms or the standard approach (closed-form solution for uniform weighting)

We explored the space of point registration problems to examine the robustness of the new algorithm. The goal was to find cases for which the algorithm "fails". By failing we mean not converging or producing error bigger than other algorithms. Because an exhaustive search of such a large space is not possible, we limited the search to the case of surgical operations. We ran computer simulations for all reasonable fiducial configurations for a medical surgery. There were no cases noted when the new algorithm presented in this chapter failed.

We then looked at the case when fiducial positions are nearly collinear, which is known to cause trouble for the majority of registration methods. We found that for some nearly linear configurations, the new algorithm fails, but for any medically applicable fiducial configuration the new algorithm converges well and produces good results.

We also looked at the time and the number of iterations required for converging. We compared those values of our new algorithm to the ones of the two selected methods. While the number of iterations is larger than for the other two methods, each iteration takes a much smaller amount of time, which leads to the smaller overall time of execution.

Increasing the accuracy of IGS (Chapter IV)

Potentially IGS might be executed with more accuracy than it is now. A set of factors influences the resulting error. To minimize the overall error, all errors of the surgical system should be minimized. One of those is tracking error. The algorithm used to track objects in modern systems is as follows: localize each trackable marker of an object, and then run point-based registration of the localized markers to a model of the configuration of markers of that object, which is stored in an object's configuration file. In the standard tracking algorithm it is assumed that the markers are localized with no anisotropy or inhomogeneity. Homogeneity is a reasonable assumption (see next paragraph for an explanation), but for the majority of tracking systems there is a high localization error in the direction from the tracking system to the object in comparison to the two perpendicular directions. Taking this anisotropy into consideration has the potential to decrease the registration error, and thus, decrease the tracking error.

In Chapter IV we present a new tracking algorithm, which takes the anisotropy created by the tracking system into consideration. The new tracking algorithm consists of two phases: (1) FLE covariance estimation and (2) anisotropic point-based registration. For FLE covariance estimation, we modified an existing estimation algorithm. FLE covariance is estimated from FRE covariance, which can be measured. For this estimation procedure, FLE homogeneity of the fiducial system is assumed. This assumption can be made because for many tracking systems the distance from the tracking system to the object is much larger than the distances between the fiducials. For FLE covariance estimation we also tested a method based on the equation relating FRE to FLE derived in Chapter II. This method is more general than the existing method (with or

without modification) and we found that it gives the same results. Because it is more general, it requires more execution time. We therefore used the modified existing algorithm in our validation experiments.

For the second phase of the tracking algorithm – anisotropic point-based registration – we used the iterative algorithm examined in Chapter III. First, we find the FLE covariance using the modified version of the existing estimation algorithm. Then we calculate the ideal weighting for the fiducials using the estimated FLE covariance. Finally, we register the localized fiducials using calculated weights in the iterative anisotropic algorithm.

We performed an experiment on real data to compare the new tracking algorithm with the standard tracking algorithm. The experiment showed that the new tracking algorithm notably improves the accuracy. Moreover, the new tracking algorithm works fast enough to be used in tracking systems in real-time.

Building a medical application for researching the effect of point-based registration in presence of anisotropic FLE (Chapter V)

In Chapter V we describe a new study in which we performed robotic mastoidectomy [78]. Mastoidectomy is a process of removing mastoid – part of the temporal bone. Nowadays this medical operation is performed only manually. The goal of the study was to use a robotic system for performing this operation. That would allow performing it more accurately and with a guarantee that no critical structures, such as the facial nerve, would be damaged.

During the robotic mastoidectomy, one fiducial frame is attached to a patient, and two more fiducial frames are attached to the robot. The three fiducial frames are tracked by the tracking system. Because the robotic system's error is very low, the major source of the error is the tracking error. In the study we used a tracking system that produces a high anisotropy in physical space. Thus, the presented robotic system should serve as an ideal platform for studying the effect of anisotropic FLE on the accuracy of a medical operation.

In Chapter VI we give conclusions and suggest some directions for future work.

9.   References

[1]   Hajnal, J.V., Hill, D.L.G., and Hawkes, D.J. Medical image registration, The Biomedical Engineering Series, M.R. Neuman, CRC Press.

[2]   Fitzpatrick, J.M., Hill, D.L.G., and Maurer, C.R. Registration. Medical Image Processing, Volume II of the Handbook of Medical Imaging, M. Sonka and J.M. Fitzpatrick, eds., SPIE Press, 447-513, 2000.

[3]   Spiegel, E.A., Wycis, H.T., Marks, M. and Lee, A.J. Stereotaxic apparatus for operations on the human brain, Science, 106(2754), 349-350, 1947.

[4]   Hounsfield, G.N. Computerized transverse axial scanning (tomography). 1. Description of system, Br J Radiol, 46(552), 1016-1022, 1973.

[5]   Lauterbur, P.C. Image formation by induced local interactions: examples employing nuclear magnetic resonance, Nature, 242, 190-191, 1973.

[6]   Hornak, J.P. The basics of MRI, http://www.cis.rit.edu/htbooks/mri/.

[7]   Woods, R.P., Cherry, S.R., and Mazziotta, J.C. Rapid automated algorithm for aligning and reslicing PET images, J. Comput. Assist. Tomogr., 16, 620-622, 1992.

[8]   Hajnal, J.V., Saeed, N., Soar, E.J., Oatridge, A., Young, I.R., and Bydder, G.M. A registration and interpolation procedure for subvoxel matching of serially acquired MR images, J. Comput. Assist. Tomogr., 19, 289-296, 1995.

[9]   Viola, P. and Wells, W. Alignment by maximization of mutual information, Proc. 5th Int. Conf. Computer Vision, 16-23, 1995.

[10] Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., and Marchal, G. Automated multi-modality image registration based on information theory, Information Processing in Medical Imaging, Bizais, Y., Barillot, C., Di Paola, R., Eds., Dordrecht: Kluwer Academic, 263-274, 1995.

[11] Studholme, C., Hill, D.L.G., and Hawkes, D.J. Automated 3D registration of MR and CT images of the head, Med. Image Anal., 1, 163-175, 1996.

[12] Studholme, C., Hill, D.L.G., and Hawkes, D.J. Automated 3D registration of MR and PET brain images by multi-resolution optimization of voxel similarity measures, Med. Phys., 24, 25-35, 1997.

[13] Maintz, J.B.A. and Viergever, M.A. A survey of medical image registration, Med. Image Anal., 2, 1-36, 1998.

[14] Maurer, C.R., Jr. and Fitzpatrick, J.M. A review of medical image registration. Interactive Image-Guided Neurosurgery, R. J. Maciunas, Ed., American Association of Neurological Surgeons, Park Ridge, IL, 17-44, 1993.

[15] Van den Elsen, P.A. Multimodality matching of brain images, Ph.D. thesis, Utrecht University, Utrecht, The Netherlands, 1993.

[16] Fright, W.R. and Linney, A.D. Registration of 3-D head surfaces using multiple landmarks, IEEE Transactions on Medical Imaging, 12(3), 515--520, 1993.

[17] Hill, D.L.G., Hawkes, D.J., Gleeson, M.J., Cox, T.C.S., Strong, A.J., Wong, W.-L., Ruff, C.F., Kitchen, N.D., Thomas, D.G.T., Crossman, J.E., Studholme, C., Gandhe, A.J., Green, S.E.M., and Robinson, G.P. Accurate frameless registration of MR and CT images of the head: Applications in surgery and radiotherapy planning, Radiology, 191, 447–454, 1994.

[18] Van Den Elsen, P.A. and Viergever, M.A. Marker guided registration of electromagnetic dipole data with tomographic images, Information Processing in Medical Imaging, A.C.F. Colchester and D.J. Hawkes, eds., 142--153, 1991.

[19] Maurer, C.R.,Jr., Aboutanos, G.B., Dawant, B.M., Maciunas, R.J., and Fitzpatrick, J.M. Registration of 3-D images using weighted geometrical features, IEEE Transactions on Medical Imaging, 15, 836–849, 1996.

[20] Maurer, C.R., Jr., Fitzpatrick, J.M., Wang, M.Y., Galloway, Jr., R.L., Maciunas, R.J., and Allen, G.S. Registration of head volume images using implantable fiducial markers, IEEE Transactions on Medical Imaging, 16, 447-462,1997.

[21] Maurer, C.R., Jr., Maciunas, R.J., and Fitzpatrick, J.M. Registration of head CT images to physical space using a weighted combination of points and surfaces, IEEE Transactions on Medical Imaging, 17, 753–761, 1998.

[22] Green, B.F. The orthogonal approximation of an oblique structure in factor analysis, Psychometrika, 17, 429–440, 1952.

[23] Schoenemann, P.H. A generalized solution of the orthogonal Procrustes problem, Psychometrika, 31, 1–10, 1966.

[24] Sibson, R. Studies in the robustness of multidimensional scaling: Procrustes statistics, J.R. Statist. Soc. B, 40, 234–238, 1978.

[25] Arun, K.S., Huang, T.S., and Blostein, S.D. Least-squares fitting of two 3-D point sets, IEEE Trans. Pattern Anal. Mach. Intell., 9, 698–700, 1987.

[26] Horn, B.K.P. Closed-form solution of absolute orientation using unit quaternions, J. Opt. Soc. Amer. A, 4, 629–642, 1987.

[27] Horn, B.K.P., Hilden, H.M., and Negahdaripour, S. Closed-form solution of absolute orientation using orthonormal matrices, J. Opt. Soc. Amer. A, 5, 1127–1135, 1988.

[28] Pelizzari, C.A., Chen, G.T.Y., Spelbring, D.R., Weichselbaum, R.R., and Chen, C.T. Accurate three-dimensional registration of CT, PET, and/or MR images of the brain, J. Comput. Assist. Tomogr., 13, 20–26, 1989.

[29] Besl, P.J. and McKay, N.D. A method for registration of 3-D shapes, IEEE Trans. Pattern Anal. Mach. Intell., 14, 239–256, 1992.

[30] Lemoine, D., Liegeard, D., Lussot, E., and Barillot, C. Multimodal registration system for the fusion of MRI, CT, MEG, and 3D or stereotactic angiographic data, Medical Imaging: Image Capture, Formatting, and Display, Proc. SPIE 2164, 46-56, 1994.

[31] Jiang, H., Holton, K.S., and Robb, R.A. Image registration of multimodality 3-D medical images by chamfer matching, Biomedical Image Processing and Three-Dimensional Microscopy, Proc. SPIE 1660, 356-366, 1992.

[32] Jiang, H., Robb, R.A., and Holton, K.S. A new approach to 3-D registration of multimodality medical images by surface matching, Visualization in Biomedical Computing, Proc. SPIE 1808, 196–213, 1992.

[33] Collignon, A., Vandermeulen, D., Suetens, P., and Marchal, G. Registration of 3D multi-modality medical images using surfaces and point landmarks, Pattern Recognit. Lett., 15, 461–467, 1994.

[34] Mascott, C.R., Sol, J.C., Bousquet, P., Lagarrigue, J., Lazorthes, Y., and Lauwers-Cances, V. Quantification of true in vivo (application) accuracy in cranial image-guided surgery: influence of mode of patient registration, Neurosurgery, 59(1 Suppl 1), ONS146-156; discussion ONS146-156, 2006.

[35] Metzger, M.C. Comparison of 4 registration strategies for computer-aided maxillofacial surgery, Otolaryngology-Head and Neck Surgery, 137, 93-00, 2007.

[36] Woerdeman, P.A., Willems, P.W., Noordmans, H.J., Tulleken, C.A., and Van Der Sprenkel, J.W. Application accuracy in frameless image-guided neurosurgery: a comparison study of three patient-to-image registration methods, J. Neurosurg., 106(6), 1012-1016, 2007.

[37] Viola, P.A. Alignment by maximization of mutual information, Ph.D. thesis, Massachusetts Institute of Technology, 1995.

[38] Woods, R.P., Mazziotta, J.C., and Cherry, S.R. MRI-PET registration with automated algorithm, J. Comput. Assist. Tomogr., 17, 536–546, 1993.

[39] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P. Multimodality image registration by maximization of mutual information, IEEE Transactions on Medical Imaging, 16, 187–198, 1997.

[40] Wells, W.M., III, Viola, P., Atsumi, H., Nakajima, S., and Kikinis, R. Multi-modal volume registration by maximization of mutual information, Med. Image Anal., 1, 35–51, 1996.

[41] Studholme, C., Hill, D.L.G., and Hawkes, D.J. An overlap invariant entropy measure of 3D medical image alignment, Pattern Recognition, 32, 71–86, 1999.

[42] Clarke, R.H. and Horsley, V. On a method of investigating the deep ganglia and tracts of the central nervous system (cerebellum), Br. Med. J., 2, 1799-1800, 1906.

[43] Horsley, V. and Clarke, R.H. The structure and function of the cerebellum examined by a new method, Brain, 31, 45-124, 1908.

[44] Brown, R. A stereotactic head frame for use with CT body scanners, Investigative Radiology, 14(4), 300-304, 1979.

[45] Kall, B.A., Kelly, P.J., and Goerss, S.J. Interactive stereotactic surgical system for the removal of intracranial tumors utilizing the CO2-laser and CT-derived database, IEEE Transactions on Biomedical Engineering, 32(2), 112-116, 1985.

[46] Kelly, P.J. Computer-assisted stereotaxis - new approaches for the management of intracranial intraaxial tumors, Neurology, 36(4), 535-541, 1986.

[47] Roberts, D.W., Strohbehn, J.W., Hatch, J.F., Murray, W., and Kettenberger, H. A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope, J. Neurosurg., 65(4), 545-549, 1986.

[48] Allen, G.S. Method and apparatus for providing related images over time of a portion of the anatomy using fiducial implants, In Office, U.S.P.a.T., ed., 1991.

[49] Labadie, R.F., Fenlon, M., and Devikalp, H. Image-guided otologic surgery, Lemke, H.U., Vannier, M.W., Inamura, K., Farman, A.G., Doi, K., Reiber, J.H.C., eds. Computer Assisted Radiology and Congress and Exhibition. Amsterdam, The Netherlands: Elsevier Science, 627–32, 2003.

[50] Lenarz, T. and Heermann, R. Image-guided and computer-aided surgery in otology and neurology: is there already a need for it? Otology and Neurology, 20(2), 143-144, 1999.

[51] Fried, M.P., Moharir, V.M., Shin, J. Comparison of endoscopic sinus surgery with and without image guidance, Am J Rhinol.,16(4):193-7, 2002.

[52] Sadoukhi, B. Image-guided surgery, http://emedicine.medscape.com/article/875524-overview

[53] Labadie, R.F., Noble, J.H., Dawant, B.M. Clinical validation of percutaneous cochlear implant surgery: Initial report, Laryngoscope, 118(6), 1031–1039, 2008.

[54] Schlaier, J., Warnat, J., and Brawanski, A. Registration accuracy and practicability of laser-directed surface matching, Comput Aided Surg., 7, 284-90, 2002.

[55] McCrory, J.J., Fitzpatrick, J.M., Willcott, M.R., Maciunas, R.J., Maurer, C.R., Jr., and Allen, G.S. Fiducial Marker, Patent No. 6,333,971, United States Patent and Trademark Office, 2001.

[56] Fitzpatrick, J.M. and McCrory, J.J. Localization Cap for Fiducial Markers, Patent No. 5,730,130, United States Patent and Trademark Office, 1998.

[57] Warren, F.M., Balachandran, R., Fitzpatrick, J.M., and Labadie, R.F. Percutaneous cochlear access using bone-mounted, customized drill guides: demonstration of concept in vitro, Otology and Neurotology, 28(3), 325-329, 2007.

[58] Labadie, R.F., Mitchell, J., Balachandran, R., and Fitzpatrick, J.M. Customized, rapid-production microstereotactic table for surgical targeting: description of concept and in-vitro validation, Int. J. Comp. Assist. Radiol. and Surg., 4(3), 273-280, 2009.

[59] Schlöndorff, G., Mösges, R., Meyer-Ebrech,t D., Krybus, W., and Adams, L. CAS (computer assisted surgery): a new procedure in head and neck surgery [in German], HNO, 37, 187-190, 1989.

[60] Parmar, A ., Grant D.G., and Loizou, P .Robotic surgery in ear nose and throat, Eur. Arch. Otorhinolaryngol., 2009.

[61] McLeod, I.K. and Melder, P.C. Da Vinci robot-assisted excision of a vallecular cyst: a case report, Ear Nose Throat J., 84, 170–172, 2005.

[62] Moghari, M.H. and Abolmaesumi, P. Point-based rigid-body registration using an unscented kalman filter, IEEE Transactions on Medical Imaging, 26(12), 1708-1728, 2007.

[63] Koschat, M.A. and Swayne, D.F. A weighted Procrustes criterion, Psychometrika, , 56(2), 229-239 1991.

[64] Batchelor, P.G. and Fitzpatrick, J.M. A study of the anisotropically weighted Procrustes problem, IEEE Workshop on Mathematical Methods in Biomedical Image AnalysisHilton Head, SC, 2000.

[65] Balachandran, R., and Fitpatrick, J.M. Iterative solution for rigid-body point-based registration with anisotropic weighting. In Miga, M.I. and Wong, K.H., eds. SPIE Medical Imaging 2009, SPIE, Lake Buena Vista, FL, 2009.

[66] Langron, S.P. and Collins, A.J. Perturbation theory for generalized Procrustes analysis, J. R. Statist. Soc. B, 47, 277–284, 1985.

[67] Sibson, R. Studies in the robustness of multidimensional scaling: Perturbational analysis of classical scaling, J. R. Statist. Soc. B, 41, 217-229, 1979.

[68] Fitzpatrick, J.M., West, J.M., and Maurer, C.R., Jr. Predicting error in rigid-body, point-based registration, IEEE Transactions on Medical Imaging, 17, 694-702, 1998.

[69] Fitzpatrick, J.M. and West, J.B. The distribution of target error in rigid-body, point-based registration, IEEE Transactions on Medical Imaging, 917-927, 2001.

[70] Wiles, A.D., Likholyot, A., Frantz, D.D., and Peters, T.M. A statistical model for point-based target registration error with anisotropic fiducial localizer error, IEEE Transactions on Medical Imaging, 27(3), 378–390, 2007.

[71] Moghari, M. and Abolmaesumi, P. Distribution of target registration error for anisotropic and inhomogeneous fiducial localization error, IEEE Transactions on Medical Imaging, 28(6), 799-813, 2009.

[72] Moghari, M. and Abolmaesumi, P. Distribution of fiducial registration error in rigid-body point-based registration, IEEE Transaction on Medical Imaging, 28(11), 1791-1801, 2009.

[73] Wiles, A.D. Navigation accuracy of image-guided interventional systems, Dissertation, Department of Medical Biophysics, University of Western Ontario, London, Ontario, Canada, 2009.

[74] Danilchenko, A., and Fitzpatrick, J.M. General approach to error prediction in point registration, IEEE Transactions on Medical Imaging, 30(3), 679-693, 2011.

[75] Danilchenko, A., and Fitzpatrick, J.M. General approach to error prediction in point registration, SPIE Medical Imaging 2010, San Diego, CA, SPIE 7625, 2010.

[76] Danilchenko, A., Wiles, A.D., and Fitzpatrick, J.M. Improved method for point-based tracking, MICCAI 2010 Lecture Notes in Computer Science, 6363, 587-594, 2010.

[77] Wiles, A.D., and Peters, T.M. Real-time estimation of FLE statistics for 3D tracking with point-based registration, IEEE Transactions on Medical Imaging, 28(9), 1384-1398, 2009.

[78] Danilchenko, A., Balachandran, R., Toennies, J.L., Baron, S., Munske, B., Webster III, R.J., and Labadie, R.F. Robotic mastoidectomy, Otology & Neurotology, 32(1), 11-16, 2001.

[79] Nathoo, N., Cavusoglu, M.C., Vogelbaum, M.A., and Barnett, G.H. In touch with robotics: neurosurgery for the future, Neurosurgery, 56(3), 421-433, 2005.

[80] Arambula Casio, F., and Padilla Castaneda, M.A. Computer assisted surgery, AIP conf. proc., 682, 38-45, 2003.

GENERAL APPROACH TO FIRST-ORDER ERROR PREDICTION IN RIGID POINT REGISTRATION

This chapter is adapted from the following paper:

- Danilchenko, A. and Fitzpatrick, J.M. General approach to first-order error prediction in rigid point registration, IEEE Transactions on Medical Imaging, 30(3), 679-693, 2011.

1. Introduction

In 2009, a new method of first-order analysis of the point-registration problem was introduced [1]. The method was developed to facilitate the statistical analysis of a phenomenon that had been observed years earlier, first by Steinmeier using phantoms [2] and more recently by Woerdeman in-vivo [3]. 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 [1] 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 with Section 2 by providing background on and precise definitions of error in these navigation systems. In Section 3 we present our new approach and derive independence and the new statistical expressions. In Sections 4 and 5 we compare our results to previously published results and to results from computer simulations. In Sections 6 and 7 we discuss our theory and its results and present our conclusions.

2. 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 non-rigidity 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 [4]:

$$\left\langle \text{FRE}^2 \right\rangle = \left(1 - 2/N\right)\left\langle \text{FLE}^2 \right\rangle, \tag{1}$$

where $\langle \cdot \rangle$ means "expected value of". Expressions for FRE for each individual fiducial have also been derived [9]. 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 fiducial.[1] The target error $\mathbf{TRE}(r)$ 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 [5], [6],

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

where d_k is the distance of r from k-th principal axis of the fiducial configuration, f_k^2 is the mean of the squared distances of the fiducials from that axis. Eq. (1) can be used to estimate the distribution of FLE from measurements of FRE for many cases. With this estimate available, Eq. (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 [7]. Eq. (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 Eq. (2) have recently been derived by Wiles [8], [22] and Moghari [10].

---

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

Since both $\langle \text{FRE}^2 \rangle$ and $\langle |\mathbf{TRE}(r)|^2 \rangle$ are proportional to $\langle \text{FLE}^2 \rangle$, 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 Figure 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. Figure 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.

Figure 1 shows schematically that, despite the fact that $\langle \text{FRE}^2 \rangle$ and $\langle |\mathbf{TRE}(r)|^2 \rangle$ are related functionally through Eqs. (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 fluctuations[2] 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

---

[2] also called "deviations"

of a linear model of the problem of rigid point registration. Recently a new approach [1]

for such an analysis was developed for the special case of isotropic, homogeneous FLE

with uniform weighting (i.e., unweighted). 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
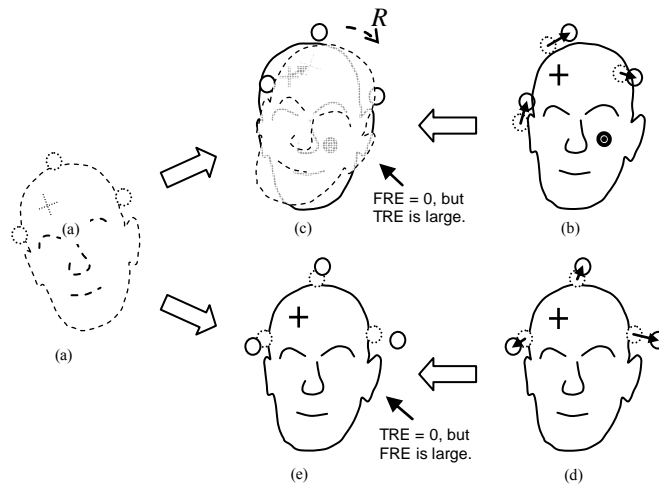
Figure 1. FRE and TRE for two example registrations. (a) Image space showing patient (dashed outline), three fiducials (dotted circles), and an anatomical target (dotted cross). For simplicity localization errors are zero. (b) Physical space showing a set of fiducial localization errors. Arrows show 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.

formulas for their respective covariances that are more general than any previous expressions.

3. 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.

The registration problem

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

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

where FRE is the weighted fiducial registration error x_i , y_i are 3-by-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-by-3 non-singular weighting matrix, which may be a function of R (see the discussion of "ideal weighting" at the beginning of "Statistical independence of FRE and TRE for ideal weighting" in Section 3 (page 52)). Clearly the solution, R, t, obtained by minimizing FRE^2 in Eq. (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 Eq. (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-by-3 identity matrix, and for this weighting FRE equals w 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 $\sum_1^N |W_i|^2 = \sum_1^N \text{trace}(W_i^t W_i) = 3$. With this stipulation, uniform weighting requires that $W_i = N^{-1/2} I$, and FRE recaptures its meaning as the unweighted root-mean square of the fiducial fitting errors and satisfies Eq. (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:

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

$R^{(0)}, t^{(0)}$ is the transformation that is sought by the minimization of FRE in Eq. (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 $\Sigma_i^{(1)}$ and $\Sigma_i^{(2)}$, 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 Eq. (3) as

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

where $\Delta \xi_i = \Delta y_i - R \Delta x_i$ is the "two-space" localization error vector for point-pair i, and its elements are normally distributed with zero means and covariances $R \Sigma_i^{(1)} R^t + \Sigma_i^{(2)}$. We define $\Sigma_i = R \Sigma_i^{(1)} R^t + \Sigma_i^{(2)}$ 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:

$$\mathbf{TRE}(r) = Rr + t - \left(R^{(0)} r + t^{(0)} \right). \tag{6}$$

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 [4]. He used this approach to investigate FRE; we and others have used it to investigate TRE [5], [6], [8], [10] and FRE [9], [18]. 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 $3N$-by-6 matrix, e is a $3N$-by-1 vector, and q is a 6-by-1 vector whose elements are $\Delta\theta_1, \Delta\theta_2, \Delta\theta_3, \Delta t_1, \Delta t_2, \Delta t_3$, where $\Delta\theta_k$ is an angle of rotation (in radians) about axis k, and Δt_k is a translation along that axis.

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

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

where we have used Eq. (4) and have defined

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

and

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

40

Using Eq. (7) in Eq. (3) gives us

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

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

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

where we have used the orthogonality of rotation matrices to go from the first line to the second. From Eq. (11), we see that, if the elements of $\Delta R^t \Delta R$ are indeed negligible in comparison to the elements of ΔR, then $\Delta R^t + \Delta R \approx 0$, which means that ΔR is antisymmetric, i.e., $\Delta R^t = -\Delta R$. Thus, the non-linear 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 [4], the approaches of the subsequent work on the statistics of TRE [5], [6], [8], 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 $\Sigma_i^{(1)}$ and all $\Sigma_i^{(2)}$, whereas for the non-linear version, we can show agreement only for those cases that are simulated. Using our definitions of ΔR and Δt in Eq. (6) gives

$$\mathbf{TRE}(r) = \left(RR^{(o)^t} - I \right) R^{(0)} r + t - t^{(0)} = \Delta R r^{(0)} + \Delta t , \tag{12}$$

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 $\Sigma_i = R^{(o)} \Sigma_i^{(1)} R^{(o)^t} + \Sigma_i^{(2)}$.

Some useful results from linear algebra

The minimization of Eq. (10), subject to $\Delta R^t = -\Delta R$ with $\Delta R R^{(0)} \Delta x_i$ 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, $\Delta\theta_1, \Delta\theta_2, \Delta\theta_3, \Delta t_1, \Delta t_2, \Delta t_3$:

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

where

$$\Delta R = \begin{bmatrix} 0 & -\Delta\theta_3 & \Delta\theta_2 \\ \Delta\theta_3 & 0 & -\Delta\theta_1 \\ -\Delta\theta_2 & \Delta\theta_1 & 0 \end{bmatrix} , \tag{14}$$

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

The meaning of ΔR can be understood by adding I to both sides of its definition in Eq. (8) and multiplying on the right by $R^{(0)}$. The result, $\left(I + \Delta R\right)R^{(0)} = R$ shows that the rotation matrix that we seek by minimizing Eq. (3) is the equivalent to the true rotation followed by a "rotation" $I + \Delta R$. Thus, $I + \Delta R$ approximates the rotation away from the truth caused by fiducial localization error. If localization error were zero, then ΔR would also equal zero. Our assumption that localization error is merely small results in the approximation, $\Delta R^t = -\Delta R$. The displacement Δr of a point r caused by that small rotation is $\Delta r = \left(I + \Delta R\right)r - r = \Delta R r$, which, because ΔR^t is equal to $-\Delta R$, can be written in the form of the familiar cross product for small rotations, $\Delta r = \Delta R r = \Delta\theta \times r$, where $\Delta\theta = \left(\Delta\theta_1, \Delta\theta_2, \Delta\theta_3\right)$. This cross product produces the exact expression for Δr in the limit of small rotation angle $\left|\Delta\theta\right| = \left(\Delta\theta_1^2 + \Delta\theta_2^2 + \Delta\theta_3^2\right)^{1/2}$. Eq. (14) can also

be derived by representing the exact rotation $I + \Delta R$ in terms of Euler angles, expanding each sine and each cosine in that representation via a Taylor's series in the angles $\Delta\theta_1$, $\Delta\theta_2$, and $\Delta\theta_3$, 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 ΔR in Eq. (14).

We now define the 3N-by-6 matrix C,

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

where

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

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

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

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

44

$$Cq = e. \tag{18}$$

With Eq. (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 [14] of C,

$$C = U \Lambda V^t, \tag{19}$$

where U is a $3N$-by-$3N$ orthogonal matrix, Λ is a $3N$-by-6 diagonal matrix whose diagonal elements Λ_{ii} are the singular values of C, and V is a 6-by-6 orthogonal matrix.[3] We will require that the six singular values Λ_{ii} of C all be nonzero. Since all the W_i are non-singular, this is equivalent to the requirement that the columns of C be independent when $W_i = I$. We can prove that they are independent via proof by contradiction. Thus, we assume instead that the columns of C are linearly dependent, which means that for some set of six constants, $a_1, a_2, ..., a_6$, $\sum_{i=1}^{6} a_i C_i = 0$, 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

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

---

[3] For convenience, we have augmented both U and Λ relative to the decomposition described in reference [14] by adding $3N - 6$ columns to U so as to make it orthogonal and adding $3N - 6$ rows of zeros to Λ.

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-by-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, all N points $x_i^{(0)}$ must be collinear. We note from "The registration problem" of Section 3 above (page 37) 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 Eq. (18) is given as follows [14],

$$q_{\min} = C^+ e, \tag{21}$$

where

$$C^+ = V \Lambda^+ U^t, \tag{22}$$

and Λ^+ is a 6-by-$3N$ 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 non-zero. The elements of the $3N$-by-1 vector

$$f_{\min} = e - C q_{\min}, \tag{23}$$

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

$$f_{\min} = \left(I - CC^{+}\right)e,$$

(24)

where I is the 3N-by-3N identity matrix. Using Eqs. (19) and (22) in Eq. (24), we have

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

(25)

For $i,j = 1, \ldots, 3N$, we have $\left(\Lambda\Lambda^{+}\right)_{ij} = \sum_{k=1}^{6}\Lambda_{ik}\Lambda_{kj}^{+} = \sum_{k=1}^{6}\Lambda_{ik}\Lambda_{kj}^{-1}$, 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 Eq. (25), and defining

$$\tilde{e} = U^{t}e$$

(26)

we find that

$$\left(f_{\min}\right)_{k} = \sum_{j=1}^{6}U_{kj}\cdot 0\cdot\tilde{e}_{j} + \sum_{j=7}^{3N}U_{kj}\cdot 1\cdot\tilde{e}_{j} = \sum_{j=7}^{3N}U_{kj}\tilde{e}_{j} \ .$$

(27)

47

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

FRE in terms of $\tilde{e}$

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

TRE in terms of $\tilde{e}$

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

$$\delta = Dq \tag{28}$$

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

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

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

$$\delta_{\min} = Dq_{\min} = DC^+e = DV\Lambda^+U^te = DV\Lambda^+\tilde{e} \tag{30}$$

Thus,

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

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

General formulation for the covariances of FRE and TRE

Eq. (31) can be employed to derive a general expression for Σ_{FRE}, the covariance of FRE, and Σ_{TRE}, the covariance of TRE. First, we define two 3N-by-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-by-3 blocks on the diagonal. Specifically, $\Sigma_{3(i-1)+j,3(i-1)+j'} = \left(\Sigma_i\right)_{jj'}$ and $W_{3(i-1)+j,3(i-1)+j'} = \left(W_i\right)_{jj'}$, for $i = 1...N$, $j = 1,2,3$, $j' = 1,2,3$, and all other elements of each matrix equal zero. With these definitions it can be seen from the definition of e given by Eq. (17) that

$$\left\langle ee^{t} \right\rangle = W\Sigma W^{t}. \tag{32}$$

Our expression for Σ_{FRE} can be gotten by noting that $\left(\Sigma_{\mathrm{FRE}}\right)_{kl} = \left\langle \left(f_{\min}\right)_{k} \left(f_{\min}^{t}\right)_{l} \right\rangle$

and then using Eq. (27), followed by Eqs. (26) and (32), to get

$$\left(\Sigma_{\mathrm{FRE}}\right)_{kl} = \sum_{j,j'=7}^{3N} U_{kj}U_{lj'}\left\langle \tilde{e}_{j}\tilde{e}_{j'} \right\rangle = \sum_{j,j'=7}^{3N} U_{kj}U_{lj'}\left(U^{t}W\Sigma W^{t}U\right)_{jj'}. \tag{33}$$

The summation over three diagonal elements corresponding to the same marker

provides an expression for an individual weighted FRE of that marker:

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

The trace of Σ_{FRE} provides a general expression for $\left\langle \mathrm{FRE}^{2} \right\rangle$:

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

The expressions in Eqs. (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 $\left(W^{-1}U\right)_{kj}$

and $\left(W^{-1}U\right)_{kj'}$ in each of these equations, resulting in

$$\left(\Sigma_{\text{fre}}\right)_{kl} = \sum_{j,j'=7}^{3N} \left(W^{-1}U\right)_{kj}\left(W^{-1}U\right)_{lj'}\left(U'W\Sigma W'U\right)_{jj'},\tag{36}$$

$$\left\langle \text{fre}_i^2 \right\rangle = \sum_{k=3i-2}^{3i}\sum_{j,j'=7}^{3N} \left(W^{-1}U\right)_{kj}\left(W^{-1}U\right)_{kj'}\left(U'W\Sigma W'U\right)_{jj'},\tag{37}$$

and

$$\left\langle \text{fre}^2 \right\rangle = \sum_{j,j'=7}^{3N} \left(U'W^{-2}U\right)_{jj'}\left(U'W\Sigma W'U\right)_{jj'}.\tag{38}$$

For isotropic weighting, *i.e.* $W_i = w_i I$, it can with some manipulation be shown that $\left\langle \text{FRE}_i^2 \right\rangle = w_i \left\langle \text{fre}_i^2 \right\rangle$.

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

$$\left(\Sigma_{\text{TRE}}\right)_{kl} = \left\langle \left(\delta_{\min}\right)_k \left(\delta_{\min}\right)_l \right\rangle = \sum_{j,j'=1}^{6} \left(DV\right)_{kj}\left(DV\right)_{lj'}\frac{\left\langle \tilde{e}_j \tilde{e}_{j'} \right\rangle}{\Lambda_{jj}\Lambda_{j'j'}}.\tag{39}$$

By applying Eqs. (26) and (32) to $\left\langle \tilde{e}_j \tilde{e}_{j'} \right\rangle$, we achieve the following result,

$$\left(\Sigma_{\mathrm{TRE}}\right)_{kl} = \sum_{j,j'=1}^{6} \left(DV\right)_{kj} \left(DV\right)_{lj'} \frac{\left(U^{t}W\Sigma W^{t}U\right)_{jj'}}{\Lambda_{jj}\Lambda_{j'j'}} . \tag{40}$$

The trace of Σ_{TRE} provides a general expression for $\left\langle \mathrm{TRE}^{2} \right\rangle$:

$$\left\langle \mathrm{TRE}^{2} \right\rangle = \sum_{k=1}^{3} \left(\Sigma_{\mathrm{TRE}}\right)_{kk}$$

$$= \sum_{k=1}^{3} \sum_{j,j'=1}^{6} \left(DV\right)_{kj} \left(DV\right)_{kj'} \frac{\left(U^{t}W\Sigma W^{t}U\right)_{jj'}}{\Lambda_{jj}\Lambda_{j'j'}} . \tag{41}$$

It should be noted that Eqs. (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 "Computer code to implement the derived formulas" in Section 8 (page 80) an implementation of Eqs. (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 two-dimensional case as well with only minor changes to accommodate the reduced dimensionality. The details are given in "Two dimensional case" in Section 8 (page 76).

Statistical independence of FRE and TRE for ideal weighting

In Figure 1, we show examples for which FRE is zero and TRE is nonzero. It can easily be seen from Eq. (5) that the number of such examples is infinite. Indeed, whenever there exists a rigid transformation R,t for which $Rx_i + t = y_i + \Delta\xi_i$ for all i,

then FRE is zero, but unless all the $\Delta \xi_i$ 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 Δx_i are equal to the same nonzero vector v_x, and all the Δy_i are equal to some other nonzero vector $v_y \neq v_x$. As a result, each $\Delta \xi_i = v_y - v_x$, the rigid transformation that minimizes FRE^2 is, $R = R^{(0)}$, $t = t^{(0)} + v_y - v_x$, and $\mathbf{TRE}(r) = v_y - v_x \neq 0$. Substituting these values into Eq. (5) reveals that each of the terms in the summation is zero. Thus, each $\mathbf{FRE}_i = 0$. There is likewise 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 Eq. (5) is the identity. One such example is also shown in Figure 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 "Derivation of ideal weighting" in Section 8 (page 70) and show that it has the form $W_i = w \times \Sigma_i^{-1/2} = w \times \left(R\Sigma_i^{(1)} R^t + \Sigma_i^{(2)} \right)^{-1/2}$, where w is a constant, which for normalized weighting is equal to $3^{1/2} \left(\sum_1^N \mathrm{trace}\left(\Sigma_i^{-1}\right) \right)^{-1/2}$. Ideal weighting is not new. It was introduced into point registration as early as 1998 by Ohta and Kanatani [15], and it

was employed by West *et al.* in 2001 [16] and Moghari *et al.* in 2009 [10], [18]. In both of these 2009 papers a first-order registration equation equivalent to Eq. (18) was derived for the special case in which ideal weighting is employed and $\Sigma_i^{(1)} = 0$.[4] As mentioned at the end of "Linearization" in Section 3 (page 42), to first order R can be replaced by $R^{(0)}$ in the expression for Σ_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 Eq. (13). Using this weighting in Eq. (32) gives us

$$\left\langle ee^t \right\rangle = W\Sigma W^t = w^2 \times \Sigma^{-1/2}\Sigma\Sigma^{-1/2} = w^2 I \ . \tag{42}$$

Since $\left\langle ee^t \right\rangle$ is diagonal, the elements e_i are uncorrelated; since they are uncorrelated and also normally distributed, they are statistically independent [1], [17]. Because the diagonal elements equal w^2, the variance of each e_i is w^2. As shown in "Statistical independence of the elements of $\tilde{e}$" in Section 8 (page 73), since U is orthogonal, the elements of $\tilde{e}$ are also independent and normally distributed with variance equal to w^2. In "FRE in terms of $\tilde{e}$" in Section 3 (page 48), we found that each component of **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 $\tilde{e}$, while in "TRE in terms of $\tilde{e}$" in Section 3 (page 48) we found that each component of $\mathbf{TRE}(r)$, which is also a component of $\delta_{\min}$, is a function of only the first 6 elements of $\tilde{e}$. Since the elements of $\tilde{e}$ are independent of each other, any functions of non-overlapping subsets of them are independent each other

---

[4] Eq. (30) of [10] and Eq. (9) of [18].

as well. It is important to note that this conclusion depends crucially on the fact that all $3N$ of the elements of $\tilde{e}$ 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 Eq. (3), is given in the linearized problem by Eq. (10) and can be expressed in terms of $f_{\min}$ as $\mathrm{FRE} = \left(f_{\min}^{t} f_{\min} \right)^{1/2}$ and that the length of the vector $\mathbf{TRE}(r)$ can be expressed in terms of $\delta_{\min}$ as $\left| \mathbf{TRE}(r) \right| = \left(\delta_{\min}^{t} \delta_{\min} \right)^{1/2}$. Since the elements of $f_{\min}$ are statistically independent of those of $\delta_{\min}$, any function of only $f_{\min}$ is statistically independent of any function of only $\delta_{\min}$. As primary examples, FRE is statistically independent of $\left| \mathbf{TRE}(r) \right|$. Since they are independent, they are also uncorrelated. We let $\mathrm{CC}(\mathrm{FRE}, \mathrm{TRE})$ be the correlation coefficient of FRE and $\left| \mathbf{TRE}(r) \right|$,

$$\mathrm{CC}(\mathrm{FRE}, \mathrm{TRE}) = \left\langle \left(\mathrm{FRE} - \mu_{\mathrm{FRE}} \right) \left(\mathrm{TRE} - \mu_{\mathrm{TRE}} \right) \right\rangle \Big/ \left(\sigma_{\mathrm{FRE}} \sigma_{\mathrm{TRE}} \right), \tag{43}$$

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

Two-dimensional space

The results of this section hold not only for three-dimensional space, but for two-dimensional space as well. The details are provided in "Two dimensional case" in Section 8 (page 76) along with the two-dimensional forms of Eqs. (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 $6 \rightarrow 3$.

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 $FRE^2 = f_{min}^t f_{min} = \sum_{k=1}^{3N} \left(f_{min} \right)_k^2$, which by Eq. (27) equals $\sum_{j=7}^{3N} \sum_{l=7}^{3N} \tilde{e}_j \tilde{e}_l \sum_{k=1}^{3N} U_{kl} U_{kj}$, and since U is orthogonal, we have $FRE^2 = \sum_{j=7}^{3N} \tilde{e}_j^2$. As noted in "Statistical independence of FRE and TRE for ideal weighting" in Section 3 (page 52), when ideal weighting is employed, the $\tilde{e}_j$ 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 [21],

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

(44)

where Γ denotes the Gamma function.[5] This probability density is identical to that derived by Sibson for the case of isotropic, homogeneous FLE with no weighting [4]. 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 [18] for fre (instead of FRE). Also, for ideal weighting, since the $\tilde{e}_j$ are independent and normally distributed with variances equal to w^2, we have $\left\langle \text{FRE}^2 \right\rangle = \left(3N-6\right)w^2$, and for normalized weighting,

$$\left\langle \text{FRE}^2 \right\rangle = 3\left(3N-6\right)\Big/ \sum_1^N \text{trace}\left(\Sigma_i^{-1}\right).$$

(45)

For the special case of homogeneous FLE Eq. (45) reduces to $\left\langle \text{FRE}^2 \right\rangle = 3\left(3-6/N\right)\Big/ \sum_k^3 \sigma_k^{-2}$, 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 Eq. (1).

The second result is an expression for $\left\langle \text{TRE}^2 \right\rangle$ for ideal weighting, first derived for the case of isotropic, homogeneous FLE in 1998 [5] and given above by Eq. (2) and

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

derived for general FLE in 2009 [10]. With ideal weighting we find from Eqs. (32) and (42) that $\left(U^t W \Sigma W^t U \right)_{jj'} = \delta_{jj'}$ (i.e., the Kronecker delta). Using this result in Eq. (41) yields

$$\left\langle \mathrm{TRE}^2 \right\rangle = \sum_{k,j=1}^{3,6} \left(DV \right)_{kj}^2 \Big/ \left(\Lambda_{jj}^2 \right). \tag{46}$$

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

The third result is an expression for $\left\langle \mathrm{TRE}^2 \right\rangle$ when FLE is inhomogeneous and anisotropic but no weighting is employed, first derived in 2009 by Wiles first for the homogeneous case [8] and then for the general case [22]. Our expression for this case is gotten simply by removing W from the expression in Eq. (41) and noting that, in the construction of the block-diagonal covariance matrix Σ, the nonzero blocks, Σ_i 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 [6]. We begin by noting that, for this case, having no weighting is equivalent to employing ideal weighting with $w = 1$. We then have $\left(U^t W \Sigma W^t U \right)_{jj'} = \delta_{ij}$, as in the second result above. Using this result in Eq. (39), we find that $\left(\Sigma_{\mathrm{TRE}} \right)_{kl} = \sum_{j,j'=1}^{6} \left(DV \right)_{kj} \left(DV \right)_{lj} \Big/ \Lambda_{jj}^2$. Performing an eigen decomposition yields $\Sigma_{\mathrm{TRE}} = V_{\mathrm{TRE}} \Lambda^{(\mathrm{TRE})} V_{\mathrm{TRE}}^t$. Since Σ_{TRE} is symmetric, V_{TRE} is

orthogonal. If we define $\eta = V_\delta^t \delta_{\min}$, then, since V_{TRE} is orthogonal and the elements of $\tilde{e}$ 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 Eqs. (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 [18]. They reduce in the homogeneous *and* isotropic case to expressions first derived in 2008 [9].

4. Validation

Our theory has three major results, which can be summarized as follows: For arbitrary weighting and arbitrary FLE, (a) Eq. (33) gives the covariance of FRE, (b) Eq. (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 Eqs. (33), (37), (41), and (46) to previously published analytic expressions for special cases. (2) We compared Eqs. (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 on a Dell Latitude D830 with an Intel Centrino Duocore 2.2 GHz CPU with 2 GBytes RAM.

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 $\Sigma_i^{(1)} = V_i^{(1)} \Lambda_i^{(1)} \left(V_i^{(1)} \right)^t$, where the columns of $V_i^{(1)}$ are the principal axes of FLE and the diagonal elements $\left(\Lambda_i^{(1)} \right)_{jj} = \left(\sigma_i^{(1)} \right)_j^2$ are 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 Eq.(4)).

(3) Generate a covariance FLE matrix $\Sigma_i^{(2)}$ for each y_i in a similar manner to that used to define $\Sigma_i^{(1)}$.

(4) Generate a weighting matrix W_i for each point.

(5) Repeat the following steps 10,000 times:

    1. Perturb each fiducial position in each space to produce $x_i + \Delta x$ and $y_i + \Delta y_i$ according to $\Sigma_i^{(1)}$ and $\Sigma_i^{(2)}$, respectively, with random, normally distributed, zero-mean components.

    2. Employ a registration algorithm to find R and t that minimize the exact registration formula of Eq. (3).

3. Calculate $\mathrm{TRE}(r)$ and FRE.

(6) Calculate $\mathrm{CC}(\mathrm{FRE}, \mathrm{TRE})$ for the 10,000 registrations of Step (5).

(7) Perform the chi-square test described in "Chi-square test for dependence" in Section 8 (page 79) 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 $\mathrm{RMS(TRE)} = \left\langle \mathrm{TRE}^2 \right\rangle^{1/2}$ and $\mathrm{RMS(FRE)} = \left\langle \mathrm{FRE}^2 \right\rangle^{1/2}$ using 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 (5), Part (5)2 depends on the weighting being used. For uniform weighting, the method of singular-value decomposition was employed (e.g., Algorithm 8.1 of reference [19]). For ideal weighting, the algorithm of [20] was employed, except that in Step 1 of that algorithm, $W_i = w \times \left(R\Sigma_i^{(1)} R^t + \Sigma_i^{(2)} \right)^{-1/2}$, as given in "Statistical independence of FRE and TRE for ideal weighting" in Section 3 (page 52). 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 (a) our expressions for TRE and FRE and (b) our claims that TRE and FRE are statistically independent and hence uncorrelated. As pointed out in [1], 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 [21]. The latter test is described "Chi-square test for dependence" in Section 8 (page 79). For all tests we chose $p = 0.05$ with Bonferroni correction applied where appropriate. These tests were also used in [1].

5. Results

Comparison to previously published expressions

We compared Eqs. (33), (37), (41), and (46) with previously published and validated analytic expressions for the first three cases and the fifth case in "Comparison with previously published derivations" in Section 3 (page 56). Regarding Eq. (33), as pointed out in "Comparison with previously published derivations" in Section 3, the simplification of that equation produced by the assumption of isotropy and homogeneity for FLE yields Eq. (45), which agrees exactly with the formula for FRE^2 derived by Sibson that case. We compared the outputs of Eq. (41) to the formulas for $\langle \mathrm{TRE}^2 \rangle$ published by Moghari [10] and Wiles [8] for a large set of inputs. Similarly we compared Eq. (37) to the formula for $\langle \mathrm{fre}_i^2 \rangle$ published by Moghari [18]. 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, …, 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 $\Sigma_i^{(1)}$, 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 Eq. (46), with the formulation of Moghari, who provided an analytic formula for the case of isotropic, inhomogeneous FLE with ideal weighting [10], and we compared Eq. (41) with W set to the identity (*i.e.*, no weighting) to the formula given by Wiles for the case of anisotropic, homogeneous FLE with no weighting [8]. We compared Eq. (37) to Moghari's formula for the case of isotropic inhomogeneous FLE with ideal waiting [16]. In every case, our outputs agreed with those of these previously published formulas (maximum difference $= 1.2 \times 10^{-10}$). In summary, our new general formulas agree with existing formulas, each of which is limited to a special case.

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 RMS(FLE) < 1 mm, and large. For the latter case we used RMS(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 hours for 15 values of RMS(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 (9), in which we calculated the CC between the values of RMS(TRE) of Steps (8) and (9), and the CC between the values of RMS(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).

Figure 2 and Figure 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 (Figure 2a). For FLE > 20 mm, correlation is statistically significant but remains negligible (CC < 0.1) even for FLE = 70 mm (Figure 2b). For uniform weighting (*i.e.*, no weighting) a very small correlation is observed, CC ≤ 0.1 that decreases with increasing N (Figure 3a). 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 ("Chi-square test for dependence" in Section 8 (page 79)) rises above a critical

value shown by the horizontal line in Figure 3(b). Such values were observed very rarely, as can be seen by the sub-critical means and, for $N \geq 20$, 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.

(a) (b)
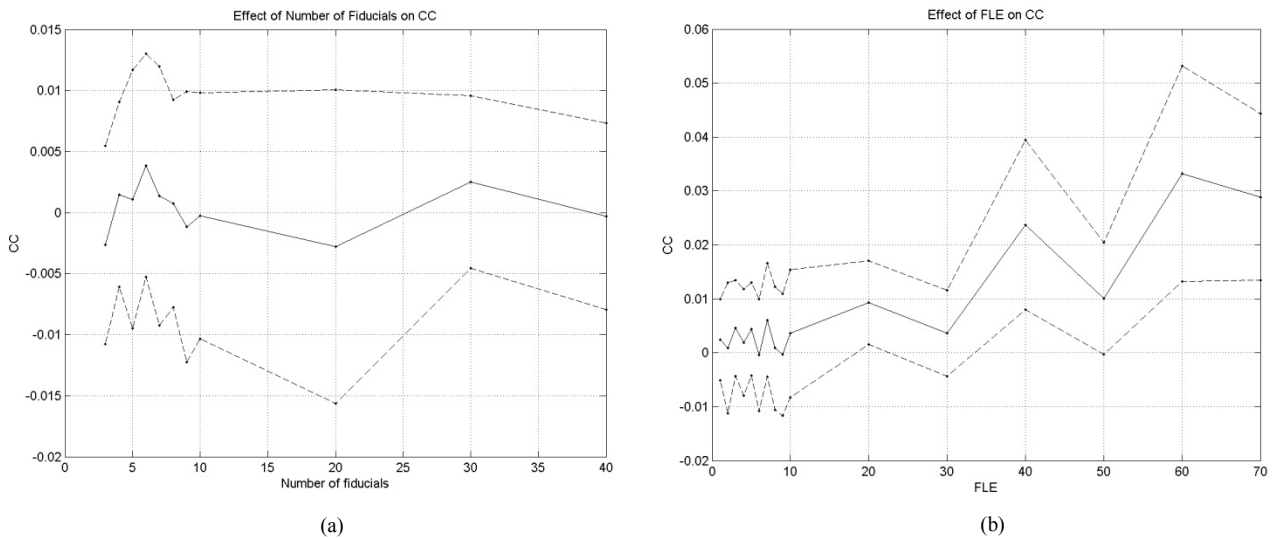
Figure 2. FRE-TRE correlation coefficient CC for ideal weighting. (a) FLE < 1 mm with varying number of fiducials. (b) Varying FLE with four fiducials. In each case, the solid line is the mean of 15 sets of 10,000 registrations for randomly selected fiducial and target positions (see text). The dashed lines are mean ± 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.
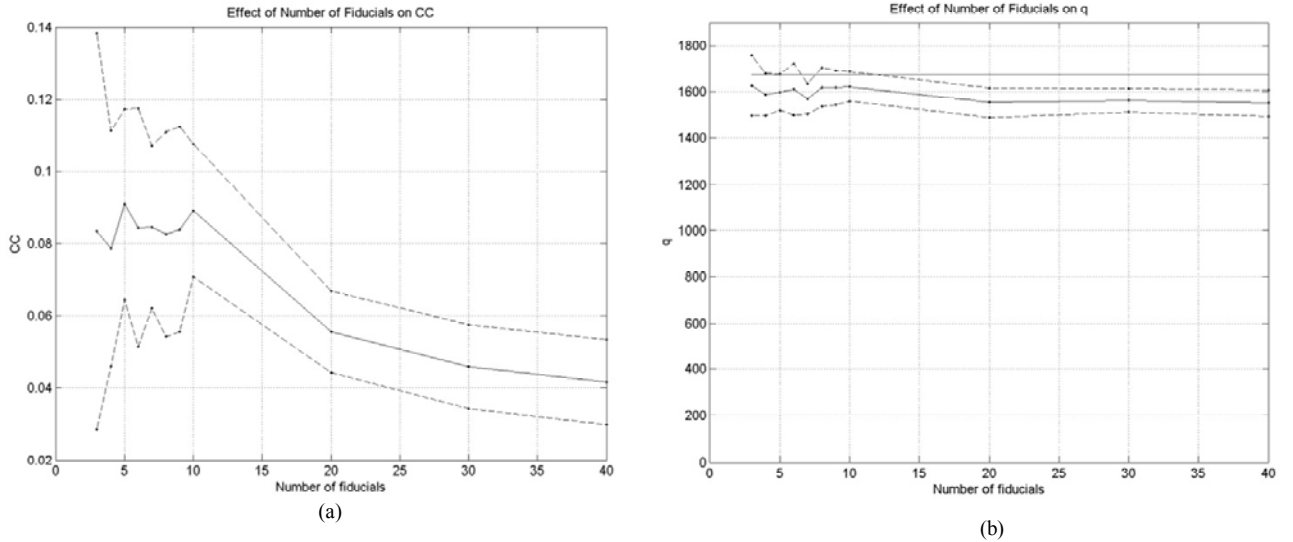
Figure 3. FRE-TRE statistics for FRE < 1 mm when uniform weighting (i.e., no weighting) is employed. (a) Correlation coefficient CC versus number of fiducials. A statistically significant but negligible (< 0.1), correlation is apparent that decreases with increasing number of fiducials. (b) The chi-square statistic q for dependence versus number of fiducials. The dashed lines are mean ± 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.

6. Discussion

We have generalized a method of first-order analysis of the point-registration problem introduced in 2009 [1] 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—Eqs. (33) and (40)—an expression for individual FRE—Eq. (34)—and expressions for $\langle \mathrm{FRE}^2 \rangle$ and $\langle \mathrm{TRE}^2 \rangle$— Eqs. (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 [5], [6], [8], [10], [18]. We have included in "Computer code to implement the derived formulas" in Section 8 (page 80) an implementation of Eqs. (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 [1]: 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 "Statistical independence of FRE and TRE for ideal weighting" in Section 3 (page 52), is not a new concept. It was introduced to point registration by Ohta and Kanatani in 1998 [15] and used by West et al. in 2001 [16] and by Moghari et al. in 2009 [10], [18] 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 Eqs. (1) and (2), from the generalizations of these expressions provided in the excellent work by Wiles [8], [22], Balachandran [9], and Moghari [10], [18], or from further generalizations given in Eqs. (33) through (38), (40), and (41), each of which relates one expectation value to another. Both Eq. (1) and Wiles' generalization of that relationship [22] can be used to estimate the FLE covariance from measurements of FRE obtained from multiple registrations based on independent fiducial localizations. Based on this

estimate, Eq. (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 for any given FLE covariance 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.

To validate our new formulation, we have compared our results to previously published formulas for FRE and TRE and to our own simulations. 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—RMS(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 [1] that estimates of fluctuations of registration accuracy, *i.e.*, TRE, based on the goodness of fit of the fiducials, *e.g.*, FRE, are unreliable.

7. 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 comparisons 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 in the exact case. These results reinforce the message delivered in [1] that for a given localization covariance, whether registration is performed with ideal or 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, bear no statistical relationship, or at most a negligible relationship, to fluctuations of individual measures of registration accuracy, *i.e.*, target registration error. Therefore, while the covariance of FRE can be estimated from multiple registrations, and while valid expressions exist relating the covariances of FRE, FLE, and TRE, any estimation of registration accuracy based on a single fiducial fit should be approached

with extreme caution both by the purveyors of guidance systems and by the practitioners who use them.

8. Appendix

Derivation of ideal weighting

The standard approach to finding the rigid transformation by minimizing Eq. (3) is based on maximum likelihood. Maximum likelihood was introduced into the point registration problem by Ohta and Kanatani in 1998 [15] and used by West et al. in 2001 [16] and by Moghari et al. in 2009 [10], [18]. It has also been invoked for intensity-based registration methods [22]. 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 $\{y_i'\}$, $i = 1,\ldots, N$ 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 $p\left(R,t,\{y_i\} \mid \{x_i'\},\{y_i'\}\right)$. 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 $p\left(R, t, \{y_i\} \mid \{x_i'\}, \{y_i'\}\right)$ with respect to R, t, and $\{y_i\}$ for fixed $\{x_i'\}$ and $\{y_i'\}$. From Bayes' law, we have that

$$
\begin{aligned}
&p\left(R, t, \{y_i\} \mid \{x_i'\}, \{y_i'\}\right) \\
&= p\left(\{x_i'\}, \{y_i'\} \mid R, t, \{y_i\}\right) p\left(R, t\{y_i\}\right) \big/ p\left(\{x_i'\}, \{y_i'\}\right),
\end{aligned}
\tag{47}
$$

and, absent prior knowledge of $p\left(R, t, \{y_i\}\right)$, we neglect its dependence on R, t, and $\{y_i\}$. Furthermore, $p\left(\{x_i'\}, \{y_i'\}\right)$ is fixed because both $\{x_i'\}$ and $\{y_i'\}$ are fixed. Therefore, we may achieve our goal by maximizing the conditional probability, $p\left(\{x_i'\}, \{y_i'\} \mid R, t, \{y_i\}\right)$ 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

$$
\begin{aligned}
&p\left(\{x_i'\}, \{y_i'\} \mid R, t, \{y_i\}\right) \\
&= c \exp\left(-\sum_{i=1}^{N}\left(\Delta x_i^t \Sigma_{xi}^{-1} \Delta x_i + \Delta y_i^t \Sigma_{yi}^{-1} \Delta y_i\right)\right),
\end{aligned}
\tag{48}
$$

where Σ_{xi} and Σ_{yi} are respectively the covariances of Δx_i and Δy_i. It can be seen from our definitions above and from minor manipulation that $\Delta x_i^t \Sigma_{xi}^{-1} \Delta x_i$ equals

$\left(Rx_i' + t - y_i\right)^t S_{xi} \left(Rx_i' + t - y_i\right)$, where $S_{xi} = R\Sigma_{xi}^{-1}R^t$, and that $\Delta y_i^t \Sigma_{yi}^{-1} \Delta y_i$ equals

$\left(y_i' - y_i\right)^t S_{yi} \left(y_i' - y_i\right)$, where $S_{yi} = \Sigma_{yi}^{-1}$. After these substitutions, we have

$p\left(R,t,\{x_i\},\{y_i\} \mid \{x_i'\},\{y_i'\}\right) = c\exp\left(-G\left(R,t,\{x_i'\},\{y_i'\},\{y_i\}\right)\right)$, where

$$G\left(R,t,\{x_i'\},\{y_i'\},\{y_i\}\right) =$$
$$\sum_{i=1}^{N}\left(Rx_i' + t - y_i\right)^t R\Sigma_{xi}^{-1}R^t\left(Rx_i' + t - y_i\right) + \left(y_i' - y_i\right)^t S_{yi}\left(y_i' - y_i\right) \tag{49}$$

Thus, in order to maximize the likelihood, we need to minimize $G\left(R,t,\{x_i'\},\{y_i'\},\{y_i\}\right)$ 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 y_i and setting all $3N$ derivatives to zero. The result is

$$y_i = \left(S_{xi} + S_{yi}\right)^{-1}\left(S_{xi}\left(Rx_i' + t\right) + S_{yi}y_i'\right) . \tag{50}$$

We note that the expression for y_i given in Eq. (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 work. In the case of isotropic, homogeneous FLE, as, for example, in Eq. (2), this expression reduces to $y_i = \left(\left(Rx_i' + t\right) + y_i'\right)/2$, i.e., a simple unweighted average.

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

$$G\left(R,t,\{x_i'\},\{y_i'\},\{y_i\}\right) = \sum_{i=1}^{N}\left(W_i\left(Rx_i'+t-y_i'\right)\right)^t W_i^2\left(Rx_i'+t-y_i'\right), \tag{51}$$

where $W_i = \left(R\Sigma_{xi}R^t + \Sigma_{yi}\right)^{-1/2}$. 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 $W_i = w\times\left(R\Sigma_{xi}R^t + \Sigma_{yi}\right)^{-1/2}$, where w is any constant. For this reason we refer to $W_i = w\times\left(R\Sigma_{xi}R^t + \Sigma_{yi}\right)^{-1/2}$ as "ideal weighting". The right side of Eq. (51) has the same form as the right side of Eq. (3), but the weighting is unspecified in Eq. (3), whereas the derivation above shows that only ideal weighting can achieve maximum likelihood.

Statistical independence of the elements of $\tilde{e}$

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

$$\tilde{e}_i = \sum_{k=1}^{M}V_{ki}e_k \tag{52}$$

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 $\langle . \rangle$, of both sides of Eq. (52):

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

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

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

$$
\begin{aligned}
& \left\langle \left(\tilde{e}_i - \langle \tilde{e}_i \rangle \right) \left(\tilde{e}_j - \langle \tilde{e}_j \rangle \right) \right\rangle \\
&= \left\langle \sum_{k}^{M} V_{ki} \left(e_k - \langle e_k \rangle \right) \sum_{l}^{M} V_{lj} \left(e_l - \langle e_l \rangle \right) \right\rangle \\
&= \sum_{k}^{M} \sum_{l}^{M} V_{ki} V_{lj} \left\langle \left(e_k - \langle e_k \rangle \right) \left(e_l - \langle e_l \rangle \right) \right\rangle \\
&= \sigma^2 \sum_{k}^{M} \sum_{l}^{M} V_{ki} V_{lj} \delta_{kl} = \sigma^2 \sum_{k}^{M} V_{ki} V_{kj} = \sigma^2 \delta_{ij},
\end{aligned}
\tag{54}
$$

where we used the fact that the elements of e are uncorrelated and equal in going from the first line of Eq. (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 $\underline{\tilde{e}}$ 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

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

where we have used the fact that the means of the e_i are zero.[6] Because each $\tilde{e}_i$ is a linear combination of the e_i, their joint density has the form (by generalizing the derivations on pages 199-201 of [17] from two variables to M variables),

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

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

---

[6] Zero 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.

$$q_{1,2,\ldots,M}\left(\tilde{e}_1,\tilde{e}_2,\ldots\tilde{e}_M\right)$$

$$=\left(\sigma\sqrt{2\pi}\right)^{-M}\exp\left(-\sum_{i=1}^{M}\left(V\tilde{e}\right)_i^2\Big/\left(2\sigma^2\right)\right)$$

$$=\left(\sigma\sqrt{2\pi}\right)^{-M}\exp\left(-\left(V\tilde{e}\right)^t V\tilde{e}\Big/2\sigma^2\right)$$

$$=\left(\sigma\sqrt{2\pi}\right)^{-M}\exp\left(-\tilde{e}^t V^t V\tilde{e}\Big/2\sigma^2\right) \tag{57}$$

$$=\left(\sigma\sqrt{2\pi}\right)^{-M}\exp\left(-\tilde{e}^t\tilde{e}\Big/2\sigma^2\right)$$

$$=\left(\sigma\sqrt{2\pi}\right)^{-M}\exp\left(-\sum_{i=1}^{M}\tilde{e}_i^2\Big/2\sigma^2\right),$$

where we have used the orthogonality of V in going from the second line of Eq. (57) to the third line. Eq. (57) can be written as a product of functions of the individual $\tilde{e}_i$ as follows: $q_{1,2,\ldots,M}\left(\tilde{e}_1,\tilde{e}_2,\ldots\tilde{e}_M\right)=\prod_{i=1}^{M}q_i\left(\tilde{e}_i\right)$, where $q_i\left(\tilde{e}_i\right)=\left(\sigma\sqrt{2\pi}\right)^{-1}\exp\left(-\tilde{e}_i^2\Big/\left(2\sigma^2\right)\right)$, 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.

Two dimensional case

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

$$\Delta R = \begin{bmatrix} 0 & -\Delta\theta \\ \Delta\theta & 0 \end{bmatrix}. \tag{58}$$

Third, *C* is a now a 2*N*-by-3 matrix, and *q* is a three-element vector whose

elements are $\Delta\theta, \Delta t_1, \Delta t_2$. *Fourth*,

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

and

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

Fifth, Λ is 2*N*-by-3, *V* is 3-by-3, Λ^+ is 3-by-2*N*, 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 Eqs. (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 $\tilde{e}$, and each component of $\mathbf{TRE}(r)$ is a function of only the first 3

elements of $\tilde{e}$. The proof that the 2*N* elements of $\tilde{e}$ are independent makes no reference

to the dimension of the space (see "Statistical independence of the elements of $\tilde{e}$" in

Section 8 (page 73)). Therefore, $\mathbf{FRE}_i$ and $\mathbf{TRE}(r)$ are statistically independent for the two-dimensional case, just as they are for the three-dimensional case.

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

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

$$\left(\Sigma_{\mathrm{FRE}}\right)_{kl} = \sum_{j,j'=4}^{2N} U_{kj}U_{lj'}\left\langle \tilde{e}_j\tilde{e}_{j'}\right\rangle = \sum_{j,j'=4}^{2N} U_{kj}U_{lj'}\left(U^tW\Sigma W^tU\right)_{jj'} \tag{61}$$

$$\begin{aligned}
\mathrm{FRE}_i^2 &= \sum_{k=2i-1}^{2i}\left(\sum_{j,j'=4}^{2N} U_{kj}U_{kj'}\left\langle \tilde{e}_j\tilde{e}_{j'}\right\rangle\right)_{kk} \\
&= \sum_{k=2i-1}^{2i}\left(\sum_{j,j'=4}^{2N} U_{kj}U_{kj'}\left(U^tW\Sigma W^tU\right)_{jj'}\right)_{kk}
\end{aligned} \tag{62}$$

$$\begin{aligned}
\left\langle \mathrm{FRE}^2\right\rangle &= \sum_{k=1}^{2N}\left(\Sigma_{\mathrm{FRE}}\right)_{kk} = \sum_{k=1}^{2N}\sum_{j,j'=4}^{2N} U_{kj}U_{kj'}\left(U^tW\Sigma W^tU\right)_{jj'} \\
&= \sum_{j=4}^{2N}\left(U^tW\Sigma W^tU\right)_{jj}
\end{aligned} \tag{63}$$

$$\left(\Sigma_{\mathrm{fre}}\right)_{kl} = \sum_{j,j'=4}^{2N}\left(W^{-1}U\right)_{kj}\left(W^{-1}U\right)_{lj'}\left(U^tW\Sigma W^tU\right)_{jj'} \tag{64}$$

$$\left\langle \mathrm{fre}_i^2\right\rangle = \sum_{k=2i-1}^{2i}\sum_{j,j'=4}^{2N}\left(W^{-1}U\right)_{kj}\left(W^{-1}U\right)_{kj'}\left(U^tW\Sigma W^tU\right)_{jj'} \tag{65}$$

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

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

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

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

Chi-square test for dependence

To detect dependence, we employ the standard chi-square test based on the contingency tables [21]. Each element, $C(i,j)$, $i=1,2,...,N_C$, $j=1,2,...N_C$, of an $N_C \times N_C$ contingency table is a count of the number of registrations for which $\text{FRE}_i \leq \text{FRE} < \text{FRE}_{i+1}$ and $\text{TRE}_j \leq \text{TRE} < \text{TRE}_{j+1}$, where the thresholds, FRE_i and TRE_j can be chosen arbitrarily. The resulting table is then used to calculate the following test statistic: $q = \sum_{i,j=1}^{N_C} \left(C(i,j) - C_0(i,j) \right)^2 / C_0(i,j)$, 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) $\chi_v^2(1-p_0)$, [7] where $1-p_0$ is the desired confidence in the truth of dependence. The number v of degrees of freedom for this problem equals $(N_C - 1)^2$. Thus, a value of q that is larger than $\chi_v^2(1-p_0)$ indicates that dependence between FRE and TRE is detected at $p \le p_0$. Failure to detect q exceeding $\chi_v^2(1-p_0)$ 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 below a Matlab function to implement our new formulas for FRE and TRE. The values that it returns correspond to our equations as follows: $(RMS\_TRE)^2 = \langle TRE^2 \rangle$ from Eq. (41), $(RMS\_FRE)^2 = \langle FRE^2 \rangle$ from Eq. (35), $Cov\_TRE = \Sigma_{TRE}$ from Eq. (40), and $Cov\_FRE = \Sigma_{FRE}$ from Eq. (33). The inputs correspond to our definitions above as follows: $X0(:,i) = x_i^{(0)} = R^{(0)}x_i$, $W(:,:,i) = W_i$, $Cov\_FLE(:,:,i) = \Sigma_i = R^{(0)}\Sigma_i^{(1)}R^{(0)^t} + \Sigma_i^{(2)}$, and $r0 = r^{(0)} = R^{(0)}r$.

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. [5], uniform weighting for homogeneous, anisotropic
FLE first solved by Wiles et al. [8], and ideal weighting for inhomogeneous, anisotropic
FLE first solved by Moghari et al. [10] , and it handles all other cases as well.

```
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));
```

9.   References

[1]   Fitzpatrick, J.M. Fiducial registration error and target registration error are uncorrelated, SPIE Medical Imaging 2009, Lake Buena Vista, FL, SPIE 7261, 2009.

[2]   Steinmeier, R., Rachinger, J., Kaus, M., Ganslandt, O., Huk, W., and Fahlbusch, R. Factors influencing the application accuracy of neuronavigation systems, Stereotactic and Functional Neurosurgery, 75, 188-202, 2000.

[3]   Woerdeman, P.A., Willems, P.W., Noordmans, H.J., Tulleken, C.A., and Van Der Sprenkel, J.W. Application accuracy in frameless image-guided neurosurgery: a comparison study of three patient-to-image registration methods, J. Neurosurg., 106(6), 1012-1016, 2007.

[4]   Sibson, R. Studies in the robustness of multidimensional scaling: Perturbational analysis of classical scaling, Journal of the Royal Statistical Society: Series B, 41, 217–229, 1979.

[5]   Fitzpatrick, J.M., West, J.B., and Maurer, C.R., Jr. Predicting error in rigid-body, point-based image registration, IEEE Transactions on Medical Imaging, 17, 694–702, 1998.

[6]   Fitzpatrick, J.M., and West, J.B. The distribution of target error in rigid-body, point-based registration, IEEE Transactions on Medical Imaging, 20, 917–927, 2001.

[7]   West, J.B. and Maurer, C.R., Jr. Designing optically tracked instruments for image-guided surgery, IEEE Transactions on Medical Imaging, 23, 533–545, 2004.

[8]   Wiles, A.D., Likholyot, A., Frantz, D.D., and Peters, T.M. A statistical model for point-based target registration error with anisotropic fiducial localizer error, IEEE Transactions on Medical Imaging, 27(3), 378–390, 2007.

[9]   Balachandran, R., Fitzpatrick, J.M. The distribution of registration error of a fiducial marker in rigid-body point-based registration, Proc. SPIE 6981, San Diego, CA, 2008.

[10] Moghari, M. and Abolmaesumi, P. Distribution of target registration error for anisotropic and inhomogeneous fiducial localization error, IEEE Transactions on Medical Imaging, 28(6), 799-813, 2009.

[11] Suess, O., Picht, T., Kuehn, B., Mularski, S., Brock, M., and Kombos, T. Neurosurgery without rigid pin fixation of the head in left frontotemporal tumor surgery with intraoperative speech mapping, Operative Neurosurgery Suppl. 2, 60, 330-338, 2007.

[12] Fitzpatrick, J.M. and Galloway, R.L. The use of fiducials in the registration of 3D images, Automedica Special Issue: Engineering Approaches to Neurological Surgery, 20, 2001.

[13] Fitzpatrick, J.M. Detecting failure, assessing success, Medical Image Registration, J.V. Hajnal, D. L. G. Hill, and D. J. Hawkes, eds., CRC Press, 117-139, 2001.

[14] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. Numerical Recipes in C: The Art of Scientific Computing, 2nd ed., Cambridge University Press, New York, NY, 1992.

[15] Ohta, N. and Kanatani, K. Optimal estimation of 3-D rotation and reliability evaluation, Lecture Notes in Computer Science, 1406, 175-187, 1998.

[16] West, J.B., Fitzpatrick, J.M., and Batchelor, P.G. Point-based registration under a similarity transform, Proc. SPIE Medical Imaging 2001, San Diego, CA, 4322, 611–622, 2001.

[17] Papoulis, A. and Pillai, S.U. Probability, Random Variables, and Stochastic Processes. 4th Edition, McGraw Hill, Boston, 266-268, 2002.

[18] Moghari, M. and Abolmaesumi, P. Distribution of fiducial registration error in rigid-body point-based registration, IEEE Transactions on Medical Imaging, 28(11), 1791-1801, 2009.

[19] Fitzpatrick, J.M., Hill, D.L.G., and Maurer, C.R., Jr. Registration, Medical Image Processing: Volume II of the Handbook of Medical Imaging, M. Sonka and J.M. Fitzpatrick, ed., SPIE Press, 447-513, 2000.

[20] Balachandran, R., Fitzpatrick, J.M. Iterative solution for rigid-body point-based registration with anisotropic weighting, SPIE Medical Imaging 2009, Lake Buena Vista, FL, SPIE 7261, 2009.

[21] Spiegel, M.R. Probability and Statistics, Schaum's Outline Series, McGraw-Hill, 116, 1975.

[22] Wiles, A.D. Navigation Accuracy of Image-Guided Interventional Systems, Dissertation. Department of Medical Biophysics, University of Western Ontario, London, Ontario, Canada, 2009.

[23] Roche, A., Malandain, G., and Ayache, N. Unifying maximum likelihood approaches in medical image registration, Int. J. Imaging Syst. And Technol.: Special Issue on 3D Imaging, 11, 71-80, 2000.

CHAPTER III

ITERATIVE SOLUTION FOR RIGID-BODY POINT-BASED REGISTRATION
WITH ARBITRARY WEIGHTING

1. Introduction

The key component of image-guided systems is the ability to register multi-modal pre-operative and intra-operative images of a patient or an atlas to the patient in physical space during the operation. Since the introduction of fiducial-based systems in image-guided surgery, point registration has become an important aspect of the whole procedure. The process of point-based registration involves identifying two sets of corresponding points in the two spaces that need to be aligned by means of a transformation, or "mapping", of one of the data sets and determining the transformation that aligns the two spaces. Each set of points consists of three or more points in a space, e.g. three-dimensional image volume or patient in physical space, each of which represent a fiducial marker.

For surgeries, such as neurosurgery, fiducial markers are rigidly attached to the patient's skull to achieve submillimetric accuracy while accessing regions close to critical structures. Fiducial markers are bone-implanted before acquisition of pre-operative images and are removed at the end of the procedure. Since the skull is rigid and the fiducials are rigidly attached to the skull, the point mapping is assumed to be a rigid transformation, and that is also the assumption in the present work.

For a typical surgical-guidance system, the two spaces for point-based registration are the pre-operative image of the patient, such as CT and MRI, and physical space. The sets of fiducial points are obtained by localizing each fiducial marker both in the pre-operative image and in physical space of the operating room. It is almost impossible to localize fiducials without appreciable error. This error in localizing the fiducials is called the fiducial localization error (FLE). If FLE is the same for all fiducials, FLE is termed "homogeneous", otherwise it is termed "inhomogeneous." FLE is called "isotropic" if it is the same in all directions, otherwise it is called "anisotropic." Due to FLE, perfect registration between the two sets of fiducial points is not possible. The resultant error in aligning corresponding fiducial points is called the fiducial registration error (FRE). In the general case, FRE is the root-mean-square (RMS) of the lengths of the displacements between corresponding fiducials after registration with each displacement multiplied by a weighting matrix. The same terminology—homogeneous or inhomogeneous, isotropic or anisotropic—is used to describe the weighting scheme. When weighting matrices are employed, FRE may be called a "weighted FRE". If each weighting matrix is the identity, then FRE is the "unweighted FRE". Registration of points in one space with points in another space is a process of finding the transformation that minimizes FRE. A least-squares approach is commonly utilized to determine the transformation. (See "Derivation of ideal weighting" in Section 8 of Chapter II (page 70) for a justification of the least-squares approach.) A closed-form solution exists for this least-squares problem when the weighting matrices are isotropic (whether or not they are homogeneous) [1], [2]. However, there is no known closed-form solution when the weighting matrices are

anisotropic, which is required for ideal registration when the statistical distribution of FLE is anisotropic.

For image-guided surgery applications, FLE in image space is made highly isotropic by using fiducials of size larger than the size of a voxel. However, FLE in physical space typically suffers from noticeable anisotropy because of the limitations of tracking techniques [3]. For example, physical positions are often acquired via optical tracking systems such as the Polaris Spectra (Northern Digital Inc., Waterloo, Ontario, Canada), whose localization error is known to be larger in the direction from the camera to the object than in the other two orthogonal directions. Furthermore, the positions are often calculated relative to a coordinate reference frame (CRF) that is rigidly attached to the patient. The use of a CRF enables compensation for patient movement relative to the tracking system during the procedure, but it tends to increase the anisotropy [3]. Thus, a solution that allows for anisotropy is of value. Such a solution is obtained by defining anisotropic (and possibly inhomogeneous) weighting matrices that are determined by the known spatial distributions of FLE and then minimizing the resultant weighted FRE. We call this procedure "anisotropic registration". While no closed-form solution exists, several approximate iterative solutions to anisotropic registration exist. Koschat [4] and Chu [5] each look at the problem of mapping two point sets with anisotropic errors as a Procrustes problem. Moghari [7] and Pennec [8] each make use of Kalman filtering to find the approximate solution. Balachandran [6] proposes a non-iterative algorithm that finds an approximate solution to the registration problem with anisotropic FLE, but anisotropy is incorporated only in the calculation of the centroids of the marker configuration and in one space.

Each iterative algorithm requires some stopping criterion. In addition, each of the algorithms above requires that additional parameters be set as well. In some cases, as few as one additional parameter must be chosen, as for example in Koschat and in Chu, but some require as many as three or four, as for example Moghari. However, there are two previously published algorithms that require no additional parameters. The first of these methods is by Ohta [9], which estimates the rotation that is optimal in the sense of achieving the lower bound of a first-order level of accuracy. The method makes use of a quaternion renormalization technique to solve the non-linear problem. However, in the case when translation is present, this algorithm loses its optimality. The second method is by Matei [10]. This algorithm extends the idea of the Ohta algorithm to keep the optimality of the found solution for cases when both rotation and translation are present. Matei interprets the anisotropic registration problem as a heteroscedastic regression problem.

In this chapter we present an alternative iterative algorithm that, like Ohta and like Matei, has no additional parameters other than a stopping threshold. This algorithm is considerably simpler and more intuitive than existing algorithms of solving the anisotropic registration problem. At each stage of iteration of the new algorithm, a non-linear problem, which has no closed-form solution, is replaced with a linear one, which has a closed-form solution. The solution of the linear problem is used to compute the closest rotation matrix for the non-linear problem and to update the fiducial points for the next iteration. We show that in the presence of anisotropic FLE, a substantial improvement in the target registration error (TRE) is achieved when the new algorithm is used instead of the commonly used closed-form solutions. We also compare the new

algorithm with Matei's algorithm using simulation and using experimental data for an image-guided surgery application. We show that for clinically-applicable cases the new algorithm improves the TRE at clinically-relevant targets relative to the other algorithms. Our choice of the two iterative algorithms for comparison is based on the fact that these two algorithms require no parameter other than a stopping threshold to be set, just like the new proposed algorithm.

2. Method

The problem of point-based registration in the presence of inhomogeneous anisotropic FLE is a problem of finding the rigid transformation that minimizes the anisotropically-weighted FRE. We let $X = \{\mathbf{x}_i\}$, $i = 1...N$, be the set of N 3-by-1 fiducial points to be transformed (e.g., points in image space), called the "moving" fiducials because they are being transformed, and $Y = \{\mathbf{y}_i\}$, $i = 1...N$, be the set of corresponding fiducial points (e.g., points in physical space), called the "stationary" fiducials. The only requirement for the fiducials is that at least three of the N fiducials are not collinear. Our goal is to find the 3-by-3 rotation matrix R and the 3-by-1 translation vector $\mathbf{t}$ that minimize weighted FRE, which is defined as follows:

$$\text{FRE}^2 = \sum_{i=1}^{N} \left| W_i \left(R\mathbf{x}_i + \mathbf{t} - \mathbf{y}_i \right) \right|^2 , \tag{1}$$

where W_i is a 3-by-3 non-singular weighting matrix of fiducial i. In this expression the weighting matrices could be functions of R. With this definition of weighted FRE, it is clear that R and $\mathbf{t}$ minimizing Eq. (1) will not be changed if all W_i are multiplied by the same factor w, and FRE^2 is multiplied by a positive factor w^2. Therefore the scale of FRE is somewhat uncertain. To remove this uncertainty, we normalize all the weighting matrices such that $\sum_1^N |W_i|^2 = \sum_1^N \text{trace}\left(W_i^t W_i\right) = 1$. The problem of minimizing Eq. (1) is made difficult by the non-linear constraint on the rotation matrix. The rotation matrix is restricted to be orthogonal, which means that equality $R^t R = I$ must be satisfied, where I is the identity matrix.

The weights account for the variation in the localization accuracy among the fiducials, a variation with respect to the direction and with respect to the markers. By definition, if $W_i = W_j$ for all i,j, then FLE is homogeneous. We assume that FLE for each fiducial is normally distributed with zero mean and can be resolved into three uncorrelated components along a set of orthogonal principal axes, with standard deviations $\sigma_{\alpha i k}$, where $\alpha = x, y$ indicates the space and $k = 1, 2, 3$ indicates the principal axis. If $\sigma_{\alpha i k} = \sigma_{\alpha i k'}$ for all k, k', the problem reduces to the isotropic problem, which can be solved by closed-form methods, both for the homogenous and the inhomogeneous cases [11]. In this chapter we treat the anisotropic case, for which there is no known closed-form solution. This is the situation that arises in surgical navigation, in which the localization error of the physical tracking system in the operating room suffers from a relatively larger standard deviation in the direction from the camera to the fiducials than in the perpendicular directions.

In general case the weighting matrix has the form

$$W_i = \left(R W_{xi}^{-2} R^t + W_{yi}^{-2} \right)^{-1/2} ,$$

(2)

where W_{xi} and W_{yi} are weighting matrices of fiducial i in the image and physical spaces correspondingly.

A special case of weighting is "ideal weighting". (See "Derivation of ideal weighting" in Section 8 of Chapter II (page 70)). It is an important case because ideal weighting maximizes the probability for the resulting transformation to be the true one [12]. The ideal weighting for the fiducial i in space α is defined as the inverse of FLE. The ideal weighting matrix for fiducial i in space α is the defined as

$$W_{\alpha i} = \left(V_{\alpha i} \operatorname{diag}\left(\sigma_{\alpha i1}^{-2}, \sigma_{\alpha i2}^{-2}, \sigma_{\alpha i3}^{-2} \right) V_{\alpha i}^t \right)^{-1/2}$$

where $V_{\alpha i}^t V_{\alpha i} = I$. The columns of $V_{\alpha i}$ are the principal axes in space α of the FLE for fiducial i. The overall weighting matrix W_i then becomes

$$W_i = \left(R V_{xi} \operatorname{diag}\left(\sigma_{xi1}^2, \sigma_{xi2}^2, \sigma_{xi3}^2 \right) V_{xi}^t R^t + V_{yi} \operatorname{diag}\left(\sigma_{yi1}^2, \sigma_{yi2}^2, \sigma_{yi3}^2 \right) V_{yi}^t \right)^{-1/2}$$

(3)

Ideal weighting was first introduced by Ohta [9]. Later it was used by Matei in [10], West in [13] and Moghari in [14], [15]. Note that the derivations and algorithms

presented in this paper work with general weighting as well as with ideal weighting. However, ideal weighting is employed in all our validation experiments.

Algorithm

The strategy of the new iterative method of anisotropic point-based registration is that the non-linear problem is replaced with a simple linear problem at each stage of iteration and the exact solution to the linear problem is found by means of linear algebra. This method is a common way of dealing with non-linear problems [16], [17]. The simplification allowing transformation of the non-linear problem into linear one involves replacing the rotation matrix by its linear approximation, which is subject just to a linear constraint. As a result of this simplification, the exact solution of the linear problem is also an approximate solution to the non-linear registration problem. The rotation matrix is then found as the matrix which is closest in the least-squares sense to the solution of the linear problem. Finally, we apply the found rotation matrix to the moving fiducials to bring them incrementally closer to the stationary fiducials. In summary, we repeat the following steps: (1) solve the linearized problem, (2) find the closest rigid transformation to the linear solution, and (3) apply the rigid transformation to the moving fiducials. The algorithm stops when a measure of the movement of the fiducials is below selected threshold. That threshold is the single adjustable parameter of the algorithm.

The inputs to the algorithm are N pairs of 3-by-1 points $\mathbf{x}_i$ and $\mathbf{y}_i$, and 3-by-3 weighting matrices W_{xi} and W_{yi} for $i = 1, \ldots, N$. At each stage of the iterative process, the rotation matrix, the translation vector, and the set of moving fiducials are updated to new values. At iteration stage n, matrix $R^{(n-1)}$, vector $\mathbf{t}^{(n-1)}$, and the points $\mathbf{x}_i^{(n-1)}$ are updated

to $R^{(n)}$, $\mathbf{t}^{(n)}$, and $\mathbf{x}_i^{(n)}$. Before starting the iterative process, a good initialization is required and is provided as the R and $\mathbf{t}$ that minimize Eq. (1) with isotropic, homogeneous FLE, *i.e.* the weighting matrices W_i are omitted. The solution is found by means of the closed-form solution based on singular value decomposition provided in Algorithm 8.1 of [11]. The found rotation matrix and translation vector are labeled $R^{(0)}$ and $\mathbf{t}^{(0)}$. Using these transformation parameters and the coordinates of the moving fiducials, the initial moving points for the iterative algorithm are calculated using the following expression: $\mathbf{x}_i^{(0)} = R^{(0)}\mathbf{x}_i + \mathbf{t}^{(0)}$, for $i = 1,.., N$. The rotation matrix and translation vector solution of the anisotropic registration problem are initialized as $R = R^{(0)}$ and $\mathbf{t} = \mathbf{t}^{(0)}$.

Following initialization, the iterative process begins. Each stage n of the iteration comprises the following steps:

1. Solve the linear problem: In Eq. (1), $\mathbf{x}_i$ is replaced by $\mathbf{x}_i^{(n-1)}$, R is replaced by an approximate rotation operator $I + \Delta\Theta^{(n)}$, where $\Delta\Theta^{(n)}$ is a 3-by-3 antisymmetric matrix with the linear constraint $\Delta\Theta^{(n)t} = -\Delta\Theta^{(n)}$. The following approximation of Eq. (1) is minimized:

$$\text{FRE}^2 = \sum_{i=1}^{N} \left(\left(I + \Delta\Theta^{(n)} \right) \mathbf{x}_i^{(n-1)} + \mathbf{t}^{(n)} - \mathbf{y}_i \right)^t W_i^2 \left(\left(I + \Delta\Theta^{(n)} \right) \mathbf{x}_i^{(n-1)} + \mathbf{t}^{(n)} - \mathbf{y}_i \right), \qquad (4)$$

where $W_i = \left(RW_{xi}^{-2}R^t + W_{yi}^{-2} \right)^{-1/2}$ is calculated using the current R.

The matrix $\Delta\Theta^{(n)}$ and the translation vector $\mathbf{t}^{(n)}$ that together minimize approximation Eq. (4), can be found exactly. The solution method is given below in "Solution of the approximate equation" in Section 2 (page 94).

2. If $n > 1$, compute $\Delta\Theta^{(n)} = \left(\Delta\Theta^{(n-1)} + \Delta\Theta^{(n)} \right) \big/ 2$ and $\mathbf{t}^{(n)} = \left(\mathbf{t}^{(n-1)} + \mathbf{t}^{(n)} \right) \big/ 2$.

3. Perform singular value decomposition of matrix $I + \Delta\Theta^{(n)}$: $U\Lambda V^t = I + \Delta\Theta^{(n)}$, where U and V are 3-by-3 rotation matrices and Λ is a 3-by-3 diagonal matrix with non-negative elements.

4. Calculate the 3-by-3 matrix $R^{(n)} = UV^t$. $R^{(n)}$ is a proper rotation matrix that is closest in the least-squares sense to $I + \Delta\Theta^{(n)}$ (see "Finding the closest rotation matrix" in Section 2 (page 96)).

5. Update the rotation matrix and translation vector: $R = R^{(n)}R$, $\mathbf{t} = R^{(n)}\mathbf{t} + \mathbf{t}^{(n)}$.

6. Updated the moving point set: $\mathbf{x}_i^{(n)} = R\mathbf{x}_i + \mathbf{t}$, for $i = 1,..,N$.

7. Calculate the relative change in points configuration:

$$\Delta X = \left(\sum_{i=1}^{N} \left| \mathbf{x}_i^{(n)} - \mathbf{x}_i^{(n-1)} \right|^2 \bigg/ \sum_{i=1}^{N} \left| \mathbf{x}_i^{(n)} - \overline{\mathbf{x}}^{(n)} \right|^2 \right)^{1/2},$$ where $\overline{\mathbf{x}}^{(n)}$ is the centroid of the points configuration $\left\{ \mathbf{x}_i^{(n)} \right\}$.

8. If $\Delta X >$ threshold, update n to $n+1$ and go to Step 1.

A Matlab (MathWorks, Inc., Natick, MA) implementation of this algorithm is provided in Section 7 (page 126).

Solution to the linear approximate problem

In Step 1 of the new algorithm, we must find $\Delta\Theta^{(n)}$ and $\mathbf{t}^{(n)}$ that minimize the approximate expression for FRE^2 in Eq. (4). We note that this minimization is equivalent to finding the least-squares solution of the set of $3N$ equations:

$$W_i\Delta\Theta^{(n)}\mathbf{x}_i^{(n-1)} + W_i\mathbf{t}^{(n)} = W_i\Delta_i^{(n-1)}, \tag{5}$$

where $\Delta_i^{(n-1)} = \mathbf{y}_i - \mathbf{x}_i^{(n-1)}$. The unknowns in these equations are the six non-zero elements of the 3-by-3 antisymmetric matrix $\Delta\Theta^{(n)}$ and the three elements of vector $\mathbf{t}^{(n)}$. We note that for an antisymmetric matrix all diagonal elements equal zero and the off-diagonal elements have the relationship: $\Delta\Theta_{ij}^{(n)} = -\Delta\Theta_{ji}^{(n)}$. Thus, there are only three independent unknown elements in matrix $\Delta\Theta^{(n)}$: $\Delta\Theta_{32}^{(n)}$, $\Delta\Theta_{13}^{(n)}$, and $\Delta\Theta_{21}^{(n)}$. To understand the meaning of these three elements, recall that $I + \Delta\Theta^{(n)}$ is an approximation of rotation R. If the angle of rotation R about its axis is small, then the movement $R\mathbf{x}_i^{(n)} - \mathbf{x}_i^{(n)}$ can be approximated as a cross-product between the axis of rotation R and vector $\mathbf{x}_i^{(n)}$. If the elements of 3-by-1 vector $\Delta\boldsymbol{\theta}$ are $\Delta\theta_1 = \Delta\Theta_{32}^{(n)}$, $\Delta\theta_2 = \Delta\Theta_{13}^{(n)}$, and $\Delta\theta_3 = \Delta\Theta_{21}^{(n)}$, then for small rotations the axis of rotation lies approximately along vector $\Delta\boldsymbol{\theta}$ and the angle of rotation is approximately equal to the length of $\Delta\boldsymbol{\theta}$. Thus, we get the following formula:

$$R\mathbf{x}_i^{(n)} - \mathbf{x}_i^{(n)} \approx \Delta\boldsymbol{\theta} \times \mathbf{x}_i^{(n)}. \tag{6}$$

The vector $\Delta\boldsymbol{\theta}$ can be used to transform Eq. (5) into the form,

$$Cq = e, \tag{7}$$

where C is a 3N-by-6 matrix, q is a 6-by-1 vector of unknowns, and e is a 3N-by-1 vector. The elements of q are defined as follows:

$$q_1 = \Delta\theta_1, q_2 = \Delta\theta_2, q_3 = \Delta\theta_3, q_4 = \Delta t_1, q_5 = \Delta t_2, q_6 = \Delta t_3. \tag{8}$$

To give expressions for calculating the elements of matrix C and vector e, we specify the jk-th element of the weighting matrix W_i in the following form: $W_{jk,i}$. We separate the third subscript with a comma to emphasize that it is not a matrix index, but a fiducial index. Additionally, we specify element j of vector $\Delta_i^{(n)}$ with two subscripts: $\Delta_{j,i}^{(n)}$. Here also we separate the second subscript with a comma to emphasize that it is not a vector index, but a fiducial index. With these notations established, after detailed inspection of Eq. (5), we get the expressions for elements of matrix C:

$$
\begin{aligned}
C_{3(i-1)+j,1} &= -W_{j2,i}x_{3i}^{(n)} + W_{j3,i}x_{2i}^{(n)}, \\
C_{3(i-1)+j,2} &= +W_{j1,i}x_{3i}^{(n)} - W_{j3,i}x_{1i}^{(n)}, \\
C_{3(i-1)+j,3} &= -W_{j1,i}x_{2i}^{(n)} + W_{j2,i}x_{1i}^{(n)}, \\
C_{3(i-1)+j,4} &= W_{j1,i}, \\
C_{3(i-1)+j,5} &= W_{j2,i}, \\
C_{3(i-1)+j,6} &= W_{j3,i},
\end{aligned}
\tag{9}
$$

and the expressions for elements of vector e:

95

$$e_{3(i-1)+j,1} = W_{j1,i}\Delta_{1,i}^{(n)} + W_{j2,i}\Delta_{2,i}^{(n)} + W_{j3,i}\Delta_{3,i}^{(n)}. \tag{10}$$

With these definitions, the solution minimizing Eq. (4) is found by solving Eq. (7) for q by using any appropriate numerical method, and then setting

$$\Delta\Theta^{(n)} = \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix} \tag{11}$$

and

$$\mathbf{t}^{(n)} = \begin{bmatrix} q_4 & q_5 & q_6 \end{bmatrix}^t. \tag{12}$$

Finding the closest rotation matrix

In Step 4 of the new algorithm, we set $R^{(n)} = UV^t$, where U and V are 3-by-3 rotation matrices obtained from the singular value decomposition of the approximate rotation matrix: $U\Lambda V^t = I + \Delta\Theta^{(n)}$. We can prove that $R^{(n)}$ found in Step 4 of the new algorithm is the closest rotation matrix to $I + \Delta\Theta^{(n)}$ in the least squares sense as follows. Define matrix $M = I + \Delta\Theta^{(n)}$. Let matrix R be the rotation matrix that is closest to M. Form the difference matrix $E = M - R$. We define operator $|A|^2$ as the sum of squared elements of A. With the above definitions we get:

$$\begin{aligned}
|E|^2 &= \sum_{i,j}\left(M_{ij}-R_{ij}\right)^2 \\
&= \operatorname{trace}(M-R)^t(M-R) \\
&= \operatorname{trace}M^tM - 2\operatorname{trace}(M^tR)+\operatorname{trace}(I).
\end{aligned} \qquad (13)$$

Matrix R that we are looking for is the one that minimizes $|E|^2$. From the last line of Eq. (13), it is easy to see that $|E|^2$ is minimized when $\operatorname{trace}(M^tR)$ is maximized. To maximize $\operatorname{trace}(M^tR)$, we employ the singular value decomposition of matrix M: $U\Lambda V^t = M$, where U and V are 3-by-3 rotation matrices and Λ is 3-by-3 diagonal matrix with non-negative elements. Substituting this for M,

$$\begin{aligned}
\operatorname{trace}(M^tR) &= \operatorname{trace}(V\Lambda U^tR) \\
&= \operatorname{trace}(\Lambda U^tRV) \\
&= \operatorname{trace}(\Lambda Z) \\
&= \sum_{i=1}^{3}\Lambda_{ii}Z_{ii},
\end{aligned} \qquad (14)$$

where $Z = U^tRV$. To get from the first line to the second line in Eq. (14), we used the property of the trace that $\operatorname{trace}(AB) = \operatorname{trace}(BA)$. We note that, by virtue of its being a product of rotation matrices, Z is also a rotation matrix. The maximum value for any element of a rotation matrix is 1. Since each Λ_{ii} is non-negative, the maximum of the sum in the last line of Eq. (14) is reached when $Z_{ii} = 1$ for $i = 1,2,3$. Thus, Z is 3-by-3 identity matrix I. Therefore, $I = U^tRV$. By multiplying this equation on the left by U and on the right by V^t, we find that $R = UV^t$ (this proof is patterned after a method published by Golub and van Loan in [18]).

To prove that matrix R is a proper rotation, we note that an improper rotation matrix always has a negative determinant. Thus, all we need to show is that the determinant of R is non-negative. Noting that $\det(AB) = \det(A)\det(B)$, we get:

$$\begin{aligned}
\det(I + \Delta\Theta) &= \det(U)\det(\Lambda)\det(V^t) \\
&= \det(\Lambda)\det(UV^t) \\
&= \det(\Lambda)\det(R).
\end{aligned} \tag{15}$$

The determinant of a diagonal matrix equals the product of the matrix diagonal elements. All diagonal elements of matrix Λ are non-negative by the definition of singular value decomposition. Thus, the sign of the determinant of R is equal to the sign of $\det(I + \Delta\Theta)$. We find an expression for the latter determinant directly:

$$\begin{aligned}
\det(I + \Delta\Theta) &= \det\begin{bmatrix} 1 & -\Delta\theta_3 & \Delta\theta_2 \\ \Delta\theta_3 & 1 & -\Delta\theta_1 \\ -\Delta\theta_2 & \Delta\theta_1 & 1 \end{bmatrix} \\
&= 1 + \Delta\theta_1^2 + \Delta\theta_2^2 + \Delta\theta_3^2.
\end{aligned} \tag{16}$$

Therefore, $\det(I + \Delta\Theta) > 0$ and correspondently the determinant of R is also non-negative.

3. Validation

Our validation is based both on simulations and on real data. For each trial in each experiment of our validation, a set $X = \{\mathbf{x}_i\}$, $i = 1...N$, of fiducial points is transformed to match a second set $Y = \{\mathbf{y}_i\}$, $i = 1...N$. In our simulations, each point is obtained from a "true" point that is perturbed by a randomly selected FLE selected from a distribution with mean zero and a known covariance. In the experiments on real data, true points are unknown, but highly accurate measurements are available to approximate true points. Five algorithms are compared. Three of these are iterative, and each iterative algorithm requires that a stopping threshold be set. This threshold serves as the minimum change in an internally defined quantity for continued iteration, such as in Step 8 of the new algorithm ("Algorithm" in Section 2 (page 91)). All experiments reported on below were run with a threshold 10^{-6} for each algorithm. For convenience in the descriptions that follow, we define "convergence" to mean reduction of the internally defined quantity to a value below this threshold and a failure to converge to mean that more than 1000 iterations were completed without convergence. To determine whether the results were sensitive to threshold, additional experiments were run with thresholds of 10^{-3} and 10^{-9} with almost identical results when convergence was achieved, suggesting that, when they converge, their relative performance has a low sensitivity to the choice of threshold.

A. Comparison by Matlab simulations of Ohta's algorithm to the new algorithm modified for zero translation.

The comparison was performed by running simulations in Matlab R2010b. Four values of *N* were chosen for the tests: *N* = 3, 4, 5, and 10. For each value of *N*, we randomly chose that number of fiducial locations in image space from a cube of edge 200

mm with center at the origin to build the fiducial set X in image space. The corresponding fiducial set Y in physical space was obtained by rotating the X fiducial configuration arbitrarily: 10, −20, and 30 degrees rotation about the x, y, and z axes. The X and Y sets represent the "true" fiducials. Note that no translation was involved in this experiment because Ohta's algorithm cannot handle translation. The elements of point sets X and Y were randomly perturbed to simulate inhomogeneous anisotropic FLE. The FLEs in image space and in physical space in the k-th direction for the i-th fiducial were drawn correspondingly from $N(0,\text{FLE}_{Xik})$ and $N(0,\text{FLE}_{Yik})$, where all FLE_{Xik} and FLE_{Yik} were randomly chosen from interval $[0,1]$. The perturbed X and Y were then registered. To have a fair comparison, we modified the new algorithm so that transformation is performed just in the form of rotation. A random target was chosen inside a cube with an edge 400 mm centered at the origin, and the TRE was computed for each registration method. This configuration (200-mm cube for the fiducials and 400-mm cube for the target) was chosen because it has been used by other authors in [7], [19], [20]. 100,000 iterations of generating true point sets X and Y and the target, generating FLE variances, perturbing X and Y, computing new registration transformations, and computing TRE values to come up with an overall RMS TRE value.

B. Comparison by Matlab simulations of the new algorithm to the closed-form solution with and without isotropic weighting, to Ohta's algorithm, and to Matei's algorithm.

We tested the new algorithm by performing computer simulations using Matlab. The same four values of N were chosen as in Section 3.A, the same protocol for choosing point sets and targets was used, and the same number of trials was run to produce the

RMS TRE value. All elements of both point sets X and Y were randomly perturbed to simulate FLE with chosen properties (described in Experiments B1, B2, and B3 below). X and Y were then registered. Five registration methods were compared: (1) the closed-form solution with W_i set to the identity matrix I for all i, (2) the isotropically weighted closed-form solution with W_i set to $\left(\sum_{\alpha=1}^{3} \sigma_{i\alpha}^{2} \right)^{-1/2} I$ for all i, (3) Ohta's algorithm, (4) Matei's algorithm, and (5) the new algorithm with W_i defined according to Eq. (3). To improve the results for Ohta's algorithm, both point sets were de-meaned. Only a rotation matrix is the result of its registration.

The new algorithm can handle arbitrary FLE for each fiducial and for each direction in the two spaces. Different experiments were performed to study the effect on TRE of using different algorithms for different FLE configurations.

Experiment B1 (homogeneous, isotropic FLE in image space and inhomogeneous, anisotropic FLE in physical space): A fiducial system with isotropic and homogeneous FLE for all fiducials in image space and different FLE in all directions for all fiducials in physical space was used for this experiment. This experiment mimics the common situation in which a pre-operative image with a small slice thickness can produce a homogeneous, isotropic FLE, while a physical tracking system produces an anisotropic FLE that is a function of the displacements between the fiducials, the camera, and possibly a CRF, and these displacements vary from fiducial to fiducial. The FLE in image space was chosen randomly from $N(0, FLE_X)$, where FLE_X was randomly chosen from the interval $[0,1]$. The FLE in physical space in the k-th direction for the i-th fiducial

was drawn from $\mathrm{N}\left(0, \mathrm{FLE}_{Yik}\right)$, where all FLE_{Yik} were randomly chosen from the interval [0,1].

Experiment B2 (homogeneous, anisotropic FLE in both image and physical spaces): This experiment mimics the common situation in which the pre-operative image with a large slice thickness produces a homogeneous, anisotropic FLE, and the displacements between the fiducials and the camera and a CRF are relatively constant from fiducial to fiducial. A fiducial system with different FLE in all directions, but equal for all fiducials in both image and physical space, was used for this experiment. The FLEs in image and physical spaces in the *k*-th direction were drawn correspondently from $\mathrm{N}\left(0, \mathrm{FLE}_{Xk}\right)$ and $\mathrm{N}\left(0, \mathrm{FLE}_{Yk}\right)$, where all FLE_{Xk} and FLE_{Yk} were randomly chosen from interval [0,1].

Experiment B3 (inhomogeneous, anisotropic FLE in both image and physical spaces): This experiment mimics the situation in which a pre-operative image whose quality varies in both direction and position (e.g., for a cone-beam CT) produces an inhomogeneous, anisotropic FLE, and (as in Experiment B1) the displacements of the fiducials from the tracking camera and a CRF vary from fiducial to fiducial. A fiducial system with different FLE in all directions for all fiducials in both image and physical spaces was used for this experiment. The FLE in image and physical space in the *k*-th direction for the *i*-th fiducial were drawn from $\mathrm{N}\left(0, \mathrm{FLE}_{Xik}\right)$ and $\mathrm{N}\left(0, \mathrm{FLE}_{Yik}\right)$, where all FLE_{Xik} and FLE_{Yik} were randomly chosen from interval [0,1].

C. Comparison of the new algorithm to Matei's algorithm using real data and by Matlab

simulations using specific fiducial configuration.

Experiment C1 (using real data from a surgical guidance application acquired with an optical tracking system): To compare the new algorithm with Matei's approach, we ran the following experiment, which has been previously described [21].

Eight retro-reflecting spheres were used as fiducial markers in physical space and were tracked in real time by a Polaris Spectra optical tracking system. No true target location is ever known exactly, thus a good estimator should be found. We used two fiducial markers frames rigidly attached to each other (Figure 1) with the centroid of four markers in one frame used as the target. The benefit of such a target is that the RMS localization error during real-time position acquisition is half that of the RMS error of an individual marker. The markers on the other frame were used for registration. Thus, in this experiment $N = 4$.
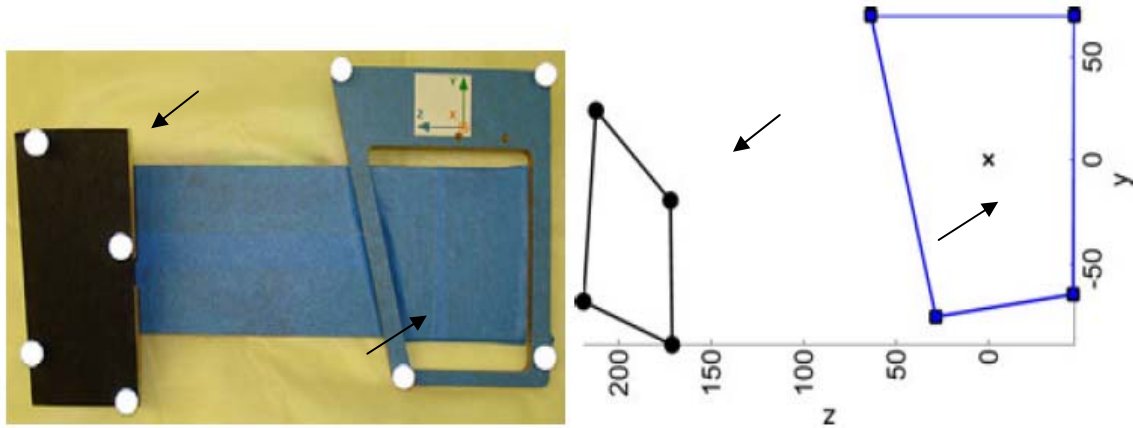

Figure 1. Photo and schematic of two rigidly attached tools. The tool whose markers were subjected to registration (a) is the one on the left. The right tool (b) was used to estimate the true location of the target, which is the centroid of the right tool (marked with x in the schematic).

Before the real-time acquisition began, a tool definition procedure was performed for the combination of the tools on the left and on the right using standard tool definition software. This procedure includes thousands of measurements of the tools' poses with an average taken over all measurements. We assume that the errors of these multiple measurements are statistically independent. Therefore, the level of FLE for each average measurement is an order of magnitude or more smaller than an individual measurement. We let Y equal the resulting average configuration of the left tool and assume that the FLE in this space is negligible. Thus, we treat the three quantities, $\sigma_{yi1}^2, \sigma_{yi2}^2, \sigma_{yi3}^2$, in Eq. (3) as being equal to zero, which is equivalent to setting W_{yi}^{-2} equal to zero in Eq. (2) and is equivalent to treating W_{yi} as infinity (specified by `Inf` in Matlab). The target is selected as the centroid of the right tool, and from the tool definition procedure the position r_0 of that target relative to Y is determined with high accuracy.

During real-time acquisition, the two rigidly attached tools are moved together within a region of about 10 cm, and both tools are tracked by the optical tracking system placed at a distance of approximately 1.5 m (recommended distance by the tracking-tool manufacturer), but the registration is performed only using the left tool, whose real-time measured configuration serves as X. Thus, X is repeatedly registered to Y. In this experiment, because the displacements between the markers and the camera do not vary appreciably from marker to marker (and there is no CRF), we assume that FLE is homogenous and anisotropic. The registration algorithms require the covariance of that FLE as input, and we used the method described in [22] to estimate that covariance. The accuracy of the registration of X to Y is then validated in terms of its agreement with movement of the measured position centroid of the right tool.

After a buffer of 200 sets of three-dimensional positions has been gathered, for every subsequent frame of measured data the following actions are performed:

1) Three-dimensional positions of all markers are gathered.

2) The target $\mathbf{r}_{ref}$ is calculated as a mean of four markers from the right tool.

3) For each of the previous 200 sets of three-dimensional positions, the tool on the left is registered (X to Y) using isotropic registration, and the algorithm from [22] is employed to estimate the covariance of FLE from the observed 200 fiducial registration errors of each fiducial

4) The estimated covariance found at Step 3 is used as input to the new algorithm to register the left tool (X to Y). The target location $\mathbf{r}_{aniso}$ is calculated by applying the found transform to $\mathbf{r}_{ref}$.

5) The estimated covariances found at Step 3 is used as input to Matei's algorithm to register the left tool (X to Y). The target location $\mathbf{r}_{matei}$ is calculated by applying the found transform to $\mathbf{r}_{ref}$.

6) The TRE vectors are calculated for the two algorithms: $\mathbf{TRE}_{aniso} = \mathbf{r}_{aniso} - \mathbf{r}_{ref}$, $\mathbf{TRE}_{matei} = \mathbf{r}_{matei} - \mathbf{r}_{ref}$.

A total of 1000 data frames are collected for this experiment. For each data frame after the first 200 buffer set, the above steps are executed and the TRE using the two algorithms are compared.

Experiment C2 (using Matlab simulations): Matlab simulations were run for a case similar to the one described in Experiment C1 to make sure that our simulations agree with real data. A simulation scheme similar to the one described in Section 3.B was

employed, but exactly the same fiducial configuration as in Experiment C1 was employed and the weights in stationary space were set equal to infinity (see above), while in moving space the FLE was homogeneous and anisotropic. The FLE in the moving space in the k-th direction was drawn from $N(0, FLE_{Yk})$, where all FLE_{Yk} were randomly chosen from interval [0,0.25]. The number of fiducials N was set to 4.

D. Comparison of the new algorithm to the closed-form solution with and without isotropic weighting, to Ohta's algorithm, and to Matei's algorithm for image-guided surgical configurations performed by Matlab simulations.

Fiducial configuration parameters that influence the results of anisotropic registration were explored. The goal of this experiment is to find out whether there are some special cases for which the new algorithm might fail. By "failing" we mean not converging at all or producing a TRE larger than the TRE using any of the following methods: (1) the closed-form solution with $W_{i\alpha\alpha}$ set to 1 for all the fiducials and directions, (2) the closed-form solution with W_i set to $\left(\sum_{\alpha=1}^{3} \sigma_{i\alpha}^2\right)^{-1/2} I$ for all i, (3) Ohta's algorithm, (4) Matei's algorithm.

The space of the parameters influencing the registration is very large. It includes FLE in all directions of each marker in two spaces, the number of markers, the size of the region from which the fiducials are selected, the size of the region from which the target is selected, and the principal axes of FLE in the two spaces. Clearly, it is impossible to carry out the complete exploration of such a large space. However, we can reduce the space by restricting our experiments to those of interest for a particular application. We have chosen surgical guidance as that application.

First, we note that if the size of the fiducial is larger than the size of the image voxel, then localization can be made highly isotropic in image space. With high quality CT and MRI scanners this becomes an easy problem. Moreover, medical images like CT and MRI are for most scanners quite homogeneous. Thus, we concentrate on homogeneous, isotropic FLE in image space. Second, in many commercial tracking systems, the distance from the camera to the fiducials is much larger than the distances between the fiducials. As a result, the displacement from the camera to the fiducials is approximately equal. This lets us make the assumption that FLE is homogeneous in physical space.

In regard to the degree of anisotropy of FLE, we note that there are three possible cases: 1) FLE is isotropic – the variances of the FLE are equal in all three directions, 2) FLE is partly anisotropic – the variances of the FLE are equal in two directions but different in the third direction, 3) FLE is fully anisotropic – the variances of the FLE are different in all three directions. For these experiments, we chose the most common scenario in image-guided surgery, which is the second case: the component of FLE in optical tracking systems is higher in the direction from the camera towards the tracked object, while the components in the two other perpendicular directions are nearly the same. In particular, the higher component is a factor of three times the other components, and those components were chosen randomly and repeatedly from the range [0.1,0.4].

We set the following constraints for the size of the regions containing the fiducials and the targets: because the typical human head can be approximated by a sphere of diameter 200 mm, we chose 200 mm by 200 mm by 100 mm cuboid region for the fiducials to represent image-guided neurosurgery applications. The region chosen for

targets is a cube with an edge of 200 mm. We also considered a more focused region in which all dimensions were reduced by half. Finally, we considered "border" cases, in which all points were chosen from the surfaces of these regions. We ran a total of 2.5 million simulations divided equally among these four scenarios.

E. Comparison of the new algorithm to the closed-form solution with and without isotropic weighting, to Ohta's algorithm, and to Matei's algorithm for nearly linear configurations performed by Matlab simulations.

There is one case that is well known to cause registration problems. It is the case in which the fiducials' positions are nearly collinear, i.e. insufficient information is available for registration in three-dimensional space. We now examine that case when all fiducials are nearly collinear. We use the same general scheme for simulations as in the previous experiments, except for the set of dimensions of the box from which we randomly draw the coordinates of the fiducials. To simulate collinear fiducials, one side of the box was set to $D = 200$ mm, and two other sides to d, where $d \leq D$ and its value is changed during the experiment. In order to characterize the degree of linearity, we introduce a heuristic measure of linearity for a fiducial configuration, which we call the linearity coefficient, l, as follows:

$$l = \frac{1 - \dfrac{d}{D}}{1 + \dfrac{d}{D}} = \frac{D - d}{D + d}.$$

This coefficient takes values from interval [0,1] for $d \leq D$. When all fiducials lie on a single line, l equals to 1. When D is equal to d, or in other words, the fiducials are chosen from a cube with an edge equal to D, the coefficient $l = 0$.

We ran linearity experiments for all configurations described in Section 3.B: 1) homogeneous, isotropic FLE in the image space and inhomogeneous, anisotropic FLE in physical space, 2) homogeneous, anisotropic FLE in image and physical spaces, 3) inhomogeneous, anisotropic FLE in image and physical spaces. The goal of this experiment is to determine how the approach to linearity affects the new algorithm's performance relative to other algorithms and to determine how linearity influences convergence.

F. Number of iterations and time required

We ran simulations to determine the number of iterations and the time required for registration. While these data are important to the feasibility of any algorithm in any application, they are most critical during real-time registration, e.g. surgical navigation. Therefore, we monitored the number of iterations and the time taken for registration for the optical-tracking application described in Section 3.C above. Furthermore, because nearly linear fiducial configurations require by far the largest number of iterations, we made similar measurements for varying levels of linearity.

For the optical-tracking application described in Section 3.C above we assumed partial anisotropic configuration wherein the FLE in two directions are the same and the FLE in the third direction equals FLE in the first two directions multiplied by a constant which we call "anisotropy coefficient". For this experiment, the anisotropy coefficient

took values from range [1,5.5]. We compare by simulations the time and the number of iterations required for the new algorithm and Matei's algorithm to converge.

For nearly linear fiducial configurations we run simulations similar to those described in Section 3.E above. We measured the time and the number of iterations to converge for the new algorithm, Ohta's algorithm, and Matei's algorithm depending on the value of the linearity coefficient.

4. Results

A. Comparison by Matlab simulations of Ohta's algorithm with the new algorithm modified to for zero translation.

Ohta's algorithm and the new algorithm were compared. In [9], Ohta's algorithm is proved to be optimal in the sense of reaching the lower bound of some level of accuracy. However, no optimality is proved for this algorithm in the case when translation is present in the transformation between the point sets. Thus, we compared the new algorithm with Ohta's optimal algorithm without translation between point sets. To make the comparison fair, we modified the new algorithm so that only rotation is searched. The results of the experiments are listed in Table 1. For all employed numbers of fiducials, RMS TRE obtained with the new algorithm are compared to RMS TRE obtained with Ohta's algorithm. The table also includes the mean and the standard deviation of the difference between TRE values of the two methods. The results in the table show that the new algorithm produces a notable improvement in TRE over Ohta's algorithm.

B. Comparison by Matlab simulations of the new iterative algorithm to the closed-form solution with and without isotropic weighting, Ohta's algorithm, and Matei's algorithm.

To compare the new iterative algorithm with closed-form solution for isotropic FLE, isotropically weighted closed-form solution for isotropic FLE, Ohta's algorithm, and Matei's algorithm, we performed Matlab simulations. We ran three experiments which differed by the properties of FLE in the two spaces. In the first experiment FLE in image space was isotropic and homogeneous, and FLE in physical space was inhomogeneous and anisotropic. After the simulations performed with the parameters configured as described in the previous section of this chapter, the new algorithm proved to be working better, giving smaller RMS TRE than the other four methods. The results of the simulations are listed in Table 2. In the second experiment FLE was homogeneous and anisotropic in both image and physical spaces. The results are listed in Table 3. The third experiment included the most general case of inhomogeneous and anisotropic FLE. The results are listed in Table 4.

The new iterative method produced an improved TRE in all experiments with all numbers of fiducials. Depending on the properties of FLE, the values of other four methods were different. For example, in the first and in the second experiment the closest results to the new algorithm were produced by the closed-form solutions. In the case of inhomogeneous and anisotropic FLE in both spaces, Matei's algorithm produced the values very close to the new algorithm. However, even in this case the new algorithm performed better. Ohta's algorithm was the worst in all cases, as might have been predicted by the fact, that it loses its optimality in the presence of a translation.

N	RMS TRE (mm)		Difference in TRE: Ohta – New (mm)	
	New algorithm	Ohta's algorithm	Mean	Std. Dev.
3	0.94609	1.06431	0.09939	0.36
4	0.71180	0.83287	0.09944	0.30
5	0.58391	0.70231	0.09905	0.26
10	0.35660	0.47331	0.09622	0.20

Table 1. TRE statistics for the modified new algorithm and Ohta's algorithm for the case when no translation is involved in the transformation.

N	New	CF	IWCF	Ohta	Matei
3	4.43255	4.62104	4.62102	4.92407	4.64518
4	1.46800	1.54842	1.54856	1.68080	1.55675
5	1.08517	1.16106	1.16161	1.31043	1.18830
10	0.62820	0.69384	0.69509	0.88736	0.78198

Table 2. RMS values of TRE (mm) for the new algorithm, closed-form (CF) solution, isotropically weighted closed-form (IWCF) solution, Ohta's algorithm, and Matei's algorithm for Experiment B1 (homogeneous, isotropic FLE in image space and inhomogeneous, anisotropic FLE in physical space).

N	New	CF	IWCF	Ohta	Matei
3	4.23894	4.55000	4.55143	4.69227	4.59405
4	1.42086	1.53653	1.54452	1.71623	1.63941
5	1.08520	1.16035	1.17155	1.34415	1.30463
10	0.65265	0.69129	0.70290	0.89194	0.89067

Table 3. RMS values of TRE (mm) for the new algorithm, closed-form (CF) solution, isotropically weighted closed-form (IWCF) solution, Ohta's algorithm, and Matei's algorithm for Experiment B2 (homogeneous, anisotropic FLE in image and physical space).

N	New	CF	IWCF	Ohta	Matei
3	4.23560	4.39486	4.39080	4.71989	4.36872
4	1.38691	1.52786	1.50594	1.57102	1.40187
5	1.00989	1.15289	1.12554	1.17779	1.02370
10	0.55087	0.69246	0.66386	0.70861	0.56667

Table 4. RMS values of TRE (mm) for the new algorithm, closed-form (CF) solution, isotropically weighted closed-form (IWCF) solution, Ohta's algorithm, and Matei's algorithm for Experiment B3 (inhomogeneous, anisotropic FLE in image and physical space).

C. Comparison of the new iterative algorithm to Matei's algorithm using real data and by Matlab simulations using specific fiducial configuration

Experiment C1: Using the algorithm described in the Section 3.C, we performed the experiment to compare the two methods. The resulting TRE comparisons are shown in Figure 2. We allowed 200 frames for initialization, thus TRE calculations started just at 200-th frame. The fiducial marker frames were translated and rotated randomly during the time of the measurements. As can be seen in the figure, the TRE of the new algorithm is always smaller than the TRE of the Matei's algorithm for this experiment with real data. (As shown in the earlier experiments Ohta's algorithm is inappropriate when translation is present. Because of the large translations involved in this experiment it performed considerably worse on this data, and is therefore omitted from these results.)

Experiment C2: Simulations were run using the scheme described in Section 3.C. In this experiment, we assumed no localization error in image space. In physical space, the FLE was considered homogeneous and anisotropic. The RMS TRE values achieved using the new algorithm and Matei's algorithm were 0.26625 mm and 0.32685 mm respectively. These values agree with the real data results from Experiment C1, which are shown as horizontal lines in Figure 2. This agreement suggests that our simulations results are reliable predictors for real applications.

D. Comparison of the new algorithm to the closed-form solution with and without isotropic weighting, to Ohta's algorithm, and to Matei's algorithm for image-guided surgical configurations performed by Matlab simulations.
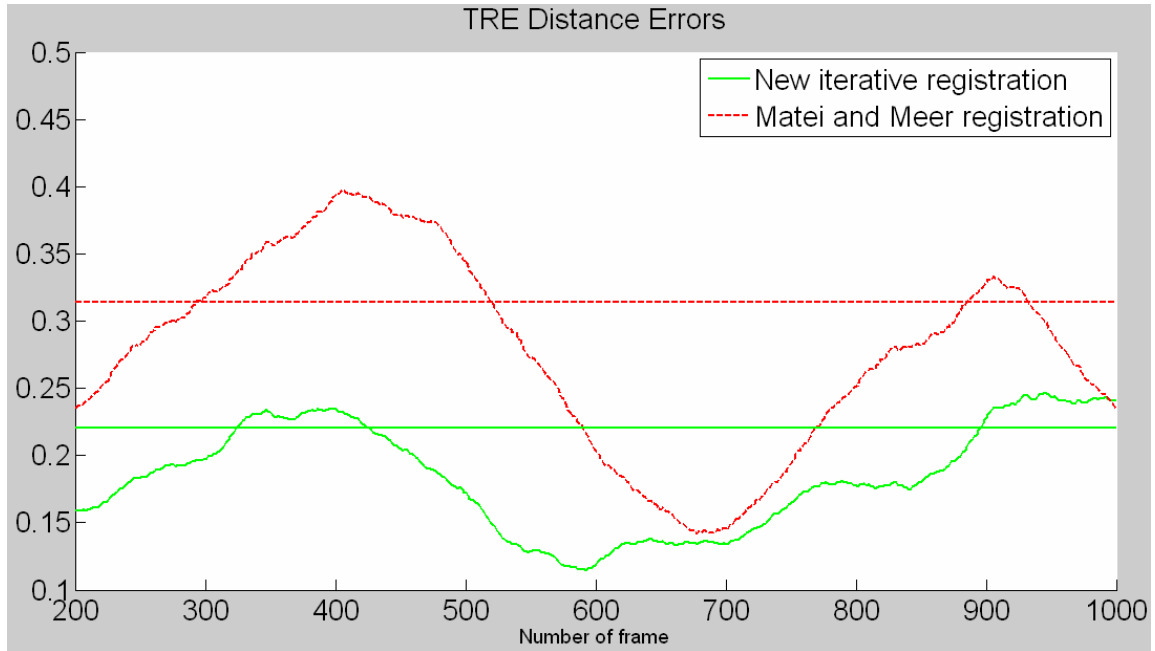
Figure 2. Results of Experiment C1: Comparison of the new algorithm with the Matei's algorithm using data collected using an optical tracking system. Horizontal lines represent the RMS TRE value over 800 frames.

The object of this evaluation was to determine whether there are cases in typical surgical guidance applications for which the closed-form solution, the isotropically weighted closed-form solution, Ohta's algorithm, or Matei's algorithm perform better than the new algorithm. Results of all the simulations showed that the new algorithm performs better than the other algorithms.

E. Comparison of the new algorithm to the closed-form solution with and without isotropic weighting, to Ohta's algorithm, and to Matei's algorithm for nearly linear configurations performed by Matlab simulations.

The results of the three experiments are shown in Figures 3 through 5. The new algorithm performs better than the other four algorithms for all values of the linearity

coefficient for which RMS TRE is ≤ 20 mm. Other algorithms produced lower TRE only for a very large linearity. In all three experiments, Ohta's algorithm, when it converges, performed best for almost linear configurations with $l \geq 0.97$. Matei's algorithm produced a smaller RMS TRE than the new algorithm in the experiment with anisotropic and inhomogeneous FLE in two spaces for $l \geq 0.93$. The closed-form solution with and without isotropic weightings performed better than the new algorithm for the experiment with homogeneous, isotropic FLE in image space and inhomogeneous, anisotropic FLE in physical space for $l \geq 0.99$. However, for these highly linear fiducial configurations for all experiments and for all methods, TRE is so large that none of the methods would provide reasonable guidance for any applications, especially clinically-relevant applications. Thus, we can conclude that for all reasonable fiducial configurations that can achieve at least moderate values of TRE (TRE < 10 mm) the new algorithm performs better than the four other algorithms.

The other goal of this experiment was to determine the effect of linearity on the convergence of the iterative algorithms. As explained above, we determined that convergence occurred only if the threshold was reached in no more than 1000 iterations. For all three experiments the new algorithm failed to converge only for linearity coefficient $l \geq 0.99$. For l = 0.99 the new algorithm failed to converge for 1.5% of registrations. For comparison, for the same linearity coefficient Matei's algorithm failed to converge for 14.4% of registrations and Ohta's algorithm failed to converge for 63% of registrations. These high failure percentages again reiterate that almost linear fiducial configurations are not applicable to providing guidance via registration.

Figure 3. RMS values of TRE (mm) for the new algorithm, closed-form solution, isotropically weighted closed-form solution, Ohta's algorithm, and Matei's algorithm with varying linearity coefficient in the fiducial configuration. Homogeneous, isotropic FLE in the image space and inhomogeneous, anisotropic FLE in the physical space were used for this experiment.

Figure 4. RMS values of TRE (mm) for the new algorithm, closed-form solution, isotropically weighted closed-form solution, Ohta's algorithm, and Matei's algorithm with varying linearity coefficient in the fiducial configuration. Homogeneous, anisotropic FLE was assumed in both image and physical spaces.
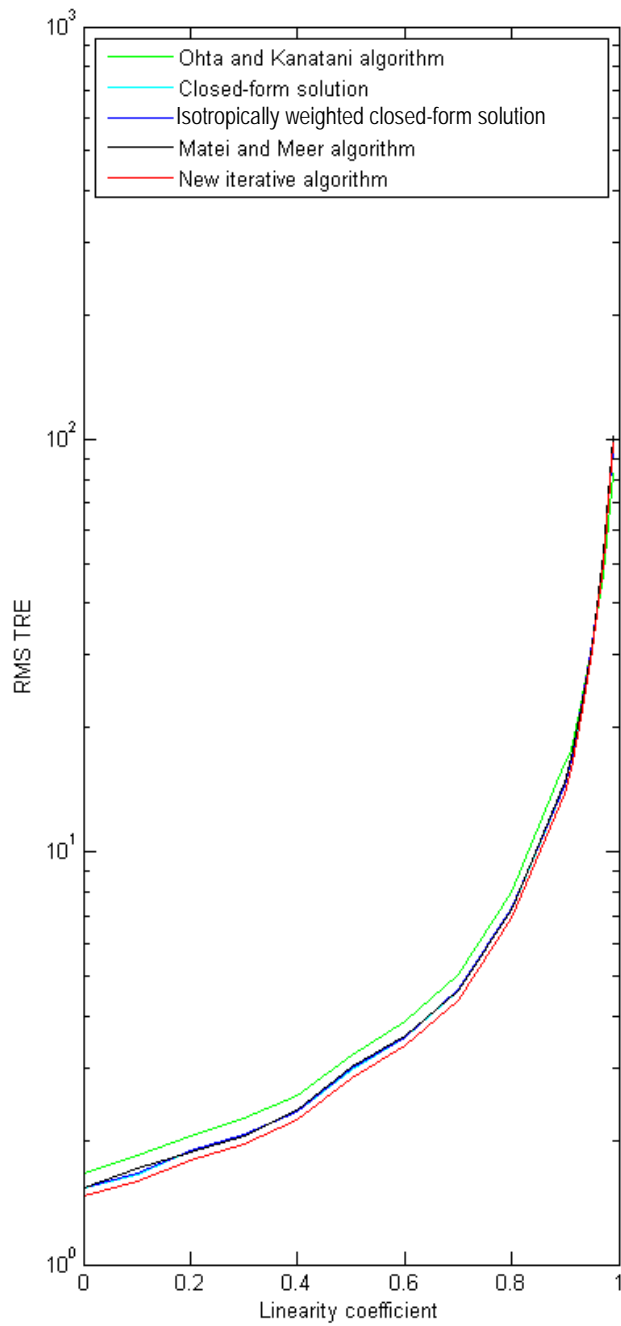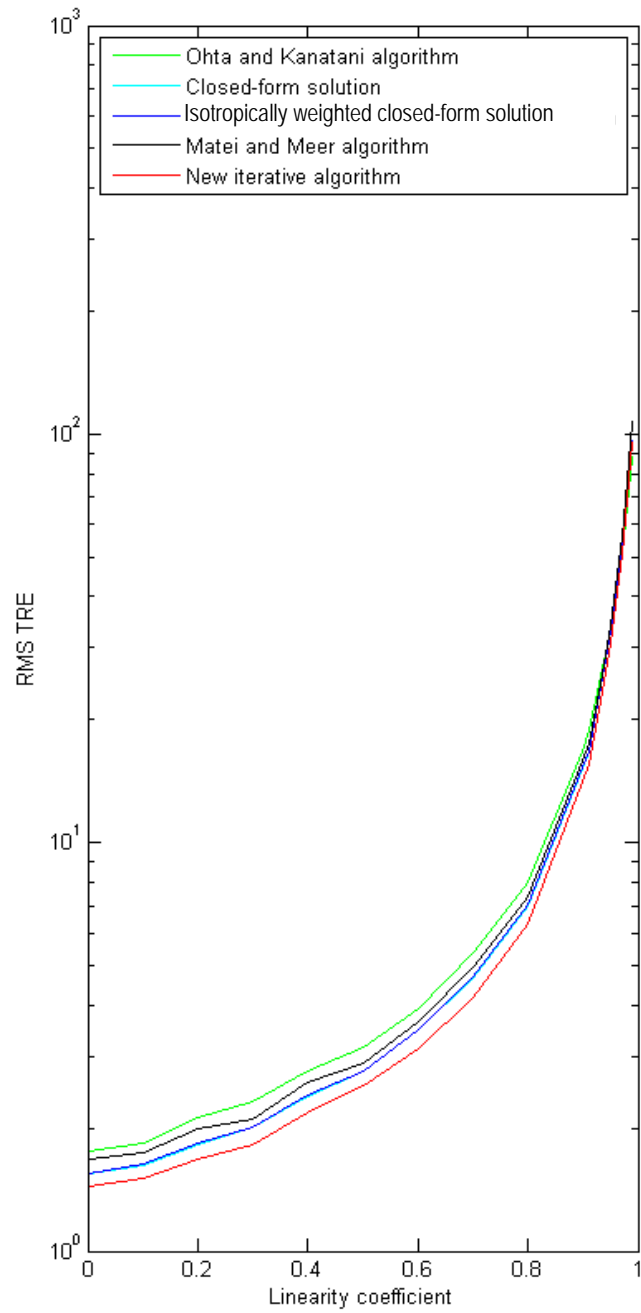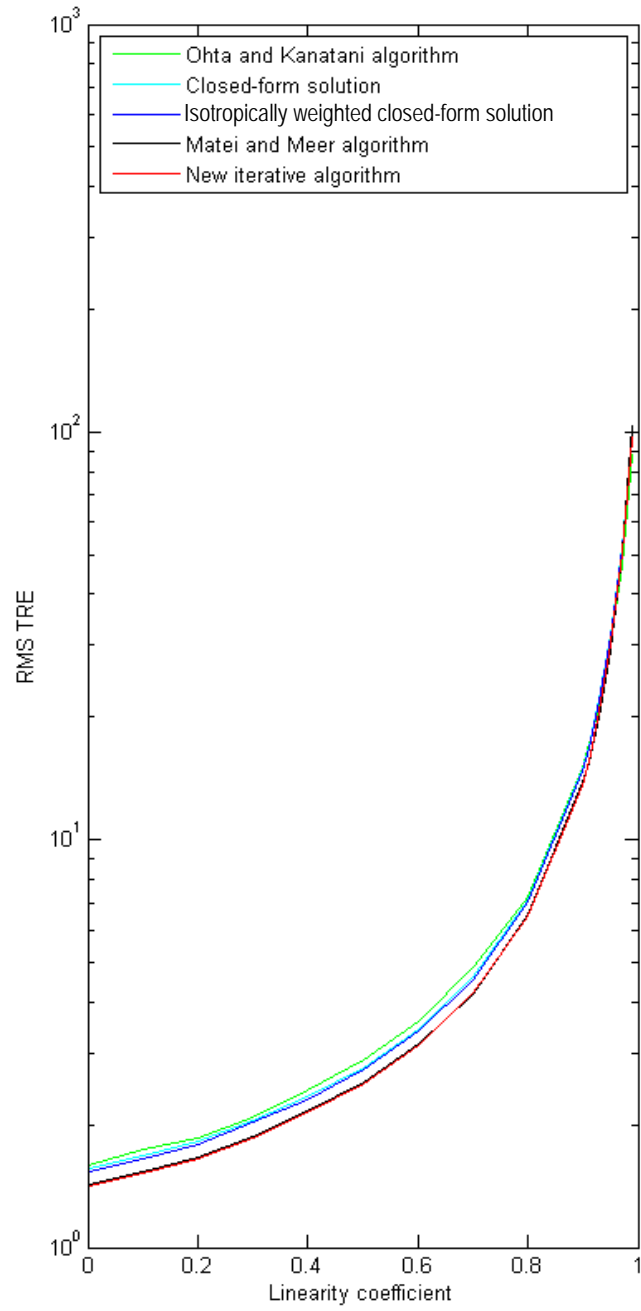
Figure 5. RMS values of TRE (mm) for the new algorithm, closed-form solution, isotropically weighted closed-form solution, Ohta's algorithm, and Matei's algorithm with varying linearity coefficient in the fiducial configuration. Inhomogeneous, anisotropic FLE was assumed in both image and physical spaces.

F. Number of iterations and time required

The results of the experiments for the optical-tracking application are presented in Figure 6 and Figure 7. These figures show with the increasing anisotropy of the fiducial system the number of iterations to converge increases and correspondingly the time for convergence also increases. The new algorithm requires more iterations to converge than Matei's algorithm. However, the average speed of executing each iteration of the new algorithm is higher than that of Matei's algorithm, and hence the overall time of registration is smaller for the new method.

Figure 6. The average number of iterations before a registration algorithm converged for the new algorithm and for Matei's algorithm as a function of the value of the anisotropy coefficient.

Figure 7. The average time of registration for the new algorithm and for Matei's algorithm as a function of the value of the anisotropy coefficient.

Figure 8 and Figure 9 present the results of the simulations for fiducial configurations that are nearly linear. The number of iterations to converge and the time of registration increase when the linearity coefficient increases. The new algorithm converges after larger number of iterations compared to Ohta's algorithm and Matei's algorithm. However, the overall time of registration using the new algorithm is comparable with the overall time using Matei's algorithm and smaller than Ohta's algorithm registration time.
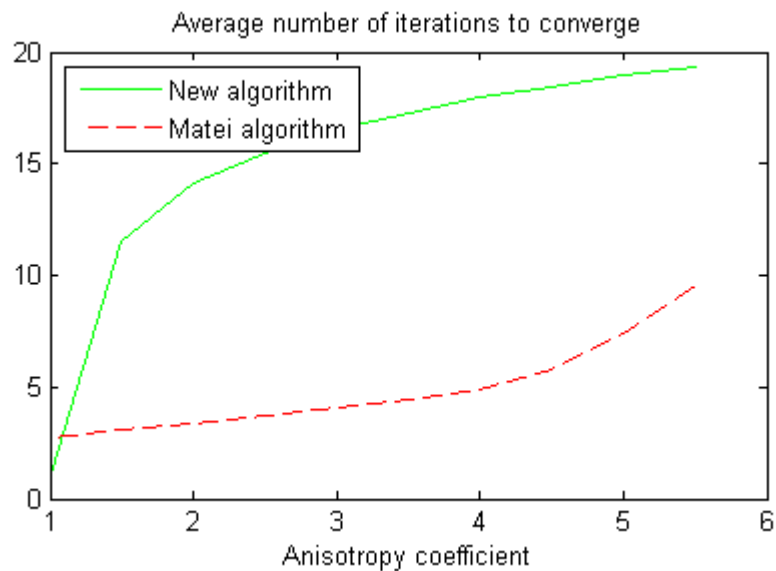
Figure 8. The average number of iterations before a registration algorithm converged for the new algorithm, Ohta's algorithm and Matei's algorithm as a function of the linearity coefficient value.
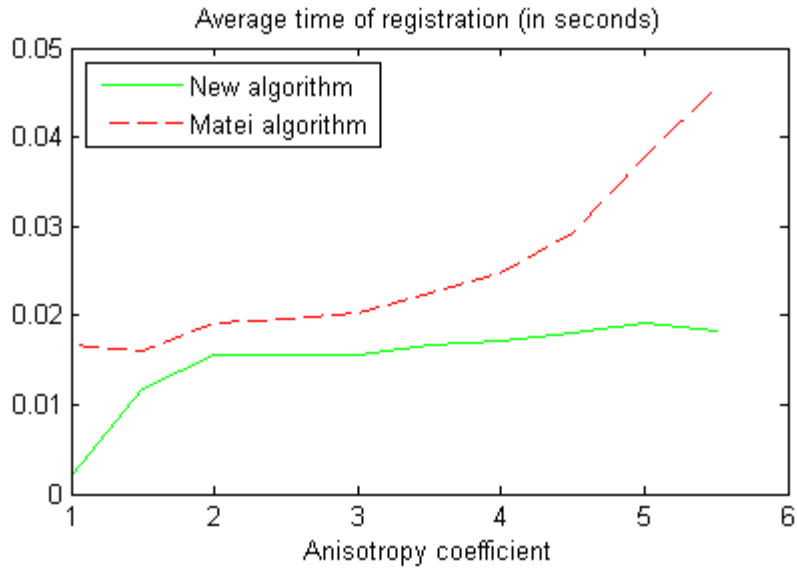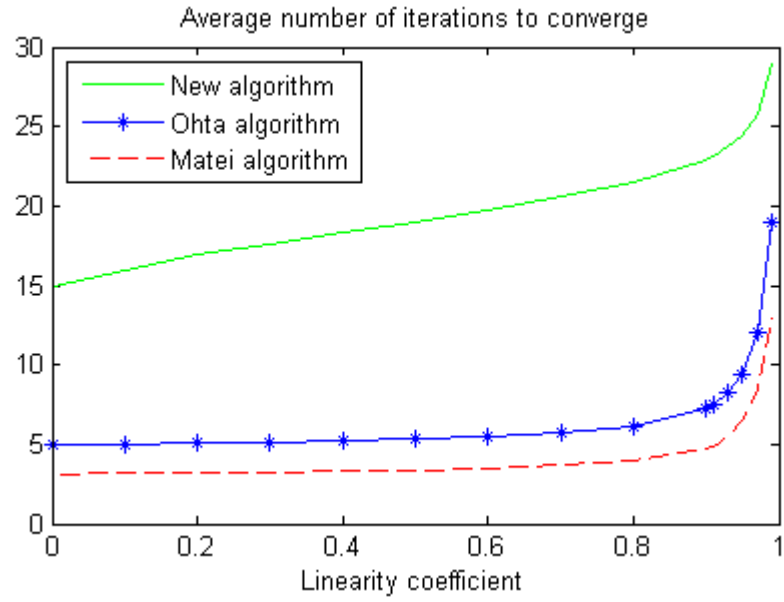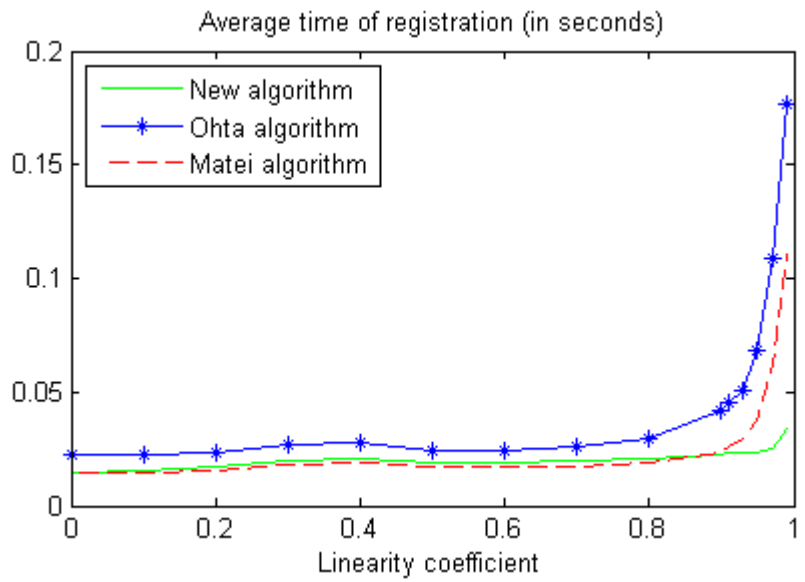

Figure 9. The average time of registration for the new algorithm, Ohta's algorithm and Matei's algorithm as a function of the linearity coefficient value.

5. Discussion

We presented in this chapter a new and relatively simple iterative algorithm for solving the point-based registration problem when anisotropic weightings are used to compensate for anisotropic FLE. In contrast to the existing solutions, the new algorithm is more intuitive, easy to understand and implement, and has no adjustable parameter other than a stopping criterion. Through simulations we have shown a substantial improvement in TRE when the new algorithm is employed for fiducial systems with anisotropic FLE instead of the commonly used closed-form solutions. We also compared our algorithm with two other iterative methods that have no adjustable parameters other than a stopping criterion. One of them, Ohta's algorithm [9], is designed only for registrations in which only rotation and no translation is involved in the point sets transformation. The second method is Matei's algorithm [10]. Our simulations showed that the new algorithm improved the TRE compared to these two methods. An especially large increase in accuracy compared to Matei's algorithm is noted when FLE is homogeneous or isotropic in even just one space. The two algorithms show the closest results when transformation is zero or extremely small. The explanation that these algorithms do not produce the best results is that they assume transformations that are very close to the identity. When the difference from the identity is negligible, both Ohta's algorithm and Matei's algorithm produce nearly the same results as the new method. However, even for those cases, the new algorithm almost always performs notably better.

Beside the simulations, we used real data from a surgical guidance application acquired with an optical tracking system. For acquiring the experimental data we used the

same approach as that used in the experiment reported in [21]. A combination of two rigidly attached frames and highly accurate measurements of the configuration of fiducials on the frames allowed us to assume that in one space FLE is zero. In the other space FLE was assumed to be homogeneous and anisotropic. With those assumptions we ran point registrations using our new algorithm and Matei's algorithm. The results of this experiment showed that the new algorithm provides much higher accuracy.

To make sure that our simulations correspond to real registrations, we ran Matlab simulations with a fiducial configuration similar to that of the experiment with the optical tracking system and defined FLE statistics to be approximately the same as in that observed in the experiment. These simulations produced results comparable to the results received after registering point sets with the real data. These comparable results suggest that all our simulations can be trusted as predictors of experiments with real data.

We explored the space of all parameters influencing the results of registration using the new method to determine whether there are some fiducial and FLE configurations for which the algorithm does not converge or performs worse than other algorithms. We limited our research just to the case of surgical guidance, which limited the fiducial configurations and their FLE statistics, and ran Matlab simulations. The new algorithm performed better than the other algorithms in those cases.

We also considered the case in which the fiducial coordinate vectors are nearly collinear. For this case other algorithms were superior near the limit of complete linearity, where RMS TRE exceeds 20 mm, but we found that for fiducial systems with linearity small enough to achieve a TRE small enough for use with surgical systems, our new method always performs better than the other methods. Finally, we found that for all such

reasonable fiducial systems our method always converged (used fewer than 1000 iterations).

Finally, we explored the number of iterations and the required time for the algorithms. We found that with the increased anisotropy or increased linearity of the fiducial system, the number of iterations to converge and the overall time of registration increase. Comparing the new algorithm to Ohta's and Matei's algorithms, the new algorithm takes more iterations to converge. However, the time of executing each iteration is much smaller for the new algorithm, so that the overall time of execution of the new algorithm is not higher than the time of performing registration using Ohta's or Matei's algorithms.

In conclusion, by means both of simulations and of real data obtained with an optical tracking system, we have shown that the new algorithm increases the accuracy of point registration in comparison to the standard closed-form solution, the isotropically weighted closed-form solution, Matei's algorithm, and Ohta's algorithm for all cases feasible during surgical guidance.

6. Conclusion

A new iterative algorithm for performing point-based registration in the presence of arbitrary (homogeneous or inhomogeneous, isotropic or anisotropic) FLE in two spaces was introduced. This new method is more intuitive than existing iterative methods and it has no adjustable parameters other than its stopping criterion. By means of simulations and by using real data acquired with an optical tracking system, we showed

that the new method produces smaller TRE. Our simulations were performed for cases that are likely to happen in surgical applications. We also compared the results of the registration of real data with registration of fiducial configuration with similar characteristics defined in simulations. The results were comparable which suggests that our simulations provide a reliable predictor for real situations.

We compared the new algorithm to four existing solutions: the standard closed-form solution with equal weighting for all fiducials, the closed-form solution with isotropic weightings, Ohta's algorithm [9], and Matei's algorithm [10]. The latter two algorithms are iterative and, like the new algorithm, involve no adjustable parameters other than their stopping criteria. Except for fiducial configurations that are far too linear to be useful, the new algorithm produced lower TRE than each of the other four algorithms in all cases tested. Comparing the new algorithm to the two closed-form methods, the maximum increase in accuracy is achieved when FLE is both anisotropic and inhomogeneous in at least one of the two spaces. Comparing the new algorithm with Ohta's and Matei's algorithms, the maximum increase in accuracy is noted when FLE is homogeneous or isotropic (or both) in at least one space. Conversely, the accuracy of the closed-form solution with and without anisotropic weightings is closer to that of the new method when FLE is isotropic and homogeneous. Ohta's algorithm and Matei's algorithm are very close to the new algorithm in the presence of anisotropic and inhomogeneous FLE in both spaces.

We analyzed the reliability of our new method with regard to the collinearity of the fiducial system. By means of Matlab simulations we showed that for realistic fiducial systems that produce a useful registration, our method performs better than the four other

tested methods. Moreover, the simulations demonstrated our new method's convergence

for such fiducial systems.

7. Appendix

The following are Matlab functions that implement the algorithm presented in this

chapter:

```
function [R,t,FRE,n,isConverged] =
anisotropic_point_register(X,Y,Wx,Wy,threshold)
% X is the moving set, which is registered to the static set Y. Both are 3
% by N, where N is the number of fiducials. Wx and Wy are a 3-by-3-by-N
% arrays, with each page containing the weighting matrix for the Nth pair
% of points in spaces X and Y. THRESHOLD is the size of the change to the
% moving set above which the iteration continues.
%
% R and t are found registration parameters, FRE is a weighted fiducial
% registration error, n is the number of performed iterations, isConverged
% shows if the method converged
%
% Creation:
% A. Danilchenko, R. Balachandran and J. M. Fitzpatrick
% December 2010
if nargin<2
    error('X and Y must be given as input');
end
if size(X,1)~=3 && size(Y,1)~=3
    error('X and Y must be 3 by N.');
end
N = size(X,2);
if size(Y,2)~=N
    error('X and Y must have the same number of points.');
end

isConverged = 1;

% Initial estimate of the transformation assumes anisotropy:
[R,t,FRE] = point_register(X,Y);
if nargin<3 % if W not given, then give the isotropic solution
    n = 0;
    return
end
if nargin<4
    threshold = 1e-6;
end
n = 0; % iteration index = 0;
config_change = Inf;
```

```
Xold = R*X+repmat(t,1,N);
W = zeros(3,3,N);
while (config_change>threshold)
    if n > 1000
        isConverged = 0;
        return;
    end
    n = n+1;
    for j1 = 1:N
        W(:,:,j1) = (R * Wx(:,:,j1)^(-2) * R' + Wy(:,:,j1)^(-2))^(-0.5);
    end
    C = C_maker(Xold,W);
    e = e_maker(Xold,Y,W);
    q = C\e;
    if n > 1
        q = (q + oldq)/2;
    end %damps oscillations
    oldq = q;
    delta_t = [q(4) q(5) q(6)]';
    delta_theta = [1 -q(3) q(2); q(3) 1 -q(1); -q(2) q(1) 1];
    [U,~,V] = svd(delta_theta);
    delta_R = U*V';
    R = delta_R*R; % update rotation
    t = delta_R*t+delta_t; % update translation
    Xnew = R*X+repmat(t,1,N); % update moving points
    config_change = sqrt(sum(sum((Xnew-Xold).^2))/...
    sum(sum((Xold-repmat(mean(Xold,2),1,N)).^2)));
    Xold = Xnew;
end
for ii = 1:N
    D = W(:,:,ii)*(Xnew(:,ii)-Y(:,ii));
    FRE(ii) = D'*D;
end
FRE = sqrt(mean(FRE));


function C = C_maker(X,W)
N = size(X,2);
C = zeros(3*N,6);
for ii = 1:N
   X0ii = [        0,  X(3,ii), -X(2,ii);
            -X(3,ii),        0,  X(1,ii);
             X(2,ii), -X(1,ii),        0];
   C(3*ii-2:3*ii,:) = [W(:,:,ii)*X0ii, W(:,:,ii)];
end


function e = e_maker(X,Y,W)
N = size(X,2);
e = zeros(3*N,1);
for ii = 1:N
   e(3*ii-2:3*ii) = W(:,:,ii)*(Y(:,ii)-X(:,ii));
end


function [R, t, FRE] = point_register(X, Y)
% This function performs point-based rigid body registration with equal
```

```
% weighting for all fiducials. It returns a rotation matrix,
% translation vector and FRE.
if nargin < 2
    error('At least two input arguments are required.');
end
[Ncoords Npoints] = size(X);
[Ncoords_Y Npoints_Y] = size(Y);
if Ncoords ~= 3 | Ncoords_Y ~= 3
    error('Each argument must have exactly three rows.')
elseif (Ncoords ~= Ncoords_Y) | (Npoints ~= Npoints_Y)
    error('X and Y must have the same number of columns.');
elseif Npoints < 3
    error('X and Y must each have 3 or more columns.');
end
Xbar = mean(X,2);  % X centroid
Ybar = mean(Y,2);  % Y centroid
Xtilde = X-repmat(Xbar,1,Npoints); % X relative to centroid
Ytilde = Y-repmat(Ybar,1,Npoints); % Y relative to centroid
H = Xtilde*Ytilde';  % cross covariance matrix
[U S V] = svd(H);    % U*S*V' = H
R = V*diag([1, 1, det(V*U)])*U';
t = Ybar - R*Xbar;
FREvect = R*X + repmat(t,1,Npoints) - Y;
FRE = sqrt(mean(sum(FREvect.^2,1)));
```

8. References

[1] Schönemann, P.H. A generalized solution of the orthogonal Procrustes problem, Psychometrika, 31, 1–10, 1966.
[2] Farrell, J.L. and Stuelpnagel, J.C. Problem 65–1: A least squares estimate of satellite attitude, SIAM Rev 8, 384–386, 1966.
[3] West, J.B. and Maurer, C.R., Jr. Designing optically tracked instruments for image-guided surgery, IEEE Transactions on Medical Imaging, 23, 533-545, 2004.
[4] Koschat, M.A. and Swayne, D.F. A weighted Procrustes criterion, Psychometrika, 56(2), 229-239, 1991.
[5] Chu, M.T. and Trendafilov, N.T. On a differential equation approach to the weighted or orthogonal procrustes problem, Statistics Computing, 8, 125–133, 1998.
[6] Balachandran, R., Fitzpatrick, J.M., and Labadie, R.F. Fiducial registration for tracking systems that employ coordinate reference frames, Proc. SPIE Medical Imaging 2005, San Diego, CA, 5744, 134–145, 2005.
[7] Moghari, M.H. and Abolmaesumi, P. Point-based rigid-body registration using an unscented Kalman filter, IEEE Transactions on Medical Imaging, 26(12), 1708–1728, 2007.
[8] Pennec, X. and Thirion, J.P. A framework for uncertainty and validation of 3D registration methods based on points and frames, Int. J. Comput. Vis., 25(3), 203–229, 1997.

[9] Ohta, N. and Kanatani, K. Optimal estimation of three-dimensional rotation and reliability evaluation, Computer Vision - ECCV 98, H. Burkhardt, B. Neumann, eds., Lecture Notes in Computer Science, Springer, 175—187, 1998.

[10] Matei, B. and Meer, P. Optimal rigid motion estimation and performance evaluation with bootstrap, IEEE Comput. Vis. Pattern Recognit, 339–345, 1999.

[11] Fitzpatrick, J.M., Hill, D.L.G., and Maurer, C.R., Jr. Registration, Medical Image Processing, Volume II of the Handbook of Medical Imaging, M. Sonka and J.M. Fitzpatrick, eds., SPIE Press, 2000.

[12] Danilchenko, A. and Fitzpatrick, J.M. General approach to first-order error prediction in rigid point registration, IEEE Transactions on Medical Imaging, 30(3), 679-693, 2010.

[13] West, J.B., Fitzpatrick, J.M., and Batchelor, P.G. Point-based registration under a similarity transform, Proc. SPIE Medical Imaging 2001, San Diego, CA, 4322, 611–622, 2001.

[14] Moghari, M. and Abolmaesumi, P. Distribution of target registration error for anisotropic and inhomogeneous fiducial localization error, IEEE Transactions on Medical Imaging, 28(6), 799-813, 2009.

[15] Moghari, M. and Abolmaesumi, P. Distribution of fiducial registration error in rigid-body point-based registration, IEEE Transactions on Medical Imaging, 28(11), 1791-1801, 2009.

[16] Candocia, F.M. Jointly registering images in domain and range by piecewise linear comparametric analysis, IEEE Transactions on Image Processing, 12(4), 409–419, 2003.

[17] Friston, K.J., Ashburner, J., Frith, C.D., Poline, J.-B., Heather, J.D., and. Frackowiak, R.S.J Spatial Registration and Normalization of Images, Human Brain Mapping, 2, 165-189, 1995.

[18] Golub, G.H. and Van Loan, C.F. Matrix Computations, Baltimore, MD: Johns Hopkins University Press, 1983.

[19] Fitzpatrick, J.M., West, J.B., and Maurer, C.R., Jr. Derivation of expected registration error for rigid-body, point-based, image registration, Proc. SPIE Medical Imaging 98, San Diego, CA, 3338, 16–27, 1998.

[20] Balachandran, R. Measurement and analysis of error in rigid-body point-based registration systems, Ph.D. dissertation, Vanderbilt University, 2008.

[21] Danilchenko, A., Wiles, A.D., Balachandran, R., and Fitzpatrick, J.M. Improved method for point-based tracking, MICCAI 2010 Lecture Notes in Computer Science, 6363, 587-594, 2010.

[22] Wiles, A.D. and Peters, T.M. Real-time estimation of FLE statistics for 3D tracking with point-based registration, IEEE Transactions on Medical Imaging, 28(9), 1384–1398, 2009.

CHAPTER IV

IMPROVED METHOD FOR POINT-BASED TRACKING

This chapter is adapted from the following paper:

- Danilchenko, A., Wiles, A.D., Balachandran, R., and Fitzpatrick, J.M. Improved method for point-based tracking, MICCAI 2010 Lecture Notes in Computer Science, 6363, 587-594, 2010.

1. Introduction

Image-guided surgery (IGS) systems have improved the standard of care in brain, spine, and orthopedic interventions by combining pre-operative medical images and virtual reality using spatial tracking technologies [1]. Recently, real-time imaging techniques, such as ultrasound and endoscopy, and robotics technology have been integrated with traditional IGS systems by tracking the imaging and robotic devices similarly to tracked surgical tools. The real-time images and robotic devices are displayed in the virtual reality environment resulting in an augmented view of the surgical target. One such application is the robotic drilling system for a mastoidectomy [2] where a section of bone is resected from behind the ear of a patient for various otolaryngology procedures. In this application, an optical tracking system (NDI Polaris Spectra, Waterloo, ON, Canada) is used to track the poses of both the patient and the robot. If the patient moves during the drilling procedure, the drilling plan is automatically updated to reflect the change in patient positioning using the real-time information provided by the optical

tracking system. This procedure serves as an example of applications in which a high degree of tracking accuracy is crucial to success. Such applications motivate the focus of this chapter, which is the development and validation of an improved method for highly accurate optical tracking.

A pose is determined by means of point registration: the configuration of N markers measured repeatedly in real time by the tracking system is registered to a model of the configuration. The method presented in this chapter is the combination of two recently developed algorithms: (i) a method to estimate the anisotropic fiducial localization error (FLE) at each of the optically tracked markers [3], which we have modified to improve its numerical stability and (ii) the registration procedure studied in Chapter III, which allows for anisotropic weighting at each of the markers [4]. We show that the FLE covariance statistics can be used as an anisotropic weighting function in the registration procedure. Accuracy improvements of $20 - 45\%$ are demonstrated and this improvement is deemed important for applications that rely on highly accurate real-time tracking such as fully automatic robotic mastoidectomy.

2. Method

The goal of this chapter is to demonstrate that the anisotropic weighted point-based registration algorithm [4] provides better results than the standard isotropic point-based registration method [5], [6]. The comparison is done using the target registration error (TRE) statistics for both registration methods.

Computing the real-time weightings

The weighting for the anisotropic registration algorithm at each marker is determined using a modification of the FLE estimates from the algorithm presented in [3]. The distance between the markers is much smaller than the distance from the centroid of the markers and the tracking system. Since the markers are identical, the FLE covariance matrix is a function only of the displacement of the markers from the tracking system, and therefore we assume that the FLE covariance matrices are the same for all markers. Thus, while FLE is expected to be anisotropic, it is assumed to be homogeneous. The tracking system captures a sequence of "frames", where a frame consists of the *x,y,z* coordinates of the *N* markers on a tracked rigid tool captured within a time short enough such that motion can be ignored (requires that the tool is moved sufficiently slowly by hand or by robot). The FLE covariance matrix is found in [3] by solving a set of six linear equations that relates the FLE statistics to the estimated FRE statistics. For each marker $i = 1,...,N$, there is one set of such equations: $A_i x = b_i$ where A_i is a 6-by-6 matrix determined by the configuration of the markers on the tracked rigid body, x is a column vector whose elements are the six independent FLE covariance components and b_i is a column vector containing the six independent fiducial registration error (FRE) components estimated from the point registrations of the previous *M* frames to the model. For an optical tool with *N* tracked markers, there are *N* sets of such equations, each of which can be solved in order to obtain an estimate of the FLE covariance matrix. However, to improve numerical stability we modify this procedure by solving an over-determined set of equations for a single FLE covariance matrix obtained by stacking the matrices and vectors such that

$$x_{avg} = \left(A_{stack}^T A_{stack} \right) A_{stack}^T b_{stack} , \tag{1}$$

where

$$A_{stack} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_N \end{bmatrix} \text{ and } b_{stack} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} . \tag{2}$$

Taking the six independent FLE covariance components from x_{avg} (with or without stacking) and rewriting them as a FLE covariance matrix Σ, the weighting matrix elements are computed as follows:

$$W = \Sigma^{-1/2} . \tag{3}$$

Alternative FLE estimation scheme

A slight disadvantage of the FLE estimation algorithm discussed above is the assumption of homogeneity of FLE. Though FLE for typical tracking systems this assumption is valid, it would be useful to be able to estimate FLE without making that assumption. We consider the estimation of inhomogeneous FLE from FRE using Eq. (33) from Chapter II. This equation is quite general. Unlike the FLE estimation method above, it accounts for inhomogeneous FLE and even allows for both nonzero inter-fiducial FRE covariance, and nonzero inter-fiducial FLE covariance. When, as in this case, inter-

fiducial FLE covariance is known to be zero, there are more equations- $\left(3N\right)^{2}$ --

represented by Eq. (33) than unknowns—6N. For example, for $N = 4$ fiducials, there are

144 equations and 24 unknowns (6N). The solution of such a system for FLE may be

approached with a least-square solution (e.g., via the pseudo-inverse or via Matlab's

backslash operator). However, because the sum on the right side of Eq. (33) goes only

from 7 to 3N, some information for FLE is lost, making the system underdetermined for

$N < 5$. For example, for $N = 4$, the rank is only 18 (i.e., less than 24). However, with the

assumption of homogeneity of FLE, this more general equation can be solved for FLE

when $N > 3$.

For this alternative FLE estimation method we ran a limited number of

experiments using the assumption of homogeneity with $N = 4$ to compare this scheme

with the FLE estimation algorithm described earlier in this section.

Obtaining a good estimate of the true target location

In the work previously presented in the literature, the algorithms were tested using

only Monte Carlo simulations, whereby the true location of the target was known exactly.

However, experimentally the true target location is never known exactly and therefore a

method of estimating the true target location is needed so that the TRE can be computed.

Here we use two rigidly connected optically tracked tools, as shown in Figure 1, to solve

this problem.

In Figure 1, the four markers on the left represent tracking markers rigidly

attached to a tracked tool (e.g., drill or hand-held pointer tool). The four markers on the

right are arranged so that, for the robotic application, the centroid (denoted by × in the

schematic) is at the tip of the drill and for the hand-held-tool application the centroid is at the tip of the probe pointer. Previous to the frame acquisitions a standard "tool definition" calibration was carried out to determine the model. The robotic arm was held stationary in several poses while the positions of the tool markers on the left and the target markers on the right were measured repeatedly (1000s of times). The average of these measurements over the various poses for each marker provides a highly accurate standard configuration, which serves as the model. The centroid of the four markers on the right is defined to be the "target" position (e.g., drill tip, or probe pointer tip). Then, during the tracking experiments (see Step 4 below), each detected configuration of the tool is registered to the standard tool. Therefore, using the right-hand rigid body to estimate the true location of the target, the target is estimated with target localization error (TLE) statistics of
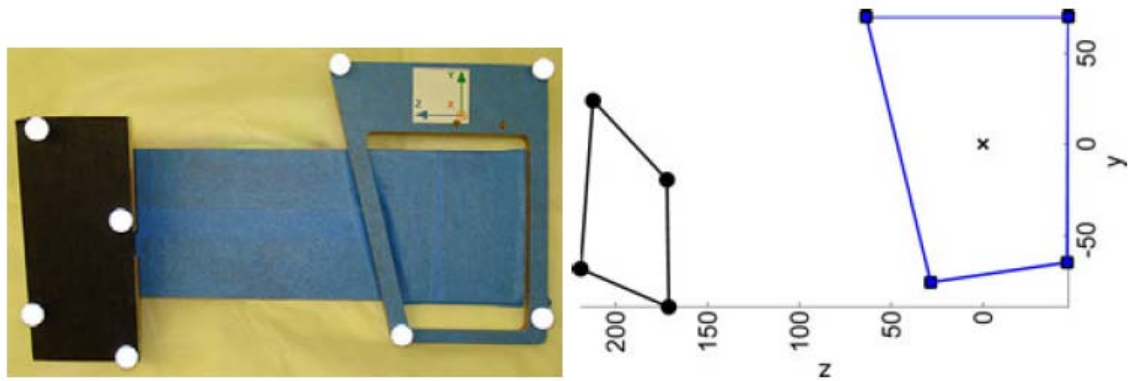
Figure 1 Photograph and schematic of the tracked rigid bodies used in the experiment. The tracked tool under test is on the left and the tracked tool used to estimate the true location of the target is on the right. The target is the centroid of the right-hand tool, which is marked by an × in the schematic. Measurements are in millimeters and coordinate frame is the local tool coordinate frame.

$$\Sigma_{tle} = \frac{\Sigma_{fle,1} + \Sigma_{fle,2} + \ldots + \Sigma_{fle,N}}{N^2}, \quad RMS_{tle} = \sqrt{trace(\Sigma_{tle})}, \quad (4)$$

135

where N is the number of markers on a tool ($N = 4$ for our example) and RMS is the root-mean-square statistic. If the FLE is indeed homogeneous across the markers, then the covariance matrix and RMS reduce to $\Sigma_{tle} = \Sigma_{fle}/N$ and $RMS_{tle} = RMS_{fle}/\sqrt{N}$, respectively.

Comparing isotropic and anisotropic registrations

The TRE statistics of the isotropic and anisotropic registrations algorithms are compared using the centroid of the right-hand tracked tool as the target ground truth. The two rigid bodies are rigidly attached to one another and moved together.

The experimental protocol is to carry out these steps for every frame of data returned from the tracking system:

1. The 3D positions of each of the markers on both rigid bodies are measured with the optical tracking system at an instance in time—a frame of data.

2. The target location, r_{ref}, is obtained by taking the mean of the four markers on the right-hand tracked tool.

3. The right-hand tracked tool is registered using isotropic registration and the FLE estimates are updated for these markers. The TLE statistics of the target are obtained using Eq. (4).

4. The left-hand tracked tool is registered using isotropic registration and the FLE estimates are updated for these markers. The target location, r_{iso}, is computed using the transform computed from isotropic registration.

5. The left-hand tracked tool is registered using anisotropic registration using the FLE estimates of the markers for the weighting as per Eq. (3) based on isotropic registrations of the $M = 200$ previous frames. The target location, r_{aniso}, is computed using the transform computed from anisotropic registration.

6. The TRE vectors for both registration methods are computed such that $tre_{iso} = r_{iso} - r_{ref}$ and $tre_{aniso} = r_{aniso} - r_{ref}$.

After 1000 frames of data are collected, the results are plotted and a set of observational statistics is computed. The results section provides a comparison of the algorithms for different dynamic paths over which the tool traveled, and also provides examples of the FLE estimates for the markers obtained with the algorithm in [3].

3. Results

Using the method described in Section 2, we compared the new tracking method, performed using anisotropic registration [4], with traditional tracking method, performed using isotropic registration [5], [6]. The tracked tools were placed at a distance of approximately 160 cm from the tracking system. We noticed that the results of the comparisons are distinct for different types of motion, thus we collected data under the following conditions:

Test A: translate approximately 10 cm parallel to Polaris' x-axis (up–down)

Test B: translate approximately 10 cm parallel to Polaris' y-axis (left–right)

Test C: translate approximately 10 cm parallel to Polaris' z-axis (front–back)

Test	TLE RMS	TRE RMS Isotropic	TRE RMS Anisotropic	TRE RMS % Diff.	TRE RMS % Diff. Corr.
A	0.08 mm	0.22 mm	0.18 mm	-19.3%	-22.6%
B	0.07 mm	0.33 mm	0.19 mm	-42.4%	-45.2%
C	0.08 mm	0.32 mm	0.18 mm	-42.8%	-46.8%
D	0.06 mm	0.34 mm	0.24 mm	-28.7%	-29.9%
E	0.08 mm	0.60 mm	0.46 mm	-23.2%	-23.7%
F	0.06 mm	0.34 mm	0.22 mm	-34.5%	-36.0%

Table 1. Results of tests A-F. All statistics are computed over the usable 800 frames. The estimate of the TLE RMS is provided using Eq. (4). The TRE RMS statistics are computed for all the distance errors in a given test. The percent difference between the two methods is shown using the isotropic RMS as the reference. Finally, noting that the TLE contributes to the measured TRE RMS statistic, see Eq. (5), we correct the TRE RMS statistics and recompute the TRE RMS percent difference.

Test D: translate in all directions

Test E: rotate in all directions

Test F: random path including translations and rotations

A total of 1000 frames of data was collected for each test. Since a sliding window of $M = 200$ frames was used to estimate the FLE, only the last 800 frames are used for statistical analysis because it takes 200 frames for the FLE estimate to stabilize.

Three key sets of results are presented in Table 1. First, an estimate of the TLE RMS is given using Eq. (4) with estimates of the FLE covariance matrices found from the same FLE estimation algorithm used to determine the weightings for the anisotropic registration algorithm. FLE was estimated for Table 1 using the over-determined system of Eq. (1).

Next, the TRE RMS computed for both the isotropic and anisotropic registrations are provided. We notice that the anisotropic TRE RMS is lower for each test and the

percent difference between the two statistics is provided[8]. Finally, we note that the TRE

RMS for each registration has a contribution from the TLE where the measured TRE

RMS can be related to the actual TRE RMS by

$$\text{RMS}^2_{\text{tre,mean}} = RMS^2_{\text{tre,actual}} + RMS^2_{\text{tle}}, \tag{5}$$

Taking into consideration this relationship, we correct the percent differences

between isotropic and anisotropic TRE RMS statistics by using the $\text{RMS}_{\text{tre,actual}}$ and

show this new percent difference in the last column of Table 1. A small increase in the

accuracy is observed with this correction.

Moving beyond the observational statistics, we provide details of the data

measured during Test F. In Fig. 2, the FLE RMS estimates are provided for (i) solving

each of the individual sets of six equations (marker 1 to 4) and (ii) solving the over-

determined system of Eq. (1). It can be seen that the over-determined system give a

solution that falls within the variation of the plots from the individual sets of equations,

demonstrating the improved numerical stability that approach provides.

---

[8] The percent difference is computed using the isotropic registration as the reference value so that $\%\text{Diff} = 100 \times \left(\text{RMS}_{\text{tre},aniso} - \text{RMS}_{\text{tre},iso}\right)\big/\text{RMS}_{\text{tre},iso}$.
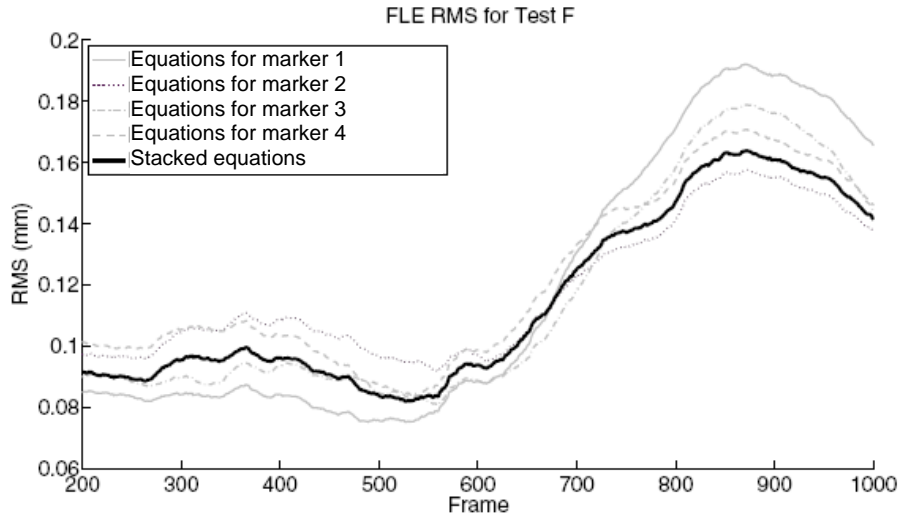
Figure 2. Sample of the FLE RMS for Test F. The RMS is computed by (i) solving the 6 equations for each marker and (ii) using the stacked, over-determined system of equations described in Eq. (1).
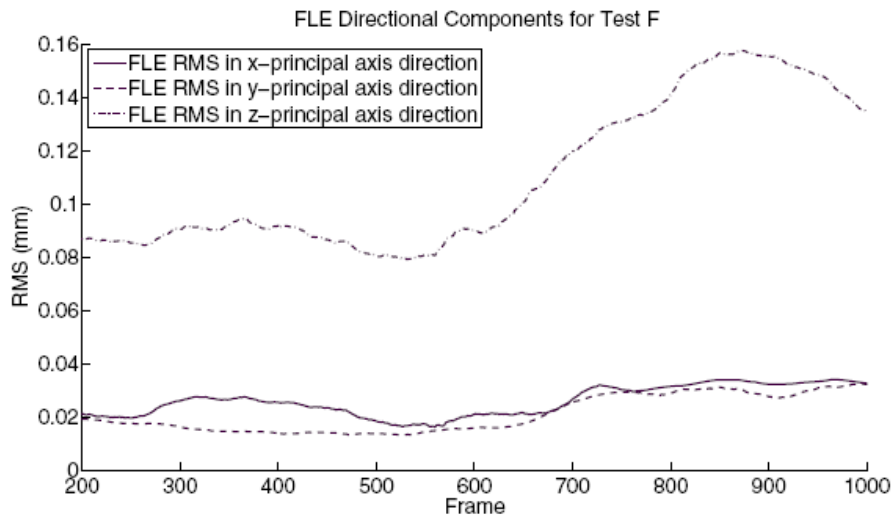

Figure 3. Sample of the FLE directional components for Test F. The directional components are singular values of the covariance matrix where the principal axes of the covariance matrix are approximately aligned with the axes of the optical tracking system.
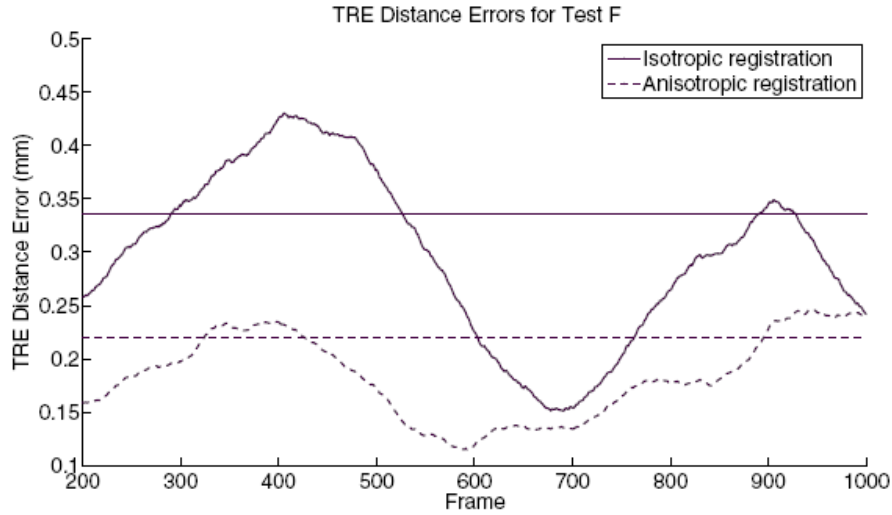
Figure 4. TRE distance errors for Test F. The distance errors are smoothed using a moving average of 200 frames to better view the trend. Horizontal lines represent the RMS for each test.

In Fig. 3, the magnitudes in the three principal directions of the average FLE covariance matrix are given. Here we observe the common behavior of optical tracking systems where one of the components is much larger than the other two directions. The direction of the higher magnitude error is along the viewing direction of the optical tracking system (z–axis of the Polaris Spectra).

The last plot in Fig. 4 shows the TRE distance errors at each frame in Test F. We clearly demonstrate here that the anisotropic registration provides a better estimate of the target location. The RMS TRE computed over the entire data collection is shown with the horizontal lines. We also performed a paired t-test for this case and found $p \ll 0.05$, suggesting that the difference in TRE between isotropic and anisotropic registration is statistically significant.

The results of the experiments comparing the modified version of the existing FLE estimation algorithm [3] to the alternative FLE estimation scheme, which employs Eq. (33) from Chapter II, showed no difference in accuracy when using the alternative

scheme. Therefore, we do not present detailed accuracy numbers for it. Not surprisingly, the time of execution of this more general scheme was higher.

4. Discussion

We have presented a modification of a recently developed method for point-based tracking of rigid bodies and have provide the first test of the method on real data. The improvement is accomplished by replacing the isotropic point-based registration method typically employed with one that takes into consideration the anisotropy of tracking system's FLE (see also Chapter III). Our method combines a modification of a recently published algorithm [3] for estimating the covariance matrix of anisotropic FLE with a novel rigid registration algorithm [4] that accommodates anisotropic weighting. The modification involves augmenting a set of linear algebraic equations to produce an over-determined set. The equations are employed to estimate an FLE covariance matrix from measured FRE components. We present evidence that this change improves the numerical stability of that estimation. Incorporating automatically generated weighting matrices optimized from the estimated FLE covariance matrices into the anisotropic registration algorithm exhibits a surprising level of accuracy, surpassing the current state of the art. Further improvements might be made by controlling for outliers or applying additional weightings during the solution of the equations, such as maximum-likelihood methods, M-estimators, regularization, etc. [7], [8].

The method was tested on data obtained with NDI Polaris Spectra® with various motions of the tracked tool in the region approximately 160 cm from the tracking system.

For a variety of motions, including pure translation, pure rotation, and combinations of both, we measured an increase in tracking accuracy in the range of $20 - 45\%$, and in every case accuracy was improved. A statistical analysis confirmed that the tracking of optical rigid bodies using our new approach is more accurate than tracking using the standard method $p \ll 0.05$.

The input to each FLE calculation is a set of previous FRE vectors collected from the most recent M frames, using the isotropic registration algorithm. We found that the size of the window, M, used for estimating the FLE statistics is an important factor in the tracking accuracy for our algorithm. The size of the window was investigated experimentally and we found that a contiguous collection of $M = 200$ frames of FRE measurements produced the best results. Larger windows gave only a very small improvement in accuracy but also increased the lag in obtaining a good estimate of the FLE statistics. With a window size of $M = 200$, one must allow the system to perform 200 data collections, before the improved tracking can commence. For the experiments described in this article, our method was implemented to run off-line in Matlab (Version 2009b, MathWorks, Inc., Natick, MA) on an Intel Core 2 Duo 2.2 GHz with 2GB of RAM while utilizing only one CPU. Each complete update, comprising FLE estimation and registration, required less than 6.4 milliseconds. For the NDI Polaris Spectra®, whose update rate is 60 Hz (period of 17 ms), our improved tracking algorithm can run in real-time. Furthermore, at 60 Hz with the NDI Polaris Spectra®, a delay of only four seconds is required to obtain the first set of 200 frames of FRE data to begin providing accurate estimates of the FLE statistics. In summary, the algorithm has promise to be

included in imaged-guided surgery applications that use point-based registration for highly accurate optical tracking.

5. References

[1] Peters, T.M. Image-guidance for surgical procedures, Physics in Medicine and Biology, 51(14), R505–R540, 2006.

[2] Danilchenko, A., Balachandran, R., Toennies, J.L., Baron, S., Munske, B., Wester III, R.J., and Labadie, R.F. Robotic mastoidectomy, Otology & Neurotology, 32(1), 11-16, 2011.

[3] Wiles, A.D. and Peters, T.M. Real-time estimation of FLE statistics for 3D tracking with point-based registration, IEEE Transactions on Medical Imaging, 28(9), 1384–1398, 2009.

[4] Balachandran, R. and Fitzpatrick, J.M. Iterative solution for rigid-body point-based registration with anisotropic weighting, Miga, M.I., Wong, K.H., eds., Proceedings of SPIE, Medical Imaging, Visualization, Image-Guided Procedures, and Modeling, Orlando, FL, USA, vol. 7261, p. 72613D. SPIE, San Jose, 2009.

[5] Schoenemann, P.H. A generalized solution of the orthogonal procrustes problem, Psychometrika, 31, 110, 1966.

[6] Horn, B.K.P. Closed-form solution of absolute orientation using unit quaternions, Journal of the Optical Society of America, A, 4(4), 629–642, 1987.

[7] Huber, P.J. Robust Statistical Procedures, SIAM, Philadelphia, 1996.

[8] Mair, E., Strobl, K.H., Suppa, M., and Burschka, D. Efficient camera-based pose estimation for real-time applications, The 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2696–3703, 2009.

CHAPTER V

ROBOTIC MASTOIDECTOMY

This chapter is adapter from the following paper:

- Danilchenko, A., Balachandran, R., Toennies, J.L., Baron, S., Munske, B., Fitzpatrick, J.M., Withrow, T.J., Webster, R.J., 3rd, and Labadie, R.F. Robotic Mastoidectomy, Otology & Neurotology, 32(1), 11-16, 2011.

1. Introduction

Taylor [1] has proposed that surgical robots be classified as either (a) surgical-assist devices, which modulate a surgeon's motions, or (b) autonomous robots (which he calls "Surgical CAD/CAM" systems to illustrate the analogy to industrial computer-aided design and manufacturing), which are programmed to replace a portion of the surgical task. Perhaps the most widely known example of a surgical-assist system is the da Vinci® Surgical System (Intuitive Surgical Inc., Sunnyvale, CA, USA), which mimics and modifies the surgeon's motions. These modifications can include elimination of tremor and scaling of motions such that large motions by the surgeon can be duplicated in a much smaller surgical field. This system has been most heavily used in urologic surgery [2]. Within the field of otolaryngology, it has been implemented for tongue-based resections to avoid splitting the mandible for access [3].

Of autonomous robots, the ROBODOC® System (Integrated Surgical Systems, Davis, CA, USA) is the most widely used and referenced. This system was designed to

bore a receiving lumen for orthopedic reconstructive joint surgery and has been used in more than 20,000 cases. In practice, the device is rigidly affixed to the long bone for the procedure. While it was shown that this system can accomplish the task better than a human operator, a class-action law suit was brought in Europe claiming that the wider surgical exposure necessary for the rigid linkage of the robot to the patient led to various complications including muscle and nerve damage, bone dislocation, and chronic pain. Because of these legal troubles, the company ceased operations in 2005 and was acquired by Curexo Medical Technologies (Freemont, CA), who received FDA approval for the ROBODOC® in 2008. The clinical status of the ROBODOC® is in question at the time of this writing.

Because surgical-assist devices are controlled directly by human manipulation, regulatory approval is more straightforward, and they have been first to achieve widespread clinical adoption. Since autonomous robots are not widely commercially available for surgical applications, few surgeons have any experience with them whatsoever. The surgical-assist device has the advantage of continuous human control, but the disadvantage that it cannot function appropriately unless the controlling human maintains continuous visual access both to the anatomy being resected and to those critical structures in the vicinity of the resection that are to be spared. The autonomous robot by contrast (which is also usually subject to continuous monitoring by the surgeon), relies on the human operator only as a backup while it follows the path prescribed by pre-operative planning. That planning, which can be performed in the surgeon's office before the patient is brought to the operating room, is based on three-dimensional tomographic images, and as a result, the autonomous robot can resect tissue that remains invisible to

the human controller, whose role is primarily to stop the process if a technical problem develops.

This scenario, in which preplanning determines the path followed by an autonomous robot, is well-suited for procedures involving rigid tissues such as bone. It is much more challenging to implement any degree of autonomy in robotic procedures involving soft tissues because of their deformation between imaging and intervention, but strategies for doing so are the topic of active research in the engineering community. As these results mature and are brought to clinical use, it is likely that we will see a blurring of the line between surgical-assist devices and autonomous robots, as limited autonomy is introduced to the former, while increasing levels of human control are incorporated into the latter. Regardless of the type or classification of robot used, the intent of incorporating robotic technology into a surgical procedure is to produce a better outcome, which can take the form of many metrics including decreased operative time, improved accuracy, smaller incision, decreased recovery time, or overall decreased cost.

In the field of otology, a core component of otological cases is the mastoidectomy, in which bone is milled away exposing but not damaging vital anatomy. Mastoidectomy lends itself to an autonomous robotic approach for two reasons: (a) the tissue to be resected is encased in rigid bone, and (b) critical anatomical features remain hidden until they are revealed by ablation. The first of these two reasons makes surgery with the autonomous robot feasible; the second makes it useful. The rigidity of bone is essential because it ensures that the three-dimensional structure of the target anatomy is the same during pre-operative imaging and planning as it is during subsequent intervention. The presence of critical hidden anatomy (nerves and other structures that must be spared but

are embedded in the bone), exploits the utility of the autonomous robot, because three-dimensional imaging pinpoints the positions of subsurface structures whose general locations can only be estimated during incremental, manual human intervention. As a result, the robot, guided by images that see beneath the surface, can safely ablate bone to which the human operator would be blind.

Once we have chosen a surgical situation, like the mastoidectomy, in which rigid bone comprises the target and critical anatomy is hidden from view, the central issue in autonomous robotics becomes the accurate determination of the single point in the three-dimensional pre-operative image that corresponds to a point in the surgical field, and vice versa. This determination of corresponding points in two spaces is termed "registration", and the success of the robotic approach hinges on the accuracy of registration from the space of the pre-operative tomographic image to the space of the intra-operative patient. This problem is not new. It has, for example, been faced for decades by stereotactic surgeons and others who have provided geometrical guidance intra-operatively based on pre-operative images [4]. To achieve it, an autonomous robot must register these two spaces either through rigid fixation, as is done by the stereotactic frame and by ROBODOC®, or via optical or magnetic tracking, such as that achieved with image-guided surgical (IGS) systems via the alignment of skin surfaces or fiducial markers.

In this chapter, we describe the use of an autonomous robot to perform a mastoidectomy using infrared tracking to monitor the motion of both the specimen and the robot. The method is applied to cadaveric temporal bone specimens. To accomplish this task, fiducial markers were implanted in the specimens, which were subsequently CT scanned. A surgeon identified the boundaries of the mastoid on the CT scan, the

trajectory was built for the robotic movement, and the robot was programmed to mill out the mastoid according to the identified boundaries. With both the robot and the specimen being tracked, any movement of the specimen during ablation was continuously compensated for by corrective motions of the robot. While the robot is subject to human control via a manual emergency stop button, it otherwise operates completely autonomously.

2. Materials and methods

Our goal is to successfully develop an autonomous robot system that can mill the regions identified by the surgeon in a CT scan. To achieve this goal, we developed the OTOBOT™ system (Figure 1), which incorporates an industrial robot, the Mitsubishi RV-3S (Mitsubishi Electric & Electronics USA, Inc., Cyprus, CA), controlled by custom software written in Matlab® and Simulink (The MathWorks Inc., Natick, MA) with real-time feedback of the robot's movement provided via a Polaris Spectra® optical tracking system (NDI, Waterloo, ON, Canada) [5]. Figure 1 shows the experimental setup of the system. We custom-built the end effector of the robot to hold a surgical drill. To enable continuous, accurate tracking of both the patient (a temporal bone specimen in the current study) and the robot, three separate coordinate reference frames each of which comprises four reflective spherical markers are attached rigidly to the robot's end effector, to the robot base, and to the patient. This arrangement enables real-time tracking of the robot movement, and hence the drill tip location, relative to the patient, whether or not the patient moves. That tracking is performed by the NDI Polaris, which is not shown but

Figure 1. OTOBOT™ robotic system set up to perform mastoidectomy on patient (temporal bone specimen in this current study). The robotic system consists of a Mitsubishi RV-3S industrial robot controlled by custom software. Coordinate reference frames with markers are attached rigidly to the robot end effector, robot base, and patient to allow tracking of the movements of the robot, drill, and patient during milling. An NDI Polaris optical tracking system (not shown here) is used to track the markers.

which faces the robot from approximately the position of the camera that acquired the picture shown in Figure 1—about 1-1.2 meters away.

While the three coordinate reference frames provide the means to determine the instantaneous positions and poses of the robot and patient relative to physical space in which the actual milling will take place, they provide no means to register that physical space to CT image space, in which the boundaries of the regions to be milled are specified. For that purpose, we employ bone-implanted fiducial markers. This choice was
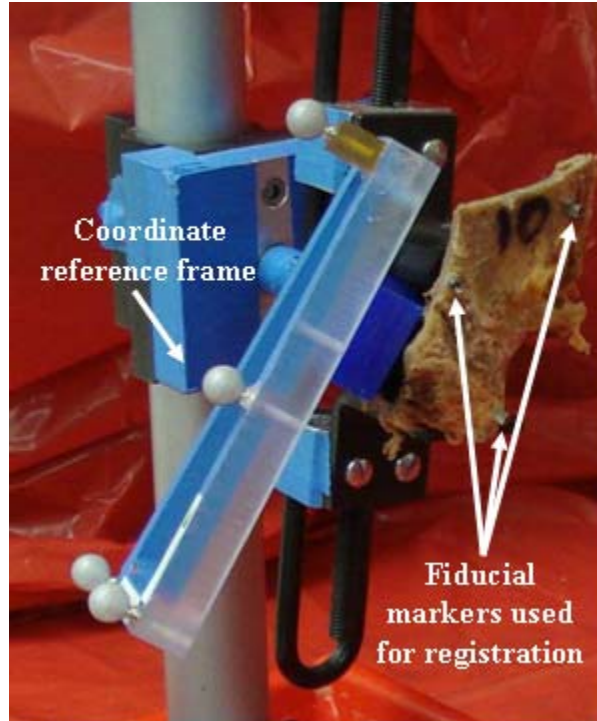
Figure 2. Temporal bone specimen with coordinate reference frame attached. Bone-implanted fiducial markers are used for registration of image and physical space.

made because bone-implanted markers are the most accurate available option, consistently enabling accuracies of 1.5 millimeters or better [6]. Three markers are screwed into bone in an arrangement surrounding the mastoid region, as shown in Figure 2. These markers, which are made of titanium, show up clearly in a CT scan and can be localized using image processing techniques. Their locations can also be acquired in physical space via a calibrated probe that is tracked by the NDI Polaris system.

Prior to the start of the milling procedure, three fiducial markers are bone-implanted into the temporal bone specimen, and a clinically-applicable temporal bone CT scan is subsequently acquired. The boundaries of the desired region to be milled are
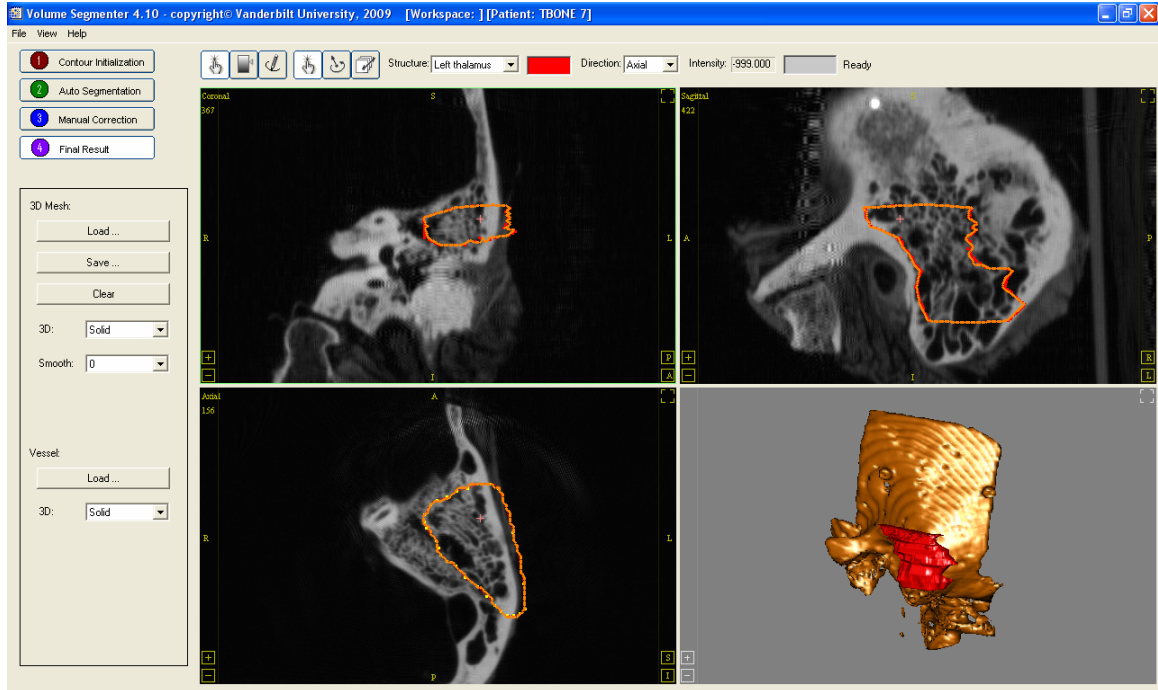
Figure 3. Screenshot of the software used for the segmentation. The surgeon contours the region to drill in the axial view of the image (bottom left). A three-dimensional shape of the region chosen for milling is displayed to the surgeon at the end of the segmentation.

contoured on this same CT scan by the surgeon using custom segmentation software (Figure 3). The surgeon outlines the boundaries on axial slices of the image while simultaneously viewing intersecting coronal and sagittal slices. The boundaries are chosen to encompass the region to be ablated while maintaining a safe distance from the critical structures to be spared, including the facial nerve, the tegmen, the sigmoid sinus, the external auditory canal, and the labyrinth. At the conclusion of the contouring phase, a three-dimensional rendering of the region that is to be milled is displayed. At this point the surgeon can make modifications to the contours if required. When the surgeon is satisfied with the segmentation, the region is transformed into a set of points that

determine the trajectory of the drill tip to be executed by the robot. For building the robot trajectory we used the algorithm developed specifically for this type of application. The description of the algorithm is presented in Section 3 of this Chapter.

The fiducial markers are localized in the CT image (may be done before or after segmentation is done). A coordinate reference frame is then rigidly attached to the specimen, and the locations of the fiducial markers are acquired in physical space using the calibrated probe. The CT image and physical space are registered by applying point-based rigid-body registration [7] to the two sets of fiducial points—one in each space. Using the registration thus found, the trajectory points for the robot are transformed from CT space to physical space, and the transformed points are given as input to the software that controls robot movement.

The specimen with the coordinate reference frame is then placed within the workspace of the robot. Care is taken to make sure that the markers in all three coordinate reference frames are within the field of view of the optical tracking system. The milling is then initiated by starting the robot-control software application. The movements of the robot and patient are continuously tracked, and the control software compensates for any movement of the patient during milling. A manual emergency stop button is available for the robot if the robot moves in an undesired fashion. A continuous computer screen update is provided showing the progress of the milling as percentage of the total points covered.

Robotic mastoidectomy was performed on three temporal bone specimens using the procedure described above. A 6 mm diameter surgical drill tip was used for the milling, and a pneumatic drill was run at pressure level equal to 100 psi throughout the
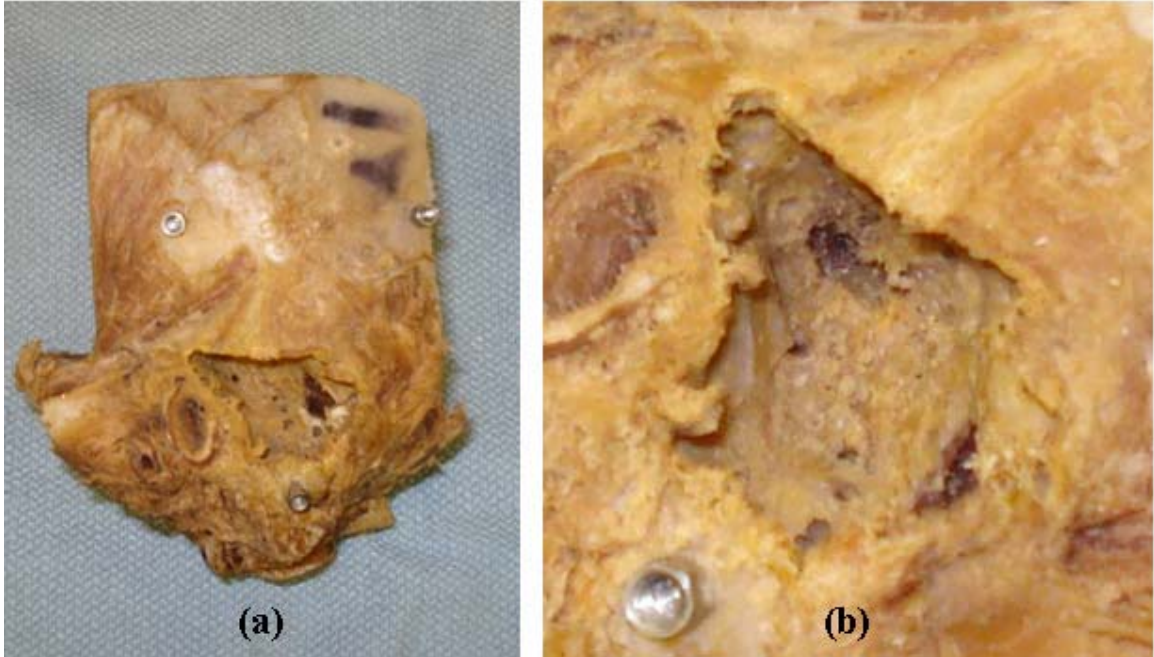
Figure 4. Temporal bone after completion of robot milling. (a) Temporal bone overview. (b) Close-up of the milled region. All three fiducials are visible in (a); one is visible in (b).

milling. We included a 1 mm safety region during milling to compensate for possible small penetrations of the drill tip beyond the region of interest (arising from small inaccuracies in registration, real-time tracking, or robot control). For this initial study, the robot was set to move at the constant speed of 2.5 millimeter per second (mm/s) during milling. We observed no undesired movement of the robot during milling, and hence we did not use the emergency stop button at any time.

3. Trajectory building algorithm

Though many algorithms exist for building trajectories for robots, we did not find any algorithms which would exactly suit our application of robotic mastoidectomy. Thus,

we built our own algorithm. There were three factors that we needed to take care of for mastoidectomy application. First, the algorithm has to perform in the shortest time possible. Second, the trajectory should cover all the points of the selected region. Third, we must approach the points on the surface of the patient's skull before we approach points below the surface points.

In our algorithm we combine the following two ideas: (1) neighborhood "marks", which are values set for each point in the region and show the neighborhood relationships of all the points in the region, and (2) the shortest distance to the outside of the skull. Here are the steps of the algorithm:

1) Select a point outside of the patient's skull in front of the region to be drilled, or in other words, select a spot from which the drill, attached to the robot, will be approaching the region. We call this point the "initial point". It will be used to calculate the shortest distances from each point of the region to the outside of the region.

2) Fill in a "Neighborhood Table" which contains the neighborhood marks for all the points. For this, first set all neighborhood marks to an infinitely big number. Then for the point in the selected region which is the closest one to the initial point, set its neighborhood mark to 1. Add this point to the "To-Process" list. Recursively process the points from the To-Process list in the following manner: if the current point has the neighborhood mark k, for all the neighbors (or points which can be achieved by iterating one of the indexes by 1) inside of the region to be drilled, update their neighborhood marks to $k+1$ and add them to "to-process" list if their

neighborhood marks are bigger than *k+1*. Remove the current point from the To-Process list and go to the next point from the list. Continue the procedure until the list is empty and all the points are processed.

3) Calculate the distances from the initial point to each point of the region and save them in the Distance Table.

4) Build the trajectory following these directions: go through all the points in the Neighborhood Table. Go only to the points not visited before. First visit the points with smaller neighborhood marks. If there are several points available with neighborhood marks that have the same value, first go to the ones having smaller distances to the initial point in the Distance Table. If at some point of time there are no unvisited neighbors, but there are still points in the region that have not yet been visited, find an unvisited neighbor by backtracking through the previously visited list of points. If such a neighbor is found, go to that point. Figures 5, 6, and 7 show a two-dimensional example of a Neighborhood Table, a Distance Table, and a resulting trajectory. Backtracking is shown in pink.

Moving just to neighbor points makes the algorithm time efficient because no time is wasted for the robot to travel through already drilled region. During neighbor-to-neighbor movement, the drill always removes part of the bone. If all points in the region are connected, then the algorithm can be shown to include in the final trajectory all points in the region. Fortunately for our application – mastoidectomy – the surgeon always selects a region in which all points are connected. Going first to points with smaller

33	32	31	■	■	28	29	30	■	■
32	31	29	28	27	26	27	28	■	■
■	■	28	27	25	23	25	27	■	■
■	■	26	25	22	22	22	25	26	27
27	26	24	22	21	18	21	22	24	26
■	24	22	20	17	15	17	20	22	24
■	23	21	18	13	10	13	18	21	23
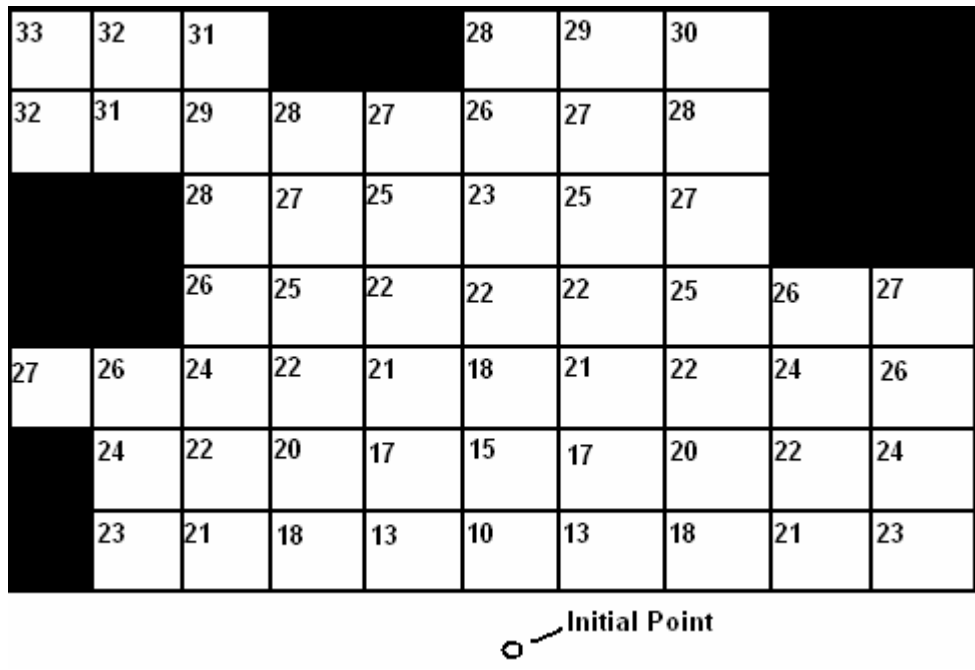
Initial Point

Figure 5. Example two-dimensional Distance Table

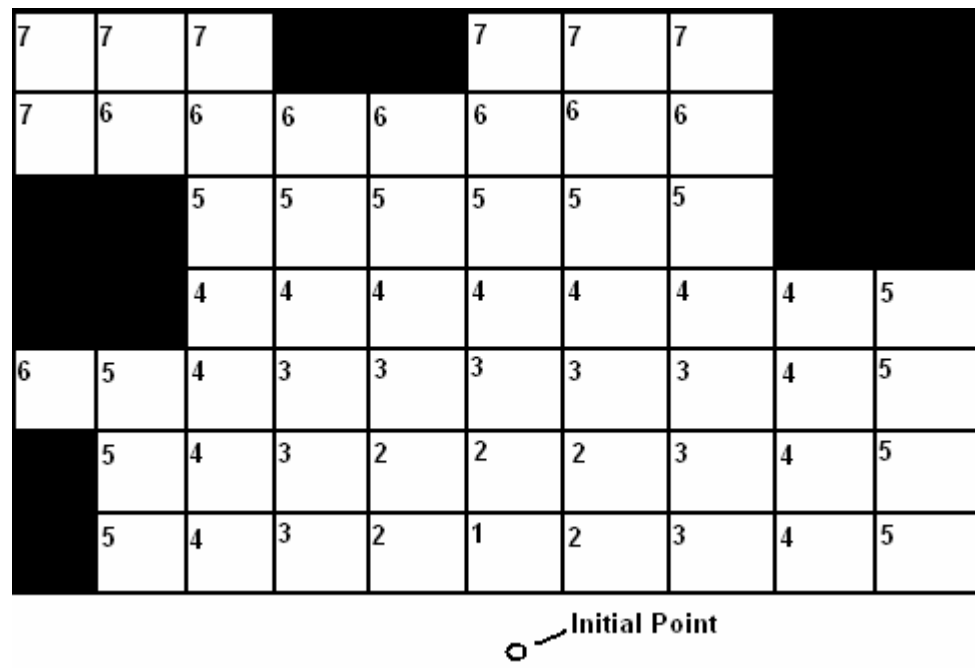| 7	7	7	■	■	7	7	7	■	■
7	6	6	6	6	6	6	6	■	■
■	■	5	5	5	5	5	5	■	■
■	■	4	4	4	4	4	4	4	5
6	5	4	3	3	3	3	3	4	5
■	5	4	3	2	2	2	3	4	5
■	5	4	3	2	1	2	3	4	5

Initial Point

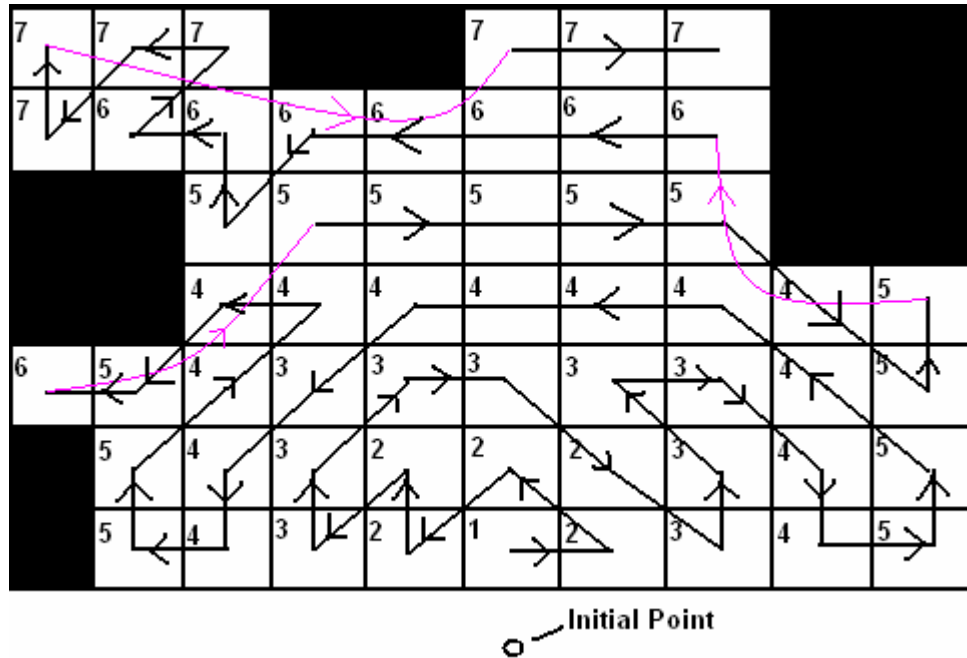Figure 6. Example of two-dimensional Neighborhood Table

157

Figure 7. Example of building the trajectory based on the Neighborhood and Distance Tables

distances from the initial point guarantees, that the points on the surface are visited before the points below the surface.

Because the edges of the region might have a complex shape, a security check is performed after building the trajectory to make sure that no "illegal" bone is drilled while going from one point of the trajectory to the other one. By "illegal" bone we mean a part of the bone that is not in the outlined region and thus should not be drilled.

4. Results

Figure 4 shows the results of the milling of a temporal bone specimen. All three milled temporal bone specimens were CT scanned after the procedure for the purpose of

analyzing the results of the milling. On analysis of the scan, no damage to any critical structure was identified. To determine whether the targeted bone was in fact removed, the pre-operative and post-operative CT scans of each bone were registered using the three fiducial markers, and the originally delineated region was superposed from the pre-operative CT scan onto the post-operative CT scan. For the three bones, we calculated the percentage of each targeted volume that was removed. For the three bones, these percentages were 97.70%, 99.99%, and 96.05%. Maximum error was identified as 0.6 mm. The time spent for completing the mastoidectomy was 14 minutes.

5. Discussion

Herein, we have described a first step along what will likely be a lengthy road towards clinical testing and implementation of mastoidectomy via an autonomous robot. This first step involved the modification of an industrial robot to perform complex milling on a cadaveric specimen under infrared tracking of both robot and specimen. In our testing on three specimens, we found that the surgeon's pre-operative plan was successfully executed by the robot with at least 96% of the targeted bone removed without damage to critical structures. Excitement at this success is tempered, however, by the realization that a great deal of work remains before this concept can be tested in the operating room. While the fundamental engineering concepts behind the robotic technique are well developed, less well studied is the translation of such concepts to clinical applications. Issues such as maintenance of sterility, logistics regarding transportation and set-up of the robot, and redundant safety constraints will need to be

incorporated, as have been done with both the da Vinci® and ROBODOC® systems, both of which took decades to go from bench top to clinical use.

We acknowledge limitations of the proposed system, most notably the lack of soft-tissue work, which comprises at least a substantial portion of any ear surgery. Working with existing technology, we designed our system to aid the surgeon by automating the most predictable component of the surgery, mastoid milling, based on CT scan of the specimen. The potential advantages of this approach include (a) reliability and positional accuracy of the robot, (b) "X-ray vision" afforded by the registration of the pre-operative CT to the intraoperative patient, and (c) possible economic benefits (e.g. reduced time of intervention, improved productivity, or other, as yet to be identified, metrics).

Regarding (a), robots are highly reliable and precise. The Mitsubishi RV-3S robot used in this study has a repeatability of 0.02 mm. In addition to high precision, robots are highly reliable in repetitive tasks, with error rates lower than humans and no risk of performance degradation due to fatigue [8]. In short, properly calibrated robots can perform specified tasks with high accuracy and efficiency. Regarding (b), coupling IGS with an autonomous robot leads to "X-ray vision" which allows such systems to see subsurface features before they may be injured. In the bone component of mastoid surgery, this "X-ray vision" offers a potentially large benefit in that the robot could be used to perform a highly accurate three-dimensional mastoidectomy—leaving a 1-2 mm margin of safety over vital anatomy—freeing the human operator to perform the more high-level fine dissection where human judgment is paramount.

Despite benefits (a) and (b), we recognize that catastrophic failure is a possibility as, for example, if the registration between the CT and the target tissue was performed incorrectly, thereby placing the robot at the wrong starting point. Given the consequences of such damage, we believe that—particularly in initial cases—there is no substitute for continuous monitoring by a trained surgeon. Put another way, the robot may be trusted to carry out the surgeon's plan only so long as a human is there to verify that it is doing what it was told. As such we envision, and are working on, human oversight systems such as having the surgeon depress the foot pedal to keep the robot moving and the drill spinning, a hand-held pause button, and the possibility of slow-motion control to reduce robot speed when in close proximity to vital anatomy.

Regarding (c), there are many possible ways in which a robotic mastoidectomy may be economically beneficial. This may include reduced time of intervention. For the experiment when we used our new trajectory building algorithm 14 minutes were required for completing the operation. The 2.5 mm/sec was chosen as an initial speed of the robot movement velocity based on our desire to carefully monitor the progress of the robot. We are confident that we can increase this speed to 5 mm/sec, as has been reported in the milling of cranial bone by others [9]. This increase in speed alone might reduce the 14 minute intervention to 10 minutes or even less.

Future economic assessments will be necessary to justify clinical use of such a system, as has been underway for the da Vinci® system, which retails for $1.4 million. Use of the da Vinci® robot for prostatectomy adds over $2,500 to the cost of each surgical intervention [10]. To date, the offsetting benefits include decreased recovery time and decreased complications [10]. Because our system is far from routine clinical

use, we can report, in terms of economics, only the cost of the system, which is approximately $40,000 (robot $19,900, infrared tracking system $15,000, and control computer $5,000). Estimating the cost of development and experimentation required to obtain FDA approval and commercialize the robot, we feel the machine may retail for $500,000. As wear on the robot, whose movements are far smaller and slower than those required for routine industrial applications, will allow it to be used on thousands of cases with minimal maintenance, we predict its cost per procedure will be below $100. Other costs accruing from this technology will be associated with markers and drapes, which per case should be below $100.

Noting that the typical outcomes associated with mastoid surgery (e.g. post operative hearing, success of tympanic membrane grafting, recurrence of disease, acquisition of speech after cochlear implant) are independent of whether the drill is held by a human or a robot, the major potential benefit offered by this technology may be reduced operative time. However, until the robot can completely replace the surgeon, its benefit will be bound by the need for the surgeon or an assistant to monitor the procedure thus tying up operating room time. At this point, we predict that the savings from reduced operating room time can be expected to more than offset the additional costs (predicted above to be $200 or less) associated with this technology. There may be other benefits that we are as yet unaware of (e.g. fewer drill bits per case, fewer complications) that may make the economic argument for this technology more compelling.

Before such a clinical system is deployed, many intermediate steps need to be taken. Our near-term future work will be focused on improving the efficiency of the path plan so that the milling time can be reduced. Medium-term future work will include

redundant safety checks including monitoring force feedback at the point where the robot grips the drill, potentiometers on robot joints (in addition to the optical encoders used in industrial robots), improvements to the user interface for the surgeon, and the development of techniques for bagging and sterilization of robot components. Only after these steps have been accomplished will our system be ready for clinical testing.

6. Conclusion

The preliminary work we have presented here, given the relatively simple and controlled conditions of a cadaveric specimen in a laboratory, shows that an autonomous robotic system is capable both of determining a trajectory and of directing a drill along that same trajectory so as to perform a prescribed mastoidectomy with 96% removal of desired bone volume and with no damage to critical structures. There is much to do, however, before such a system can be considered for clinical use. The major issue is to establish a level of patient safety that is at least at the level of current clinical practice. This safety level will require that a surgeon remain in charge of the procedure including sitting at the surgical bed, monitoring the robot's progress, and stopping the system if a problem develops. This work is only a start, but it provides encouragement that, with a plan based on pre-operative tomographic images that give information about what is below the bone surface, an accurate registration to transfer that information to the patient in the operating room, and an obedient robot continuously monitored by the surgeon, robotic mastoidectomy is technologically feasible.

7. References

[1] Taylor, R. A perspective on medical robotics, Proc. IEEE, 94(9), 1652–1664, 2006.

[2] Guru, K.A., Hussain, A., Chandrasekhar, R., Piacente, P., Bienko, M., Glasgow, M., Underwood, W., Wilding, G., Mohler, J.L., Menon, M., and Peabody, J.O. Current status of robot-assisted surgery in urology: a multi-national survey of 297 urologic surgeons, Can. J. Urol., 16(4), 4736-4741, 2009.

[3] Hockstein, N.G., Nolan, J.P., O'Malley, B.W., Jr., and Woo, Y.J. Robot-assisted pharyngeal and laryngeal microsurgery: results of robotic cadaver dissections, Laryngoscope, 115(6), 1003-1008, 2005.

[4] Fitzpatrick, J.M. The role of registration in accurate surgical guidance, Journal of Engineering in Medicine, 224(5), 607-22, 2010.

[5] Baron, S., Eilers, H., Munske, B., Toennies, J.L., Balachandran, R., Labadie, R.F., Ortmaier, T., and Webster, R.J., III. Percutaneous inner-ear access via an image-guided industrial robot system, Journal of Engineering in Medicine, 224(5), 2010.

[6] Maurer, C.R., Jr., Fitzpatrick, J.M., Wang, M.Y., Galloway, R.L., Jr., Maciunas, R.J., and Allen, G.S. Registration of head volume images using implantable fiducial markers, IEEE Transactions on Medical Imaging, 16, 447-462, 1997.

[7] Fitzpatrick, J.M., Hill, D.L.G., and Maurer, C.R. Registration, Medical Image Processing, Volume II of the Handbook of Medical Imaging, M. Sonka and J.M. Fitzpatrick, eds., SPIE Press, 447-513, 2000.

[8] Paul, R.P. and Nof, S.Y. Work methods measurement--a comparison between robot and human task performance. International Journal of Production Research, 17(3), 1979.

[9] Engelhardt, M., Bast, P., Lauer, W., Rohde, V., Schmieder, K., and Radermacher, K. Manual vs. robotic milling parameters for development of a new robotic system in cranial surgery, International Congress Series, 1268, 533–538, 2004.

[10] Lotan, Y., Cadeddu, J.A., and Gettman, M.T. The new economics of radical prostatectomy: cost comparison of open, laparoscopic, and robot assisted techniques, J. Urology, 172(4),1431-1435, 2004.

[11] Menon, M. Robotic radical retropubic prostatectomy, BJU Int., 91(3), 175-180, 2003.

CHAPTER VI

CONCLUSION AND FUTURE WORK

1. Conclusion

In this work we have developed methods for working with image registration when fiducial localization error is arbitrary. Our methods work when FLE is homogeneous or inhomogeneous – same or different for each marker, and isotropic or anisotropic – same or different in all directions for each marker. First, we developed methods for estimating TRE and FRE. The uniqueness of our formulas is that they work for any fiducial configuration and any FLE configuration. We also looked at some special cases that have been investigated by others and simplified our formulas for those cases. In addition, we researched the question of FRE and TRE independence. We showed that contrary to a wide-spread belief, a small FRE in a given case does not imply a small TRE. Thus, using FRE as a measure of registration accuracy should be done with a caution. Second, we developed a new method of registration when FLE is anisotropic or isotropic. This new method is more intuitive than the existing ones and performs with more accuracy in less time. Third, we developed a new tracking method. This method uses our new registration algorithm combined with FLE estimation algorithms. As FLE estimation algorithms, we first used an existing method, and then we used the expression that we derived for estimating FRE from FLE, but we used it "in reverse" instead to estimate FLE from FRE. Finally, we developed a robotic system for performing a medical procedure—

the mastoidectomy. This system allows for patient movement via tracking using a standard tracking algorithm.

In Chapter II we derived formulas for estimating TRE and FRE. The input parameters for the estimation equation are the fiducial configuration and FLE configuration. Our formulas work for any arbitrary fiducial configuration, for any FLE configuration – FLE can be homogeneous or inhomogeneous and isotropic and anisotropic – and for any set of weightings. Previously several methods existed for estimating TRE and FRE. However, all of them dealt with some particular FLE and/or some particular set of weightings. The uniqueness of our formulas is that they have the same form for all FLE configurations and all weightings. We compared our formulas to the previously published expressions. Our formulas produced the same results that the previously published expressions did. We also compared the results of our formulas with Matlab simulations. The simulations showed the same values, which supports the correctness of our formulas. Finally, by means of simulations we showed that FRE and TRE are independent. This result is important because of a wide-spread erroneous belief that FRE can be a good measure of accuracy of registration. In fact, using FRE as a measure of accuracy must be done with extreme caution.

In Chapter III we developed a new iterative algorithm for performing point-based registration. This algorithm allows an arbitrary FLE configuration – FLE can be homogeneous or inhomogeneous and isotropic and anisotropic. Our new algorithm is more intuitive and, moreover, more accurate and faster than other existing algorithms. Our method has just one adjustable parameter – the stopping criteria. By means of Matlab simulations and using real data acquired with a tracking system we showed that our new

method works better than four other existing methods. It produces a smaller error and converges more quickly compared to other iterative methods (no closed-form methods handle arbitrary FLE). Our simulations were performed for the cases most likely to happen in surgical applications. A decreased TRE was observed in all those cases. We also performed simulations to research the reliability of our new algorithm with regards to the collinearity of the system. For all fiducial configurations usable in real-life applications, our new method converges better and produces smaller TRE compared to the four other methods we studied.

In Chapter IV we developed a new tracking method. We combined an FLE estimation algorithm with a new registration method presented in Chapter III. For estimating FLE we used first an existing FLE estimation algorithm and then an expression for calculating FRE derived in Chapter II. The advantage of our new tracking method over the standard tracking method used in commercial systems is that our algorithm takes into consideration FLE anisotropy. Our new tracking algorithm was tested on the real data with a considerable improvement in TRE observed during the experiments.

In Chapter V we presented our newly developed robotic system for performing mastoidectomy. Mastoidectomy is a medical operation of removing part of the human temporal bone called mastoid. Nowadays this operation is performed by humans only. The difficulty of this operation is the presence of several critical structures in the area close to a temporal bone. Thus while the operation is performed by a human surgeon there will always be a probability of an error and damaging one of the critical structures. To eliminate that probability of an error, we built a robotic system for performing this

operation. The robotic system consists of an industrial robot, an optical tracking system, a real-time module, and a host machine. Before the operation, CT scans of the patient are acquired. Then a surgeon outlines the region to be drilled using special software. Afterwards based on the outline a trajectory is built using the new trajectory building algorithm. We built our new trajectory building algorithm to deal with the specific demands of the mastoidectomy procedure – all points of the outlined region should be visited, points on the surface should be visited before points below them, and time for visiting all points should be as small as possible. We combined two ideas – neighborhood marks and distance marks in our new algorithm. We performed several experiments drilling temporal bones. The results of the experiment show that on the average 97% of the outlined region was drilled and no critical structures were damaged during robotic drilling. Using our new trajectory building algorithm it took only 14 minutes to remove a mastoid which beats the average time of a human surgeon (~ 20 minutes). We believe this time can be reduced even more.

2. Future work

In Chapter III our new algorithm was compared to two existing iterative methods. Our criterion for the choice of methods for comparison was the number of adjustable parameters required for each algorithm. In both algorithms we selected and our new algorithm there is only one parameter to adjust – a stopping criterion. However, several more methods exist to perform registration in presence of anisotropic and inhomogeneous

FLE which require adjusting more than one parameter. As future work our new method can be compared to those methods using simulations or real data.

The registration algorithm presented in Chapter III was validated using Matlab simulations and one experiment with real data. However, more real data based validation is needed to be performed for this algorithm. One of the ways to perform such a validation would be using some very accurate measuring devices, such as FARO Arm, and less accurate devices.

As it was mentioned in Chapter IV, all current FLE estimation algorithms assume that FLE is homogeneous. Though it is almost true for many tracking systems, having an inhomogeneous FLE estimation algorithm would be of value. One possible solution can be to use Eq. (33) from Chapter II. Without any additional information it can be used for estimating homogeneous FLE. For estimating inhomogeneous FLE, either a larger number (> 4) of fiducials is required or some additional information is required. Thus, experiments with a larger number of fiducials or with the determination of restrictions that would allow the use of Eq. (33) from Chapter II with fewer than five fiducials can be done in the future.

To compare our new tracking method presented in Chapter IV with the traditional tracking method used in commercial tracking systems on a real-life application, robotic mastoidectomy procedure can be performed using both methods. For this matter two identical objects should be created. Both of them should be drilled with the same trajectory. Different tracking methods should be used during the objects tracking. After drilling, the drilled regions of both objects should be compared to the outlined region. Based on those result a conclusion can be made about the quality of the new tracking

algorithm in real-life applications. Finally, methods to allow effective monitoring of the

robot's progress through its trajectory should be developed.