[[1]](#footnote-1)

Calibration and characterization of electromagnetic position and orientation trackers

First A. Author, *Fellow, IEEE*, Second B. Author, and Third C. Author, Jr., Member, IEEE

*Abstract*—foo

*Index Terms*—Enter key words or phrases in alphabetical order, separated by commas. For a list of suggested keywords, send a blank e-mail to [keywords@ieee.org](mailto:keywords@ieee.org) or visit <http://www.ieee.org/organizations/pubs/ani_prod/keywrd98.txt>

# INTRODUCTION

T

HIS document is a document

Project summary:

The ILEMT (In-Loop ElectroMagnetic Tracker) has the specific aim of developing a tracker with adequate speed, resolution and accuracy to be used within a feedback system that stabilizes handheld surgical instruments. For this particular use, the primary focus is on measurement speed and low noise which is primarily due the (open) hardware design.

A magnetic position tracker measures the *pose* (position and orientation) of a *sensor* with respect to a magnetic *source*. [2D single axis drawing]

A single field measurement is not sufficient to determine the pose, so the source incorporates distinctly modulated electromagnet source coils, and the sensor often measures the local magnetic field along independent axes. Frequently , but other combinations are possible. For each pose, each source field experiences a characteristic magnetic coupling into each sensor coil, giving distinct couplings, which can be represented as the coupling matrix , where is the signed coupling from source coil into sensor coil . The sign represents direction of the magnetic field vector w.r.t. the sensor coil.

We notate a pose as the linear transform which maps the source coordinate frame onto the sensor coordinates. [picture of coordinates] This can be represented as a 4x4 linear homogenous transform matrix. Let be the *measurement function*: the actual coupling measured in pose . In practice, depends on magnetic interactions that are intractable to model. Instead, we resort to a *measurement model* This is a physical model (such as the ideal dipole) which is parameterized by . Calibration is the process of determining

During calibration we measure for many different known poses spread across the position and orientation workspace. Then a general nonlinear optimizer is used to determine the that minimizes:

Where the matrix 2-norm is the sum of element-wise squares:

This least-squares formulation permits us to use a least squares solver (here Matlab lsqnonlin). These solvers give a much more robust convergence than general function minimization, especially when the derivatives of are unknown and the initial value for is inaccurate.

This is an excursion into solution methods, which we need at least a bit on later. But this does work here. Magnetic tracking is an *inverse problem*. Once we have chosen a measurement model that we expect to approximate the true source/sensor coupling, it is straightforward to apply the physics to predict what coupling we should expect to see in a given sensor pose. But the physics does not tell us how to solve the inverse problem: given a measurement , what is ? Some measurement models permit a closed-form pose solution using linear algebra [cite], but in general it is again necessary to use nonlinear optimization, this time finding which minimizes .

One way to appreciate the challenge of measurement modelling is that we expect to achieve an accuracy which is much smaller than the size of the source and sensor. In the tested ILEMT configuration, the source is ??? mm, and the sensor ??? mm, while the specified accuracy is 200 µm. Due to the physics of magnetism, this is not as problematic is it seems. Getting ahead of ourselves with the dipole model? The magnetic field drops off rapidly, , and with increasing the field also rapidly converges toward the dipole field of an ideal magnetic point source. At source to sensor distances greater than 10x the source size the dipole deviation is ???. In ILEMT, to minimize noise, the sensor distance is around 5x the source size, which makes higher order models more useful. r/d, dipole discussion elsewhere

Sometimes the source and sensor are identical, but the common practice is to make the source much larger, with a correspondingly increased drive current. Then the sensor can be made much smaller while still maintaining the same signal to noise ratio. Since the source is much larger than the sensor, the dipole deviation of the source tends to dominate the error of the measurement model.

# Related work

## Calibration

In work that calibrates an EMT using a magnetic model, nonlinear optimization, such as the approach we describe, is usually used, but specifics on the model parameters, and especially the optimization strategy, are lacking.[cite]

Some EMTs use other sensing principles and source modulations [cite Ascension, rotating field thing, etc], but we are concerned with the most common approach where the source coil has a sinusoidal drive voltage, and the sensor uses inductive pickup coils.

## Evaluation

Methods for EMT testing are necessarily discussed in work that characterizes the accuracy of a new EMT. Ground truth accuracy? Test patterns?

In the medical literature it is common to characterize tracker accuracy, both to compare performance of different products, and also to test the suitability of a particular EMT for a specific type of medical procedure. The tracker calibration is a given, provided by the manufacturer, and embedded into the product’s proprietary firmware and calibration data. Usual practice is to use a manual fixture (or *phantom*) for sensor positioning. This limits the number of poses that can be tested; frequently only one sensor orientation is used, leaving us guessing about the position error caused by sensor rotation (rotation translation coupling). In medicine it is indeed a good thing to evaluate the accuracy of an available tracker in a particular context. While not comprehensive, these relatively simple procedures do reveal significant variation across EMTs and their use environments.

## Output correction

Once tracker output poses have been collected across a range of ground truth poses, it is also common to undertake to correct the tracker error. [cite] This differs from magnetic calibration in that it once again assumes a pose output, and fits a model to the pose error. This can have considerable benefit in the particular calibrated environment, since it compensates for interference from metal objects and electrical sources. [cite VR]

# Positioning for calibration:

Calibration and accuracy testing require placing the sensor in many known poses, in both translation and rotation. This is our ground truth, and must have significantly higher accuracy than the expected tracker performance. The ILEMT microsurgical application is particularly demanding. For ILEMT, the accuracy as a fraction of the source to sensor distance specification is similar to other EMTs, but the workspace is smaller, so the absolute accuracy required for calibration is also more demanding. To characterize an accuracy of 200 µm we require a ground truth position accuracy of 20—70 µm. Expressed as an angular error at 200 mm, this is 100—350 microradians, or 0.006—0.020 degrees. [this is sensitivity related, but sensitivity to ground truth error is just geometry. Sensitivity to calibration or model error is less obvious. Gather or rationalize sensitivity discussion]

Compounding the difficulty, the sensor positioner must also be non-metallic to avoid interfering with the measurement. More precisely, any magnetic or highly conductive material must be sufficiently far away from the workspace. Fortunately, this metallic interference drops as , so if the closest interferer is at 3.2x the source/sensor distance, then the interference will be reduced by 1000, which can usually be neglected.

One approach has been to use a nonmetallic manual positioning fixture [cite duplo, medical paper]. This procedure is tedious, which tends to limit the number of poses used. Rotation measurement has often been ignored, perhaps in part because of the combinatorial explosion created when rotation and translation are both varied.

For a credible performance evaluation, it is necessary to position the sensor in multiple rotations, as well as across the translation workspace. It can be possible to calibrate a low order measurement model using only translation data, but it then remains unclear what accuracy can be expected with the sensor is rotated. Unsurprisingly, we have found that calibrations done with translation only *do* perform worse when there is also angular motion.

For 3D translation and a single rotation axis, magnetic isolation can be achieved by placing the sensor on a nonmetal pole extending away from the large metal components in the positioning stage. [cite Polaris, Johnson & Johnson] The use of this pole does create stronger demands on the ideality of the stage translational motion, with minimal incidental rotation (stage runout).

We use this approach, and then achieve additional rotations by using manual fixtures to get multi-axis 90 degree rotations of the source and sensor. Given the primary concern for sensor position error (compared to rotation), the accuracy of the source rotation fixture is particularly important, since source rotation error is converted to sensor position error by the source to sensor radius. Broadly, rotation of the sensor helps to identify the sensor response, while rotations of the source identify the pattern of the source field. The source and sensor responses are independent, so it is not necessary to evaluate the full cross product of source and sensor rotations.

[What calibration data patterns do we use?]

[pictures, more details of stage, fixtures, discussion of using surface plates and machining measurement tools, move with backlash, stage performance test stats?]

# The measurement model

## The dipole approximation

Magnetic field

Fig. 1: Dipole model. A source coil is located at with moment . The sensor located at measures the source magnetic field vector .

*Source*

*Sensor*

The dipole is the simplest physical model for the field created by a magnetic point source. For distance from the source , and any actual source coil of size , as , the field converges to a dipole. The dipole approximation [blah blah cite, figure out] is good at .

We will only discuss dipole calibration. It is simple, and often works well in magnetic trackers. Also, for calibration the details of the model matter very little. Where the particular magnetic model does become important is in implementing the pose solution, which is outside our scope.

In 3D Euclidian space, a dipole possesses only two parameters: location and vector moment . [figure] If the sensor is located at , then the field is:

(so dipole)

where , , and is the permeability of the medium.

is a vector, which a single sensor coil would measure as a scalar voltage along some sensor moment . The analogous dipole form of the sensor response is:

(se dipole)

this response is simply the scalar product of the sensor moment with the field. [could spell out the equation a bit more, eg. substituting in sensor model.]

## Coordinate notation

These basic magnetic models are coordinate free, but implementation of measurement model across multiple coordinate systems and multiple source and sensor coils requires more notation. [use ~ rather than ^ for the model?]

In the tracker kinematics, there are two independent parts: the source and the sensor, abbreviated *so* and *se.* Each has its own coordinate system.If an arbitrary point is represented in the source coordinates, then it is , while in the sensor coordinates it is . Our general notation is:

where , and optionally designates a sub-part. For example, the tracker output is the sensor pose with respect to the source coordinates, which could be written as .

We will index the source and sensor coils by Here, is simply a name, not a variable. This is intuitive in the common case the source and sensor coils roughly coincide with the axes of the corresponding coordinate system. We represent the dipole parameters as two 3x3 matrices and , where the coil locations and moments are concatenated as column vectors:

The source parameters are and , and the sensor has and . Since these quantities move with the part, the calibration parameters are determined in the corresponding coordinate system, eg. , but when transformed to source coordinates the matrix becomes .

Fig. : Pose diagram for kinematics of the measurement model. The pose maps from source coordinates to sensor, but each coil also has an independent location and moment within its frame.

Source

Sensor

z

y

x

Coils

z

y

x

Coils

## The multidimensional measurement model

When we apply the dipole model, we establish a common coordinate system by compute the sensor parameters in source coordinates, applying the linear transform :

If is represented by a 4x4 linear homogeneous transform matrix this is implemented by matrix multiplication.

is a measured 3x3 coupling matrix, while , the coupling predicted by the measurement model. The magnetic model eqn so dipole, se dipole gives a scalar coupling between a single source and sensor coil. To get the full 3x3 we apply the magnetic model to all combinations of source and sensor coils.

## Concentricity and orthogonality:

Frequently the source and sensor coils are wound around a common center (*concentric*), in a cube or sphere configuration. [picture] As well as being more compact than multiple separate coils, this also simplifies the pose solution problem [cite Kim18, and everyone back]. It might additionally be assumed that the coils are orthogonal, precisely aligned on the source and sensor coordinate axes.

But even if the coils are fabricated to approach the concentric orthogonal ideal, there is always some deviation, so accuracy can be improved by adding calibration parameters to model the coil position and orientation error. This is in part a tradeoff between making the coupling measurement obey a simple measurement model, versus adopting a more complex measurement model. But highly non-concentric coil arrangements may also be useful [cite planar source work], in which case we must model the actual coil configuration.

## Fixture transforms:

When calibrating to a high accuracy, one challenge is that the calibration setup itself has unknown kinematic parameters. What exactly is the pose of the source and sensor with respect to the coordinate systems of our calibration fixtures? And what is the pose of the source fixture with respect to the sensor fixture?

EMTs are mainly used for measurement of relative motion, and neither the source nor the sensor have a precise mechanically defined coordinate system. That is, the accuracy is not defined with respect to any directly measurable marks or edges on the source and sensor. We could assume that the coordinates are nominally centered within the source or sensor, and aligned with mounting holes or other features, but we still can’t mechanically measure to a point inside an object.

Fortunately, we can evade this problem because when our concern is purely with relative motion, then any centering or alignment error matters very little. We introduce additional *fixture transforms* which absorb the unknowns of the calibration setup. [picture of the full pose diagram] The calibration parameters are augmented with the fixture transforms, and the calibration optimization solves for these together with the parameters needed for the measurement model [Does measurement model include fixture transforms or not? Probably not, but then we need another Greek letter or something.]

Specifically, we need to model ground the truth kinematics so that by independent means we know the correct tracker output .This is done by a sequential combination of known motions and fixture transforms . The can be thought of as the joints in a robot arm, and the are the fixed links that connect those joints. The are precisely known, and are varied during the calibration data collection, while the are constants that are initially only imprecisely known. The calibration positioner has three variable links, the source fixture, the stage, and the sensor fixture, so there are three pairs of , giving this equation for the kinematic chain:

# Optimization methods

Fig. : Pose diagram for calibration position ground truth, including fixture transforms.

Source

Sensor

Combining the magnetic and fixture parameters, we have:

giving 54 optimization variables in the general case. There are several reasons why we do not simultaneously optimize all these variables:

1. Stepwise optimization, gradually increasing the optimization state space, greatly helps convergence when initial values are poorly known.
2. There are redundant degrees of freedom in the model; these variables tend to “run away” in opposing directions, to no effect, but preventing convergence.
3. Depending on the test data, and the result calibration desired, we need to restrict the optimization in other configurable ways.

Control of the optimization is implemented by the bounds specified on the state. An optimization variable is initialized to a desired value, and then forced to not change by limiting the bounds to that value. Which values are “frozen” in this way is configured as needed, without any change in the optimization state space or objective function.

## Initial values

Optimization will not converge, or will be unreasonably slow, if the initial values are “too far off”. This difficulty is vaguely defined, and in practice a certain amount of trial and error (and head-scratching) can be expected. When available, the best initial value is a previous calibration result. If there is a change in the setup then it is worthwhile manually modifying an old calibration to approximate the expected new calibration.

Some initial values are non-critical. Coil formulas or magnetic simulations can predict the gain of the source and sensor. This is useful for coil design but is not needed for the calibration optimization. If the coil arrangement is very roughly concentric and orthogonal, and gain is not too far from 1, then it will often work well to initialize with .

The most critical initial state is the fixture transforms, since these affect the sign and gross rotation of the predicted coupling . When the fixture includes large rotations, it is easy to make an error in manually encoding the 3DOF rotation of a setup.

## Stepwise optimization

In practice it works well to use a rough guess for initial values, then first optimize the fixture transforms to find out which way is up, then add sensor moments , and then proceed to a full optimization. If optimization fails to converge to a reasonable residual, especially if there is unexpected asymmetry between source or sensor coils, check for hardware faults such as sign inversion on source or sensor signal paths.

## Redundant degrees of freedom

### Magnetic parameters vs. fixture transforms

If all of the parameters are free, then the source and sensor coordinate frames are undefined. We chose to tie the coordinate frames to the source and sensor coils by setting:

(66)

This places the source and sensor origins at the z coils, with z axes aligned. To constrain the rotation about z, we force the x coil y component to 0 also: [confusing, change 123 to XYZ also?]

### Source vs. sensor gain

The source and sensor gains (moment magnitudes) are not independently measurable; we only observe the product of the two. So we must fix either the source or sensor gain ( and in eqn (66)). We initially force , attributing all of the system gain to the sensor.

### Data degrees of freedom vs. fixture transforms

Data collection is time-consuming, and optimization runtime increases with the input data, so it is useful to do preliminary calibrations using reduced data. It is particularly convenient to omit fixture motions, since these are manual. But when some degrees of freedom are not exercised in the calibration motion pattern, then it is impossible to uniquely identify all of the fixture transform variables, so some must be frozen. The source fixture cannot be determined without source motion , nor can be fully determined without . These underdefined fixture transforms do not affect the pose solution since the only exist to absorb unknowns in the calibration setup.

## Reduced models

We may choose to fix model parameters to get a reduced order output model. In a concentric model, all the coil locations are zero: .

## Pose representation:

In both the pose solution and in the calibration optimization the optimizer state includes 6DOF poses. The three translation DOF cause no problems, but many representations for 3DOF rotations (such as the rotation matrix and quaternions) have more than three elements. These excess degrees of freedom are normally eliminated by invariants such as orthonormality of the rotation matrix, but the optimizer does not know how to do this. Euler three-angle representations do not have excess degrees of freedom, but suffer from gimbal lock and numerical instability. We have had good results using a *rotation vector* representation during optimization. This is closely related to the axis-angle representation , where is the unit vector for the axis of rotation, and is the rotation in radians. The corresponding rotation vector is . The complete representation for a pose is then: . This representation is converted to a linear homogenous matrix when the linear transform must be applied.

## Result accuracy/termination

What residual can be expected? What result accuracy is necessary? Sensitivity? Error is normalized

# Pose space error correction

When high accuracy is required over a well-defined workspace, it is possible to measure the error at many poses, and then apply a correction to the pose solution. This can be done via an interpolated lookup table. [cite Polhemus, VR work] One advantage of this is that it is possible to correct for environmental magnetic interference, but it can also correct intrinsic errors in the tracker calibration. The main shortcoming of the lookup table approach is that it requires a very large number measurements, and also can’t be extrapolated beyond the calibration volume.

In microsurgery the accuracy requirement is high, but the workspace is inherently constrained. We have found that the Direct Linear Transformation [cite] on the position significantly improves position accuracy.[reference result table] [Equation for transform here?] Since this is implemented by a 4x4 transform matrix, it is fairly low order, so does not require additional calibration data. The linear correction also has a stable extrapolation, and remains available beyond edge of the calibrated volume, degrading gradually.

## Pose solution

Need something here. Optimization, concentric, kim18, initialization, UKF, much is outside our scope, but need to at least say how we got the pose results we are testing. Hemisphere.

# Performance evaluation

## Test configuration

Dipole/premo configuration description, workspace. Minimum viable description of ILEMT hardware. r/d workspace regieme? Sample calibration?

## Methods

Specifically how we ran the calibration? Could walk through the calibration step buildup, starting with very rough initial value. Valuable, but also redundant with initial values/stepwise discussion. Good exercise to re-bootstrap the calibration.

Also, need to describe the calibrations which reported results are based on, and how those calibrations were done. Residuals?

Data hygiene, discuss test patterns here rather than earlier?

## Position and rotation error

Report error as magnitude of the position and rotation error vectors, max and RMS

[how to find pose error using transform matrices? Maybe we don’t need as much detail on testing as on calibration?]

[Figure: workspace error vector plot]

[Table: summary stats, no source fixture, no source fixture + linear, with source fixture. Concentric?]

Table : error vs. source type

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Calibration type | Dipole approximating source | | | | Concentric source | | | |
| XYZ (mm) | | RxRyRz (degrees) | | XYZ (mm) | | RxRyRz (degrees) | |
| RMS | Max | RMS | Max | RMS | Max | RMS | Max |
| (default) | 0.322 | 0.828 | 0.280 | 0.640 |  |  |  |  |
| Corrected | 0.202 | 0.454 | 0.280 | 0.640 |  |  |  |  |
| Concentric | 28.259 | 73.798 | 20.135 | 57.215 |  |  |  |  |

## Linearity

Measuring position error on the sort of wide-spaced data used in calibration does not directly test what relative accuracy can be expected during fine scale motion during fine scale motion. It would require an entirely impractical number test points to densely cover the 6DOF workspace. We *expect* that the error will vary smoothly across the workspace, so that a small motion will experience less error than the worst-case full-workspace error, but we have little idea the degree to which this is so.

Table : effect of source fixtures

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Calibrate data: | Test data: source fixtures? | | | | | | | |
| source fixtures? | No source fixture | | | | Source fixture | | | |
| No | 0.322 | 0.828 | 0.280 | 0.640 | 1.378 | 3.867 | 0.662 | 1.758 |
| Yes | 0.431 | 1.110 | 0.330 | 0.829 | 0.700 | 2.316 | 0.431 | 1.090 |
|  | RMS | Max | RMS | Max | RMS | Max | RMS | Max |
|  | XYZ (mm) | | RxRyRz (degrees) | | XYZ (mm) | | RxRyRz (degrees) | |

Another shortcoming of the absolute error measurement is that it does not characterize the amount of cross-coupling between the measurement DOF, in particular the spurious rotation caused by translation, and vice-versa.

An alternate approach to error is *linearity*, and in particular Integral Non-Linearity (INL) and Differential Non-Linearity, as used in analog-digital converter specifications. The error experienced over a small motion is (differential) non-linearity, a deviation in the derivative of the response. The integral non-linearity is the total deviation of the measurement from the desired response (absolute error).

To implement linearity testing we abandon the idea of densely sampling the workspace, and instead make dense sweeps, varying only one pose component at a time. This doesn’t solve the problem that it is impossible to test all poses, but does densely test *some* part of the workspace, and also gives a different way to present results.

What is actually measured is the INL, which has to be differentiated to get the DNL. To minimize noise introduced by the differentiation we use a Savitzky-Golay filter (.

[Figure: on-axis INL/DNL. Off axis? DNL is confusing]

[Figure: cross coupling?]

[Table: max INL, DNL by axis, cross coupling]

# Conclusion

We have described in detail EMT calibration methods that can achieve [blah performance over blah workspace]. We expect that most of these methods have been used before; the point is to gather them in one place, establish that they work, and give some understanding of why they work. The primary research innovation is the characterization of tracker performance with greater precision and comprehensiveness than seen before. Characterization of tracker linearity and the rotation/translation cross coupling are particularly important for understanding the performance achievable during small motions.

[might go in related work] A huge number of magnetic trackers have been developed in the nearly 50 years that this technology has existed. EMTs have been a niche technology, with optical and computer vision techniques being more easily applied, often more accurate, and increasingly much less expensive due to the high production volume of cameras and wide availability of computer vision libraries. The niche of EMTs remains those uses where clear sightlines cannot be guaranteed.

But the use of EMTs is often not considered even when they might be appropriate because of the need to largely reinvent the technology each time. While the ILEMT tracker has specific aims for medical application, we hope that our open-source signal processing code and hardware designs will help more engineers succeed in applying EMTs across a broader scope. It is only the small size of the market that causes standalone EMTs to be so expensive.

[estimate of number of trackers, number of papers on technology development, number of application papers] it is a reasonable graduate student project to implement an EMT, and then develop refinements of signal processing or source and sensor configuration.

1. This paragraph of the first footnote will contain the date on which you submitted your paper for review. It will also contain support information, including sponsor and financial support acknowledgment. For example, “This work was supported in part by the U.S. Depart­ment of Com­merce under Grant BS123456.”

   The next few paragraphs should contain the authors’ current affiliations, including current address and e-mail. [↑](#footnote-ref-1)