[[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

The ILEMT (In-Loop ElectroMagnetic Tracker) has the 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 achieved primarily in the hardware design (cite lame paper), and not by calibration methods.

A magnetic position tracker measures the *pose* (position and orientation) of a *sensor* with respect to a magnetic *source*. This is a 6 DOF (Degree Of Freedom) measurement. Pose is the linear transform which maps the source coordinate frame onto the sensor coordinates, which can be represented as a 4x4 linear homogenous transform matrix.

A single field measurement is not sufficient to determine the pose, so the source incorporates distinctly modulated electromagnet source coils, and the sensor 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 with respect to the sensor coil.

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 smaller while still maintaining the same signal to noise ratio. With this relatively large source, its magnetic aberrations tend to dominate the measurement model error.

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 also be supposed that the coils are orthogonal, precisely aligned on the source and sensor coordinate axes. But even when 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. Highly non-concentric coil arrangements can also be useful [cite planar source work], in which case we must model the actual coil configuration.

Some EMTs use other sensing principles and source modulations [cite Ascension, rotating field thing, etc], but here we consider the most common approach, where the source coil has a sinusoidal drive voltage and the sensor uses inductive pickup coils. These methods are applicable to any EMT that operates based on pairwise measurement of coupling between source and sensor axes.

# 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]

more

## Positioning

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. It may be possible to calibrate a low order measurement model using only translation data, but it then remains unclear what accuracy can be expected when the sensor is rotated. Compounding the difficulty, the sensor positioner must also be non-metallic to avoid interfering with the measurement.

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 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] (robot arm, coordinate measuring machine, etc.) Using such distance-based magnetic isolation it is impractical to achieve positioning along all rotation axes. Some sort of precise non-metallic rotary positioner is necessary.

## 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. Frequently only one sensor orientation is used, leaving us guessing about the position error caused by sensor rotation. While not comprehensive, these relatively simple procedures do reveal significant variation across EMTs and their use environments. cite

## 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 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. [cite VR]

# The calibration problem

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 choosing a suitable anddetermining

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

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

The calibration optimization state will include 6 DOF poses. The three translation DOF cause no problems, but many representations for 3 DOF 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 chosen 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:

|  |  |
| --- | --- |
|  | (1) |

A pose or linear transform may be represented as either this vector or as a linear homogenous matrix, but must be in matrix form when the linear transform is to be applied.

# The measurement model

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 40 mm and the sensor 15 mm, while the desired accuracy is . Due to the physics of magnetism, this is not as problematic is it seems. For distance from the source , and source coil of size , as increases to , the field converges to a dipole.

## The dipole approximation

The dipole is the simplest physical model for the source magnetic field, yet often works well in magnetic trackers, so we will use it here. For calibration the details of the model matter little, and higher order models can be used. Where the particular magnetic model does become important is in implementing the pose solution, which is outside our scope.

A dipole possesses only two parameters: location and vector moment (Fig. 1). If the sensor is located at , then the field is:

|  |  |
| --- | --- |
|  | (2) |

where , , and is the permeability of the medium. is a vector, which each individual sensor coil measures as a scalar voltage along some sensor moment . The dipole form of the sensor response is the scalar product of the sensor moment with the field:

Fig. : Pose diagram for kinematics of the measurement model. The pose maps from source coordinates to sensor. Each coil has an independent location and moment within its frame. These are grouped into the matrices.

Source

Sensor

z

y

x

Coils

z

y

x

Coils

Magnetic field

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

*Source*

*Sensor*

|  |  |
| --- | --- |
|  | () |

.

## Coordinate notation

These basic magnetic models are coordinate free, but we must introduce multiple coordinate systems to parameterize the full measurement model across multiple source and sensor coils. 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:

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

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

## The multidimensional measurement model

The dipole model (2) (3) gives a scalar coupling between a single source and sensor coil, while we must find , the full 3x3 predicted coupling matrix. The source and sensor parameters are defined in separate coordinate systems (Fig. 2). We establish a common frame by transforming the sensor parameters into the source coordinates. This is done by applying as a linear transform (matrix multiplication):

We then iterate over all combinations of source and sensor coils to find the pairwise couplings in .

## 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 where is the source fixture with respect to the sensor fixture?

EMTs are mainly used for measurement of relative motion; neither the source nor the sensor have a precise mechanically defined coordinate system. We can exploit this to remove the fixturing uncertainty by introduce additional *fixture transforms* which absorb the unknowns of the calibration setup. The calibration parameters of the measurement model are augmented with the fixture transforms, and the calibration optimization solves for these unknowns also

Specifically, we need to model the ground truth kinematics so that by independent means we know the correct tracker output .This is done by a sequential combination of fixture transforms and known motions . The can be thought of as the joints in a robot arm, and the are the fixed links that connect those joints:

|  |  |
| --- | --- |
|  | (5) |

The are known, and are varied during the calibration data collection, while the are constants that are initially only imprecisely known. The “stage motion” is the combination of the four motorized motions (X Y Z ) with the sensor fixture motion (. We mechanically align the stage axis with the sensor fixture origin, so that the sensor fixture motion can be combined into the stage motion without any intervening unknown **.**

# Optimization constraints

Combining the magnetic and fixture parameters, we have:

We do not optimize all 54 potential variables simultaneously. One reason is that there are redundant degrees of freedom in the model. These variables can “run away” in opposing directions, to no effect, but preventing convergence. (See also ref appendix.)

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

|  |  |
| --- | --- |
|  |  |
|  | (6) |
|  |  |

This places the source and sensor origins at the Z coils, with coil Z axes aligned to the Z coordinate axes. To constrain the rotation about Z, we force the X coil Y component to 0 also:

These constraints cause the fixture transforms to absorb any pose offset between the Z coil and the fixture.

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: or in (6). We initially force , attributing all of the system gain to the sensor. (The pattern of forced values may be clearer in Table I sample calibration.)

Fig. 5: Pose diagram for kinematics of measurement error. is the tracker output and is the error in this output. is the ideal perfectly accurate pose, which is not directly known. Allowing for stage error, we construct from the stage kinematics, and compare to , instead findingthe calibration error . This differs from by the stage error , so does not represent the full uncertainty of the measurement.

Source

Sensor

EMT output

Stage kinematics

True pose

# Pose solution and error correction

Magnetic tracking is an *inverse problem*. We can calibrate a measurement model to predict the coupling we expect to see in sensor pose , but this does not tell us how to solve the inverse problem: given a measurement , what is the best estimate of the pose ? 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 .

In microsurgery the accuracy requirement is high, but the workspace is inherently constrained. 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. 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.

As with the calibration of the measurement magnetic model , it is important that output correction calibration data cover the intended workspace, but because of its ad-hoc non-physical character it tends to generalize even more poorly outside of the calibration volume than does the magnetic model. This is particularly true when using an interpolated lookup table [cite Polhemus, VR work].

General linear transformation is a simple method for transforming spatial data using matrix multiplication. Since this is implemented by a 4x4 transform matrix. It is fairly low order, so will not over-fit the calibration data, and gives has a stable extrapolation beyond edge of the calibrated volume, degrading gradually. A general transform matrix differs from a linear homogenous matrix in that all 16 elements are unconstrained. It is neither orthogonal nor normalized, so is not a rigid body transformation, and can compensate scaling, shearing and trapezoid distortions. An arbitrary point is corrected to :

|  |  |
| --- | --- |
|  | (7) |
|  | (8) |

During calibration we must determine the which minimizes the mismatch between every calibration point and its ground truth position. General nonlinear optimization can be used, but we have found the simple Direct Linear Transformation algorithm to work well here (although it is known to suffer numeric weakness).[cite] Only position is corrected using this method. See [Table II].

# Characterizing tracker error

To determine the accuracy of a calibrated EMT we compare the output to the true pose . We already require a precise positioner to collect the calibration data, so precise test data is also easily collected, however the methods of finding the error are worth clarifying. Fig. 5 is a pose diagram of the error kinematics. is the error of the pose measurement , so:

|  |  |
| --- | --- |
|  |  |
|  | () |

is itself a 6DOF transform. If the error is computed as a transform matrix by (9) and converted to the eqn. (1) rotation vector format, then the translation and rotation parts represent the error in a somewhat interpretable form. However, it is often preferable to reduce the 6DOF error to translation and rotation magnitudes by taking the vector magnitude of the 3DOF pose parts. This error metric avoids all difficulties with angle wrapping but does not tell us the error direction.

## Uncertainty of stage position and tracker measurement

What we wish to determine is the error that can be expected in measurements made by the EMT. This typical error is the measurement *uncertainty*. Fig. 5 reminds us that we do not know the true , only constructed using (5), so our actual error computation is:

and each have their own uncertainties, so is an uncertain measurement of the desired . (The fixture transforms can be regarded as defining the coordinate frames, so have no uncertainty.)

Let be the uncertainty of a measurement , and the combined uncertainty of a composite measurement, then in standard measurement practice: Cite

|  |  |
| --- | --- |
|  | () |

This can be explained by supposing that are the standard deviations of independent normally distributed random variables.

We wish to find the uncertainty of the pose measurement: . If we knew the true error then we could estimate as the standard deviation (RMS value) of this error across the test points. Let be the uncertainty estimate from the statistics of the available . alone underestimates , but we can compensate by incorporating the stage uncertainty:

|  |  |
| --- | --- |
|  | () |

## Systematic and random error

Error can be divided into *systematic* (repeatable) error and *random* error. The entire calibration procedure is an attempt to minimize the systematic error in the tracker measurement. The stage has its own systematic and random positioning errors. Since the same stage is used for both calibration and testing any systematic error in the stage is present in both the calibration and test data. If the calibrated measurement model incorporates systematic stage error (such as scale factor), then the true error becomes correlated with the stage error .This correlation simultaneously increases the true error and reduces the measured calibration error . In the extreme, if random stage error is minimal and perfectly models both the magnetics and the systematic stage error then from (11) will go to zero. Even so, (11) remains reasonable since .

When choosing the model it is important to consider that if has many degrees of freedom (such as with a lookup table) then it will overfit to include the stage error, and this increases the true error even as it reduces the calibration error. If the stage is the only position reference then this effect is invisible. A lower order model may appear to give worse error, yet actually be more accurate.

If the calibration data has far more DOF than does then random positioning error will be averaged out in the calibration, while random error during testing will increase . This is the reverse of the situation with systematic error, since true error does not increase, but does.

## Magnetic error in the test setup

Any magnetic or highly conductive material (mainly metal) must be excluded from the test setup since this is another source of systematic error. As with stage error, this increases the true error seen in other magnetic environments, yet may go undetected by the calibration error . Avoiding metal is difficult because standard positioning devices such as stages and robot arms almost invariably contain a considerable amount of metal. Fortunately, any 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. Metallic interference is greatest when the interferer is near the source or sensor, or in between them, so this should be avoided. A way to test for the presence of magnetic interference is to attach the source and sensor to a handheld bar and move this about the workspace, finding the distance from each component that causes a detectable change in output.

## Stage and fixture design

In the Fig. 3 calibration stage we achieve a separation by placing the sensor on a fiberglass pole, but this in itself only permits rotation about a single axis *Rz*. The manual source and sensor fixtures in Fig. 4 create additional 90º rotations of the source and sensor. We found that the black granite is slightly magnetic, so we raise the source fixture above the surface plate. The pink granite plate (Starrett) in the source fixture does not have any detectable magnetic effect. The sensor fixture allows any sensor axis to be aligned with the stage *Rz*, permitting automated testing on all axes, while the source fixture only supports certain 90º rotations.

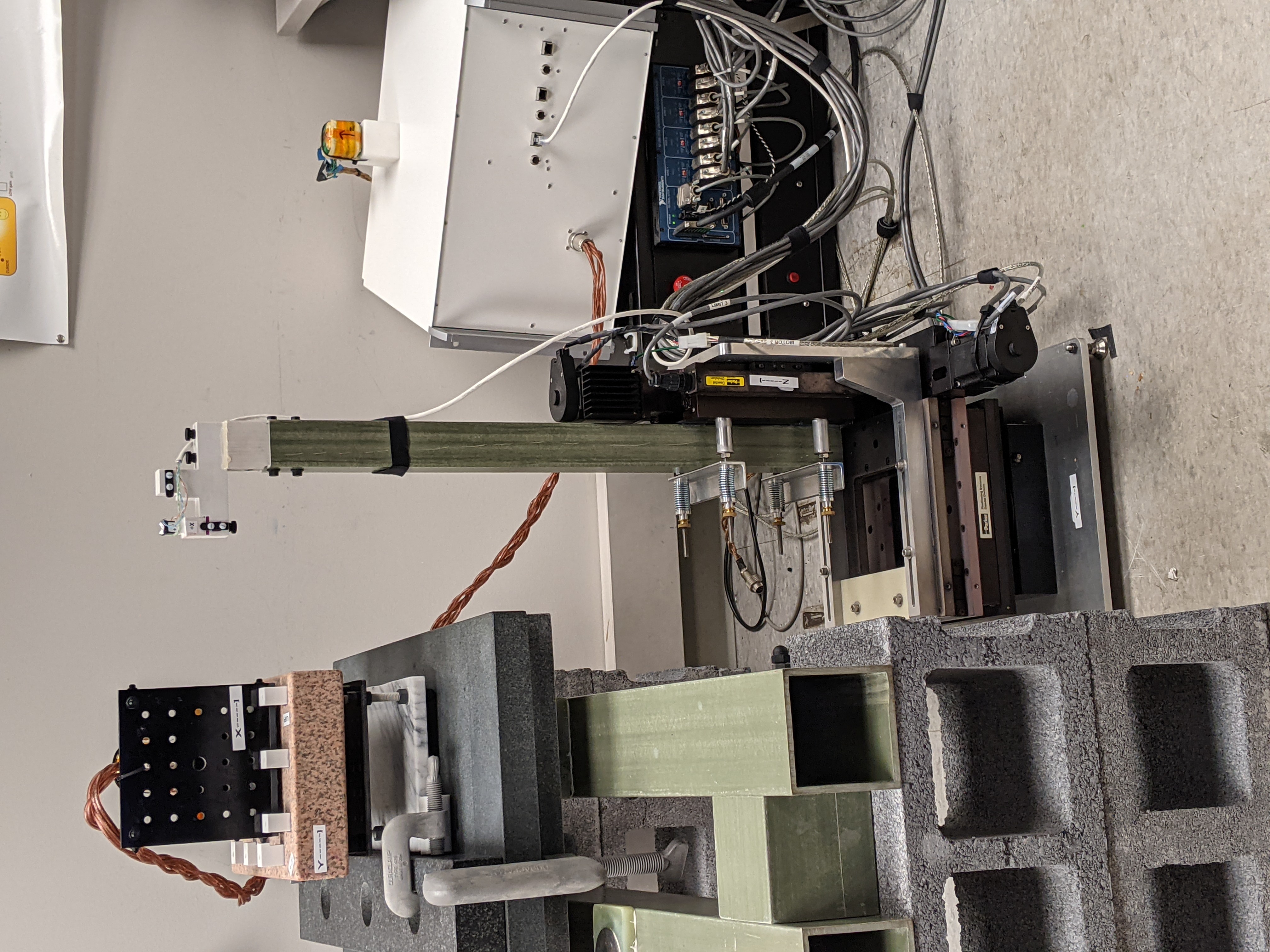


Fig. : The stage provides automated motion over a 100mm cube translation workspace, with rotation about Z. The fiberglass pole extends the sensor away from the metal components in the stage so as to avoid interference with the magnetic measurement.

Sensor

Source

Stage: XYZ + *Rz*

Fig. : The source and sensor fixtures permit precise 90-degree rotations of the source and sensor, which enables testing across the entire workspace, supplementing the limited translation and single rotation axis of the stage. Thee surface plate permits large precise planar and straight-line sliding motion.



Sensor

Sensor fixture

Stage pole

Source fixture

Source

Surface plate

Source slider

We wish to minimize the stage error , and we must also estimate , the uncertainty of the stage position.

The magnetic isolation pole has a large 0.7 m moment which converts any off-axis stage motion into XY position error

The stage components are Parker Daedal 806004CT (Z axis), 310062AT (XY axes), and 20801RT (Rz axis). These are driven by closed loop microstepping motors with 6400 count encoders. The nominal resolution of 0.8 µm and 5.4 µrad, but actual position error is far larger.

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

## Data collection patterns

The calibration and test data can and *should* be different. In any sort of modelling, the model should be evaluated with different data than that used to fit the model. If not, performance may be exaggerated by overfitting. Yet the calibration data must cover the same sort of variation that will be seen in operation (and during testing). Also, in EMT practice, because of the strong constraints of the dipole model, it works well to use smaller data for calibration than testing, which speeds the calibration optimization.

[fixture position thumbnail pictures]

For calibration, the motion pattern we use is a 100 mm cube with a 3x3x3 grid of test points, and five Z axis () rotations spaced across degrees, or 135 points in each fixture configuration. The test data has a 5x5x5 grid with the same rotations, giving 625 points. The sensor fixture has three orientations, so that each axis of the sensor is successively aligned with the stage axis. This pattern is collected in each of the source and sensor fixture rotations. There are four source fixtures and three sensor fixtures.

Broadly, rotation of the sensor helps to identify the sensor response, while rotations of the source identify 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. Since one configuration is in common, this requires six repeats of the stage motion pattern, with manual fixture repositioning in between. With this fixture motion the full test data is 3750 points. Even with this coarse grid spacing it is impractical to densely sample across a 6DOF workspace.

## Tested hardware

Source drive and sensor readout are provided by the ILEMT open source EMT reference design. This hardware is optimized for low noise and high measurement speed rather than cost. [cite] Typical performance with the tested sources is RMS position noise at 1500 samples/sec (500 Hz bandwidth), at . This a frequency-domain multiplexed AC design, where each source coil continuously emits a sinusoidal field at a distinctive frequency (7.5 kHz, 10.5 kHz and 13.5 kHz).

The coupling recorded for each tested pose is the mean of 1500 measurements, so random variation in the coupling due to broadband noise is negligible. Non-repeatable variation in the coupling is dominated by drift. The collection of calibration and test data takes approximately 8 hours, so drift over this time implicitly included in the calibration error measure. To minimize this drift we powered up the tracker at least 2 hours prior to data collection.

The only sensor used is the Premo Magnetics 3DCC10-A-0600J. This is a fairly large 17x15x14 mm sensor, which improves the signal to noise ratio. While we have characterized performance with a variety of sensors, we find that with its smaller size relative to the source, the sensor has less effect on the accuracy. Using only one sensor simplifies our presentation.

Results are given for two different source designs: *concentric*  and *dipole-approximating*. The concentric source follows the most common design for EMTs: three orthogonal coils wound around a ferromagnetic cube core. The dipole-approximating source uses three separate coils, each with the geometry from Ref. [1], which gives a close approximation of a dipole field. Because the coils are distinct, this source is non-concentric. Both sources are hand-wound prototypes, and likely have less ideal fields than might be seen with a production source.

Fig. : The concentric source has three orthogonal windings around a single ferrite cube core, giving a high field strength for its size. The dipole-approximating source uses three air-core coils with a geometry that gives a field more closely resembling a magnetic dipole at close ranges.



Dipole-approximating

Concentric

50 mm

For both sources the coil diameter is , which with a testing radius of gives . Operating at this relatively short range reduces measurement noise, but also increases the difficulty of accurate calibration because the deviation from a dipole model can be significant. [cite] ???.

These sources and sensor were chosen to meet ILEMT requirements, but from the viewpoint of calibration techniques, many different source and sensor types might be used. The purpose of calibration is to model whatever source and sensor combination is chosen.

## Calibration

These results use three calibration types: the *default* condition uses the general dipole magnetic model, while *concentric* restricts the source and sensor coil positions to their coordinates origin (), and *corrected* applies the DLT output correction to XYZ (section blah). Also, the data may include source fixture motion, or may only use a single source fixture (sensor always on the same side of the source).

Table I is a sample calibration for the dipole-approximating source. These are the and matrices from (4) (coils are columns). Note the values forced to 0.0 and 1.0 to avoid excess degrees of freedom (see §**Error! Reference source not found.**. This source is highly non-concentric, with 45mm offset. The sensor is designed to be concentric, but calibration finds position offsets similar in magnitude to the position accuracy we obtain below. Source and sensor non-orthogonality are fairly small, but still significant. The source X Z deviation is 1.6°, which would correspond to a 5.5 mm error at 200 mm.

Table I

Sample calibration (dipole approximating source)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Source position (mm) | | | Source moment | | | Sensor position (mm) | | | Sensor moment | | |
| X | Y | Z | X | Y | Z | X | Y | Z | X | Y | Z |
| 45.266 | -0.514 | 0.0 | 0.950 | 1.0E-02 | 0.0 | 0.144 | 0.051 | 0.0 | 0.158 | 4.2E-04 | 0.0 |
| 1.249 | 45.380 | 0.0 | 0.0 | 0.947 | 0.0 | 0.110 | 0.116 | 0.0 | 0.0 | 0.157 | 0.0 |
| -43.764 | -42.830 | 0.0 | -2.6E-02 | 1.7E-03 | 1.0 | -0.223 | -0.185 | 0.0 | 2.4E-03 | 1.9E-03 | 0.161 |

## Position and rotation error

Blah blah fig. 6

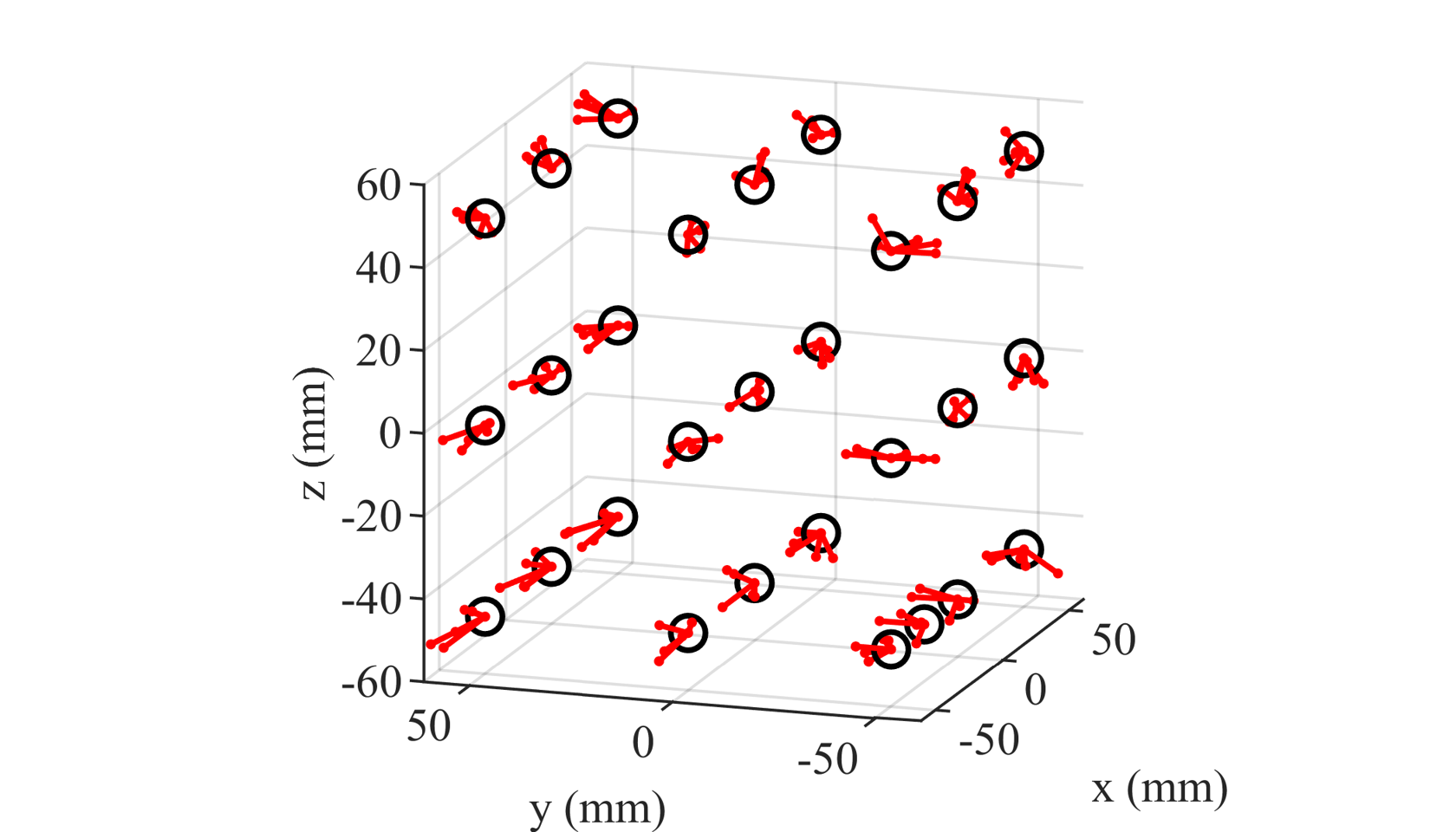


Fig. : Translation error across the workspace, under sensor rotation. The error vectors are exaggerated 35x. Each point has 5 error vectors because there are 5 rotations. Note the considerable effect of rotation on translation error.

Table II shows the rotation and translation error of the two calibrated source as the calibration type is varied. The errors reported are RMS and maximum vector magnitudes of the translation error (XYZ) and the rotation vector of the rotation error . Here both the calibration and test data were taken with three sensor fixtures, but no source fixture rotation. With the general calibration, the dipole-approximating source gives significantly better accuracy than the concentric source (empirical support for ref. blah). Also, the linear correction significantly improves the XYZ accuracy. Since this source is not concentric, the accuracy is terrible using a concentric calibration model. For the concentric source, the concentric model works has 2x the error of the general dipole model, which for some uses may be an acceptable sacrifice in order to enable a more efficient pose solution.

Table II

measurement error vs. Calibration type and Source design

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Calibration type | Dipole approximating source | | | | Concentric source | | | |
| XYZ (mm) | | RxRyRz (degrees) | | XYZ (mm) | | RxRyRz (degrees) | |
| RMS | Max | RMS | Max | RMS | Max | RMS | Max |
| (default) | 0.335 | 0.931 | 0.219 | 0.577 | 0.452 | 1.304 | 0.275 | 0.642 |
| Corrected | 0.202 | 0.525 | 0.219 | 0.577 | 0.283 | 0.767 | 0.275 | 0.642 |
| Concentric | 28.305 | 73.636 | 20.273 | 56.885 | 0.877 | 2.577 | 0.511 | 1.415 |

## Effect of source fixtures on accuracy

What happens when the calibration data does not cover the operation workspace? While we understand that a mismatch is not good, it may not be appreciated how severe the accuracy degradation may be. Table II shows the effect of source fixture rotation, present in either the calibration data, the test data, or both. This is with the concentric source and default calibration type (no correction, non-concentric calibration). When calibrated without source rotation, but tested with source rotation, the (no, yes) condition, translation error increases by 10x. Adding source rotation to the calibration improves performance in the (yes, yes) condition, but at some cost to performance in the (yes, no) condition. (The dipole-approximating source performed badly across source fixture motion, with the pose solution frequently failing to converge.)

Table III

measurement error vs. source fixture rotation

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Test data: no source fixtures | | | | Test data: source fixtures | | | |
| Calibrate data: | RMS | Max | RMS | Max | RMS | Max | RMS | Max |
| source fixtures? | XYZ (mm) | | RxRyRz (degrees) | | XYZ (mm) | | RxRyRz (degrees) | |
| No | 0.452 | 1.304 | 0.275 | 0.642 | 4.559 | 12.345 | 1.825 | 5.216 |
| Yes | 0.878 | 2.096 | 0.539 | 1.385 | 2.271 | 5.931 | 1.097 | 3.401 |
|  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Test data: no source fixtures | | | | Test data: source fixtures | | | |
| Calibrate data: | RMS | Max | RMS | Max | RMS | Max | RMS | Max |
| source fixtures? | XYZ (mm) | | RxRyRz (degrees) | | XYZ (mm) | | RxRyRz (degrees) | |
| No | 0.452 | 1.304 | 0.275 | 0.642 | 4.559 | 12.345 | 1.825 | 5.216 |
| Yes | 0.878 | 2.096 | 0.539 | 1.385 | 2.271 | 5.931 | 1.097 | 3.401 |

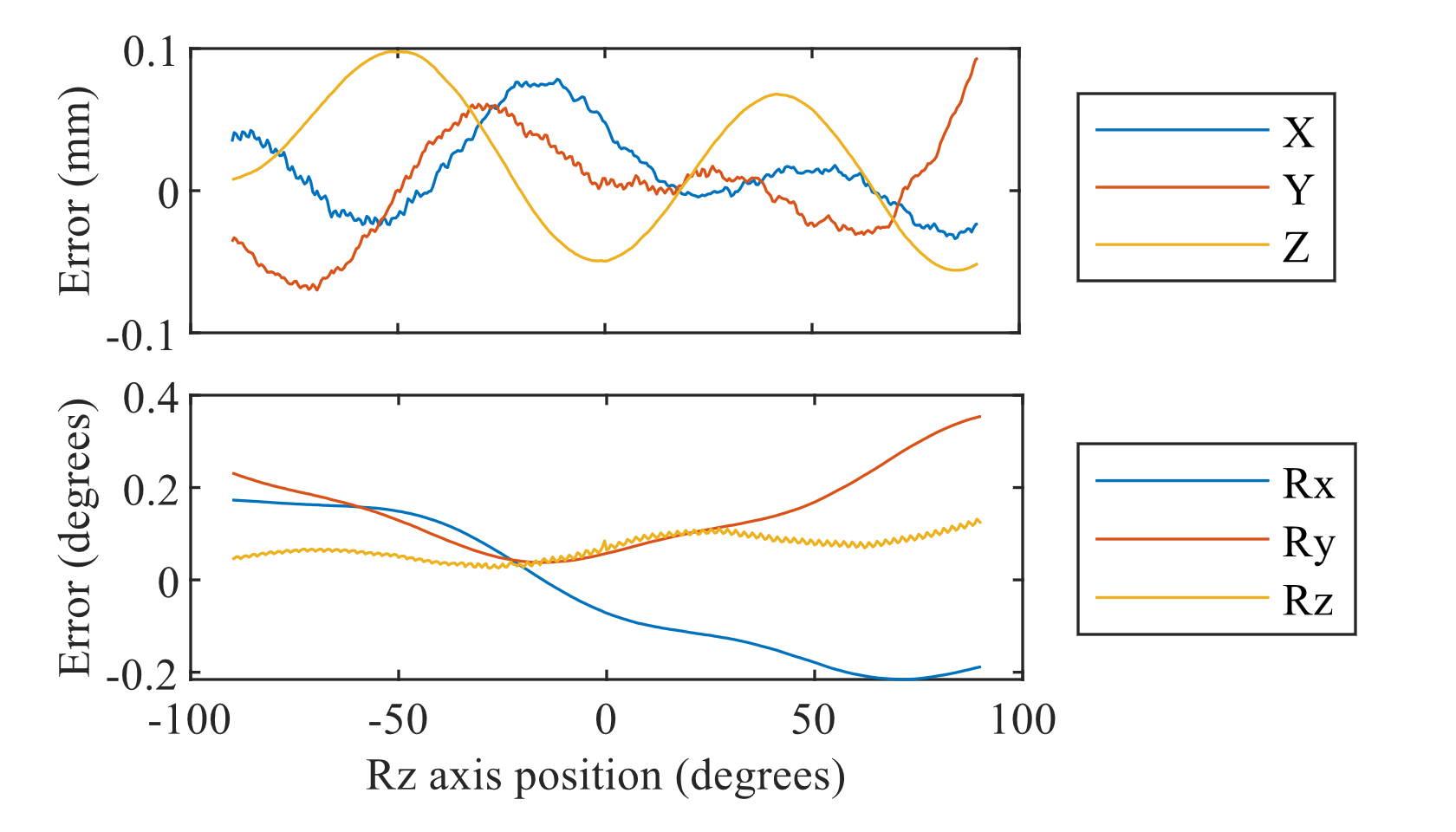
Calibrating across a larger workspace increases performance in the larger workspace, but at the cost of lower accuracy in a more restricted workspace. Since sensor motion is usually 6DOF, it is at a minimum necessary to calibrate and test in 6DOF. While it might not be required in some uses, commercial magnetic trackers usually support operating in any hemisphere of the source, which requires source calibration in different orientations. It is an entirely unreasonable assumption that accuracy will be acceptable in parts of the workspace which have not been tested.

## 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. In the grid data used in the above testing, the XYZ increment is 2.5 cm, while the position error may be as much as 100x smaller. It would require an entirely impractical number test points to densely cover the 6DOF workspace. We expect a degree of *linearity*: 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. To test this we abandon grid sampling, 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 enables useful decomposition of the error.

One shortcoming of the sparse grid error measurement is that it does not characterize the measurement cross-coupling. Any motion of the sensor will result in a change of all 6 measured DOF. We would hope that the largest change represents the true motion, but there are also spurious changes along directions that the sensor did not actually move in (Fig. 7). To simplify presentation we split these errors into *direct error* and *cross error,* and show the vector magnitude of each. Direct error is the XYZ error caused by XYZ motion (or error caused by motion), while cross error is cross coupling between rotation and translation.

Fig. : Pose measurement errors on all six axes in response to a rotation sweep on the Rz axis alone. (a) is the *cross error* (here translation caused by rotation), while (b) is the *direct* *error* (rotation error caused by rotation).

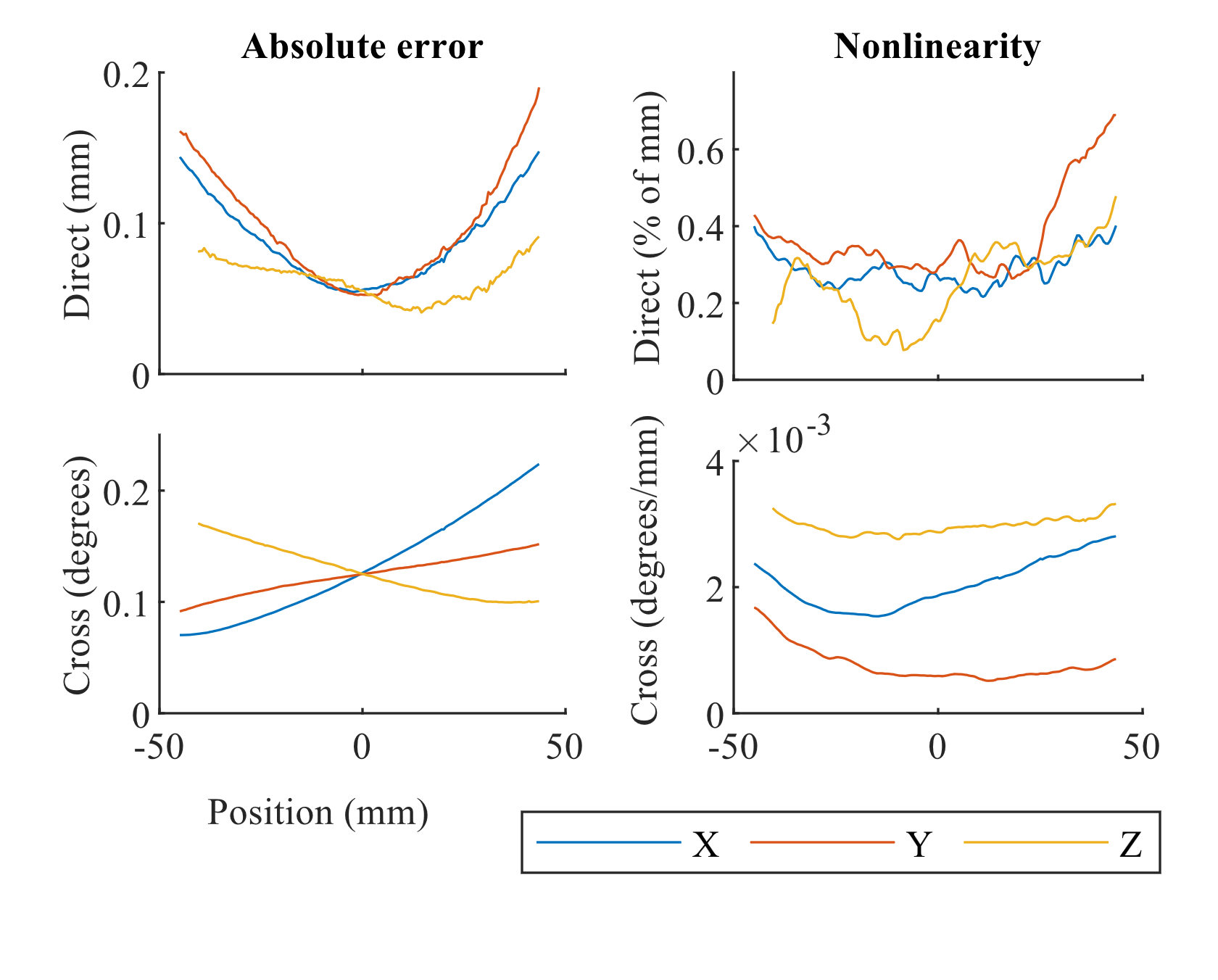


(a)

(b)

Fig. 8 shows the absolute error and nonlinearity resulting from sensor motion sweeps on all 6 axes. Measurements are at 0.5 mm, 0.5 degree increments over 100 mm, 180 degrees. This uses the dipole approximating coil, without source fixtures (as in Table I corrected condition). We quantify the linearity by the derivative of the error across the workspace (the magnitude of the error gradient). The linearity at a point tells us the relative error that can be expected across a small motion. The direct nonlinearity is dimensionless, and expressed as a percentage of the motion. While this resembles a scale factor error, it includes off-axis motion as well. Cross nonlinearity is the magnitude of the gradient of translation error with respect to rotational motion, and vice versa, with units mm/° or °/mm. To minimize noise we differentiate using a Savitzky-Golay filter (.

Fig. : Error and nonlinearity as each axis is swept across the workspace, one at a time. The absolute error is the vector magnitude of the 3DOF rotation or translation error caused by motion on the axis in the legend. Nonlinearity is the magnitude of the 3DOF derivative (gradient) of the error. As magnitudes, these plots are all nonnegative. The “U” shape of absolute error represents the response passing near an error minimum, not curvature in the overall response.



**Error from translation**

**Error from rotation**

In Table IV each Fig. 8 curve is summarized by RMS and max, and then these measures are combined by max across the curves (driving axes) in each plot. Absolute error is computed as in Table II, but is now significantly lower. This is mainly because the sweep data only varies one axis, so is on average closer to the center of the workspace, where error is lower. Note how the magnitude of the rotation-to-translation cross error is similar to the direct translation error. This reinforces the point that we cannot characterize EMT translation accuracy by using a translation-only test pattern.

Table IV

Cross coupling and nonlinearity during axis sweeps

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | Absolute error | | Nonlinearity | |
|  |  | XYZ | RxRyRz | XYZ | RxRyRz |
|  |  | mm | degrees | Percent | |
| Direct | RMS | 0.104 | 0.227 | 0.39% | 0.44% |
| max | 0.190 | 0.408 | 0.69% | 0.67% |
| Cross | RMS | 0.139 | 0.095 | 0.0030 | 0.0050 |
| max | 0.224 | 0.174 | 0.0033 | 0.0092 |
|  |  | degrees | mm | °/mm | mm/° |

Conclusion

How does accuracy compare to reports in the literature?

The dipole-approximating source had better accuracy than the concentric source when the source was not rotated. The concentric coil is more compact, and works better with in the presence of source rotation. Being nearly concentric, it also and permits a simpler pose solution, at some loss in accuracy.

You have to test effect of sensor rotation even if you are only concerned with translation because the cross-coupling can be large.

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.

Appendix I: calibration procedures

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.

(here Matlab lsqnonlin)

We have also seen this general pattern as the workspace is generalized from XYZ + only by adding sensor fixtures .

1. Stepwise optimization, gradually increasing the optimization state space, greatly helps convergence when initial values are poorly known.
2. 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 could be used to predict the gain of the source and sensor. While useful for coil design, it 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 . It is easy to get it wrong when attempting to visualize and numerically represent the 3DOF rotation of a setup that includes large fixture rotations.

## 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 finally proceed to a full optimization. If optimization fails to converge to a reasonable residual (ref section below or move), 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.

### Data degrees of freedom vs. fixture transforms

Data collection is time-consuming, and optimization runtime increases with the input data, so we wish to make preliminary calibrations using reduced data. It is particularly convenient to initially 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.

## Calibration residue

Table V shows the calibration residue for the various calibrations above. This is a measure of the quality of fit between the calibration data our chosen , and roughly predicts the attainable measurement accuracy. We can easily see that the (non-concentric) dipole-approximating source performs very badly with a concentric calibration, that adding source fixture motion is going to degrade accuracy, and that even with the concentric source it is useful to use non-concentric calibration.

Table V

Calibration residues, RMS per point

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Source type | |
| Calibration type | Source fixtures | Dipole approx | Concentric |
| (default) |  | 0.54% | 0.52% |
| Yes | 0.76% | 1.16% |
| Concentric |  | 20.8% | 0.71% |
| Yes | 36.6% | 1.89% |

However, this residue is over coupling matrices, and the pose solution’s sensitivity to coupling error varies according to the specific geometry of the test pose. If the magnetic configuration is unfavorable then the pose error is magnified (cite source placement). This likely explains why pose solution for the dipole-approximating source failed to converge when source fixture motion was present, even though the residual is smaller than for the concentric source.

There is no single answer to what an acceptable calibration residue is, but the measurement linearity is unlikely to be better than the calibration residue, and changes in which reduce the residue usually result in higher measurement accuracy.

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)