High-precision evaluation of electromagnetic tracking

David Kügler[1] · Henry Krumb[1] · Judith Bredemann[2] · Igor Stenin[3] · Julia Kristin[3] · Thomas Klenzner[3] ·
Jörg Schipper[3] · Robert Schmitt[2] · Georgios Sakas[1] · Anirban Mukhopadhyay[1]

Abstract

Purpose   Navigation in high-precision minimally invasive surgery (HP-MIS) demands high tracking accuracy in the absence of line of sight (LOS). Currently, no tracking technology can satisfy this requirement. Electromagnetic tracking (EMT) is the best tracking paradigm in the absence of LOS despite limited accuracy and robustness. Novel evaluation protocols are needed to ensure high-precision and robust EMT for navigation in HP-MIS.

Methods   We introduce a novel protocol for EMT measurement evaluation featuring a high-accuracy phantom based on LEGO®, which is calibrated by a coordinate measuring machine to ensure accuracy. Our protocol includes relative sequential positions and an uncertainty estimation of positioning. We show effects on distortion compensation using a learned interpolation model.

Results   Our high-precision protocol clarifies properties of errors and uncertainties of EMT for high-precision use cases. For EMT errors reaching clinically relevant 0.2 mm, our design is 5–10 times more accurate than previous protocols with 95% confidence margins of 0.02 mm. This high-precision protocol ensures the performance improvement in compensated EMT by 0.05 mm.

Conclusion   Our protocol improves the reliability of EMT evaluations because of significantly lower protocol-inherent uncertainties. To reduce patient risk in HP-MIS and to evaluate magnetic field distortion compensation, more high-accuracy protocols such as the one proposed here are required.

Keywords   Electromagnetic tracking · High-precision surgery · Metallic distortion compensation · Evaluation protocol · LEGO® phantom

Introduction

Electromagnetic tracking systems (EMTS) are currently the best option for continuously tracking surgical instruments in minimally invasive surgeries (MIS), where no line of sight is available. However, advantages of MIS do not translate to high-precision MIS (HP-MIS), because navigation without line of sight is not accurate enough [10,21]. Typical HP-MIS such as electrode placement (neurosurgery [19]), vestibular schwannoma removal and cochlear implantation (temporal bone surgery [16]) require navigation accuracies

below 0.5 mm. Current analyses and systems do not provide capable protocols to verify this. Yet, to improve the accuracy for HP-MIS, we require a better understanding of EMTS evaluations and their inherent reliability. Therefore, we propose (1) a novel evaluation protocol for EMTS targeting HP-MIS and (2) an analysis of its inherent uncertainty contributions with a novel level of rigor and detail.

While previous standardized EMTS evaluation protocols [1,3,5,8,11,13,17,20] quantify the tracking performance of EMTS, none of these protocols address the challenges specific to HP-MIS. Their drawbacks for HP-MIS are threefold: (1) lack of transparency for the protocol's inherent uncertainty, (2) no evaluation of small offsets (less than 30 mm) and (3) lack of analysis of relative tracking accuracy. Because of their impact on clinical safety for HP-MIS, we explain these drawbacks starting with *the protocol's uncertainty*. When combining multiple sources of measurement uncertainty, the overall uncertainty is dominated by the largest individual

✉ David Kügler
    david.kuegler@gris.tu-darmstadt.de

1   Department of Computer Science, Technischer Universität
    Darmstadt, Darmstadt, Germany

2   WZL Aachen, RWTH Aachen, Aachen, Germany

3   ENT Clinic, University Düsseldorf, Düsseldorf, Germany

Table 1 Positioning accuracies and grid sizes of EMTS evaluation protocols; u_{95} (95% uncertainty) $= 2\sigma$ (standard deviation)

Approach	Positioning uncertainty (95%)	Grid size
Board		
Hummel et al. [4,5,12]	Estimated[a]:0.102−0.248 mm	55 mm
Koivukangas et al. [8]	Estimated[a]:1.0 mm	20 mm
Haidegger et al. [3,17]	Claim/reproduced[b]:0.004/0.320 mm	88 mm
Proposed	0.021 mm	8 mm
3D		
Wilson et al. [20]	Estimated[a]:0.14−0.252 mm	$Y-Z$[c]: 10 mm
Lugez et al. [11]	Estimated[a]:0.54−1.22 mm	$Y-Z$[d]:12.5 mm
Reichl et al. [13]	Estimated[a]:0.506 mm	25 mm

[a] Estimates following standard guides [6]

[b] See Fig. 7 (95% = 2std. deviations)

[c] X (depth) irregular and not documented

[d] X not calibrated nor fixed

contribution. For small offsets and high precision, an in-depth analysis of the source is required to discriminate between EMTS and protocol. Unfortunately, Franz et al's guidelines for a "good protocol" [1] include neither evaluation nor discussion of the accuracy limits of the protocol itself (e.g., repeatability of sensor position [6]). We compare protocol uncertainty of related work considering effects such as fitting tolerances and machining accuracy in Table 1. For *small offsets*, the dependence of EMTS accuracy on offset distance [5] is unverified. Finally, *absolute tracking accuracy* is not relevant to the clinical scenario [1] (only relative accuracy is), since absolute positions cannot be measured in the absence of a common reference coordinate frame.

In this paper, we introduce a novel protocol consisting of a LEGO® phantom and an evaluation procedure suitable for high-precision evaluations. To avoid registration errors, we compare distances between positions. Our EMTS protocol is optimized for: (1) relative position tracking accuracy, (2) a small work region (most distances < 50 mm), (3) high precision (25x accuracy requirements) and (4) quasi-static motion characteristics. Because LEGO® phantoms do not achieve high precision for orientations with good repeatability, orientation errors are not part of this protocol. In fact, high-precision surgeries using instruments with sensors at the tip typically do not require sub-degree accuracy.

We establish the advantages of our high-precision protocol in multiple evaluations. A detailed analysis explains positioning uncertainties and protocol reliability. We evaluate the impact of our protocol in laboratory, c-arm and operating theater environments using a trakSTAR 3D Guidance EMTS. Finally, we present results of a rudimentary machine learning technique to compensate magnetic field distortion. High-precision protocols and compensation techniques like the one presented here may reinvigorate efforts to find more f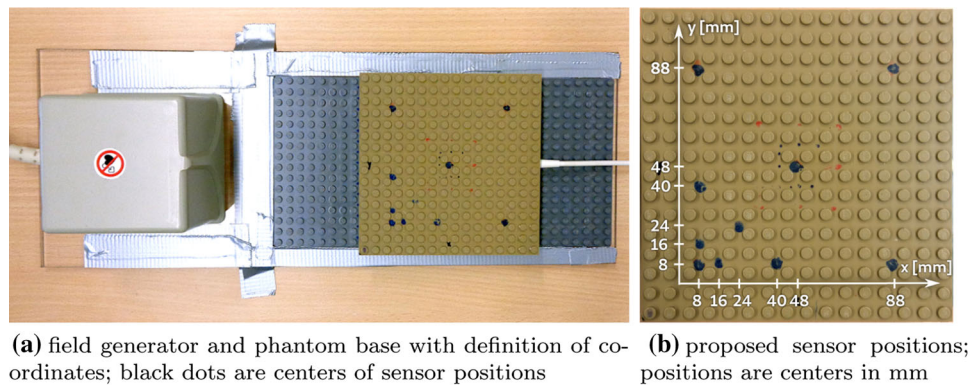lexible and sophisticated tracking methods for clinically viable HP-MIS. The specific design choices of the protocol proposed here are partly motivated by translation of EMTS navigation for HP-MIS from the laboratory to clinical applications.

Related work

Since Franz et al. [1] give a comprehensive overview on all aspects of electromagnetic tracking from systems and evaluation protocols to magnetic distortion compensation, we limit the related work to recent developments most similar to this work.

While several evaluation protocols have been proposed over the years, the Hummel's board [5] has gained the position of a quasi-standard evaluation procedure. Here, a carrier block with the EMTS sensor fixed to it is positioned at different locations on an acrylic board. The ground truth position of the sensor is assumed from the computer-controlled manufacturing machine's accuracy. A separate evaluation of the positioning (fitting tolerances of carrier block), hole position and the procedure's repeatability is not performed.

Haidegger et al. [3] have used LEGO® bricks to position the sensor. The authors assume "an inherent fitting accuracy of 2 μm" [3] for LEGO®. We were not able to reproduce this assumption (Sect. 4.1). Wilson et al. [20] presented a three-dimensional phantom with holes at varying depth. Some evaluations used EMTS Stylus [8,11] for positioning. Using a reference tracking system, Reichl et al. [13] constrain the accuracy by said tracking system (e.g., optical, OTS), which requires registration between coordinate systems. They show the dependency of the accuracy on the distance from the center of the tracking volume. The constraints on positioning uncertainty and grid size for positioning are collected in Table 1.

Fig. 1 Phantom and evaluation setup

(a) field generator and phantom base with definition of coordinates; black dots are centers of sensor positions

(b) proposed sensor positions; positions are centers in mm

There are two strategies for the reduction in magnetic field distortion on EMTS [1]: passive, such as shielding, and active typically based on machine learning [7]. The passive techniques translate well into clinical practice for specific distortion sources (e.g., c-arm), because no calibration is needed. Active methods on the other hand have also shown promise, but require environment-specific initialization. Current developments include the online estimation of magnetic field distortions due to metallic objects and compensation based on Kalman filter [15].

No research features an evaluation of the accumulated phantom uncertainty or results for small offsets between positions (< 20 mm). Since positioning uncertainties are not significantly better than the reported EMTS's uncertainty, it is hard to establish the source of the uncertainty. This further limits the evaluation of magnetic field distortion compensation techniques.

Methods

For our evaluation protocol, we design a highly precise phantom and a simple, quick and repeatable evaluation procedure. The *Phantom Design* (Sect. 3.1) is the combination of base board (Fig. 1a) and sensor block (Fig. 3a, LEGO® bricks with primary sensor). The *Evaluation Procedure* (Sect. 3.2) consists of four steps outlined in Fig. 2: phantom calibration, positioning uncertainty estimation, electromagnetic measurements and evaluation. Finally, we describe the learning-based model for compensation (Sect. 3.3).
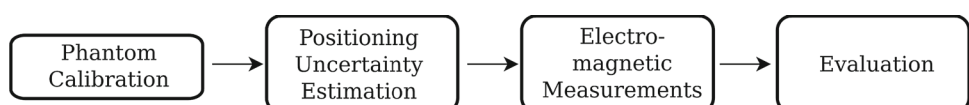
Phantom design

In order to make positions reproducible with high precision (i.e., low uncertainty), we improve previous LEGO® phan-

toms [3,17]. The base of our phantom is reinforced by bricks for stability (Fig. 1) with the reference sensor embedded (Fig. 3c). The largest uncertainty reduction is achieved by improving the fitting between base and sensor block (Fig. 3b improved to Fig. 3a). The primary sensor (tracked object) is rigidly attached to the sensor block independent of the design.

Evaluation procedure

We introduce two additional steps into the evaluation procedure as outlined in Fig. 2, which are not considered in the state-of-the-art protocols. "Phantom Calibration" and "Phantom Uncertainty Estimation" assess and ensure the reliability of the protocol. Figure 1b shows the 10 positions (marked with black dots), which are selected for sensor block placement to acquire measurements. They provide a good range of values to characterize the dependency on offsets (displacement distances of sensor block). Since we use distances between calibrated sensor positions on the base board for the procedure, *all evaluations are based on distances* and their errors. Varying relative board/field generator positioning increases diversity.

Phantom calibration In order to verify and correct the positioning of the sensor block (and by extension the sensor) on the base board, we measure its positioning with a reference system (Carl Zeiss Micura coordinate measuring machine, CMM). By repeating (10x) the placement (de- and re-attachment) of the sensor block on the base, we determine systematic and stochastic contributions of the process of positioning. Since we rely on offsets between different positions, absolute values are defined w.r.t. an arbitrary yet consistent origin. Assuming a 8-mm grid, our calibration represents

Fig. 2 Evaluation procedure steps

(a) multi-brick sensor block at position (8,8)

(b) single brick on 8 mm grid-design similar to [3, 17]

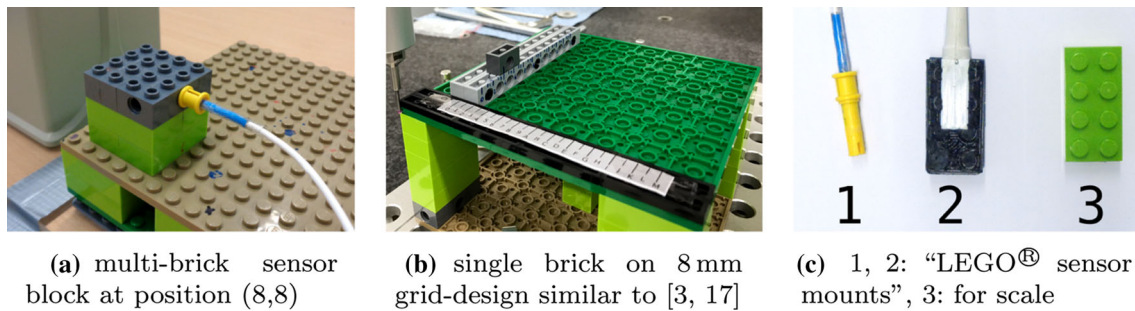**(c)** 1, 2: "LEGO®" sensor mounts", 3: for scale

Fig. 3 Prototype phantom designs and sensor fixation

position-dependent corrections. These corrections minimize systematic positioning errors on the board.

Positioning uncertainty estimation In the evaluation of measuring protocols, one always has to consider how much to trust the underlying reference standard. In this protocol, the chain of reference standards is:

EMTS measurement < phantom position < CMM measurement < CMM calibration phantom. Therefore, the uncertainty of the positioning depends on the determination of underlying reference standards and includes uncertainties from all steps.

We summarize relevant standard guidelines to express uncertainty in measurements [6,18]. Assuming processes such as measurement, geometry and placement to be Gaussian-distributed not only simplifies calculation and understanding, but also often represents good approximations, e.g., for a CMM [18]. For Gaussians, the combination of multiple processes with standard deviation σ_i is $\sigma = \sqrt{\sum \sigma_i^2}$. The 95% uncertainty u_{95} (inclusion margin for 95% of samples) is approximated by $2 \cdot \sigma$. The estimation of a mean value from multiple measurements incurs errors represented by an uncertainty contribution of $\sigma/\sqrt{N-1}$. While estimated on limited information, we derive the positioning uncertainty of the state-of-the-art protocols (Table 1) by the same rules. We follow the chain of reference standards from CMM calibration phantom to EMTS phantom and list contributing factors as well as derived uncertainties.

In addition to statistical testing, we provide theoretical justification for the approximation of distances by a Gaussian in Sect. 4.1. Since we measure 2d positions on the board in a common coordinate system instead of distances between positions, we obtain distances by variation of individual position measurements. This yields a mixture of shifted Rice distributions, which can be approximated by a Gaussian.

Electromagnetic measurements A software tool guides the user through electromagnetic tracking (EMT) measurements to increase efficient acquisition. It instructs the operator to place the sensor block at a specific position, collects 100 measurement samples and median filters position records

to eliminate jitter. Movement of the evaluation phantom is compensated by working in reference to the reference sensor position and orientation. To enable comprehensive evaluation and compensation, the full *dataset of collected position records* for all sensors is saved. Finally, the base board is rotated and placed into different regions of the EMTS's measurement volume and the acquisition repeated.

Evaluation We define the error to be the difference of the offset (Euclidean) between phantom and median EMTS position. The former is calibrated by CMM measurements for the proposed design to compensate for the LEGO® board's inaccuracies. We obtain distances for all possible position combinations of all position records collected in sequence.

For statistical evaluation, we group errors by the length of the corresponding offsets' uncalibrated distances on the base board. The absolute coordinates w.r.t. the field generator's origin are not required, because the *offset distance error* or *relative distance error* is independent from absolute positions.

Compensation

The distortion of the magnetic field is addressed by different methods [7], which typically involve a training or calibration procedure. Since the expected performance in learning algorithms depends on the quality of annotations to train, we implement a rudimentary compensation scheme to test the impact of higher-accuracy phantom data. The compensation function $g(x, \omega)$, a position- and sensor-dependent cubic correction polynomial, operates on the absolute position x. We minimize sensor-dependent model weights (ω, ω_r) w.r.t. the tracking error of the phantom protocol ($\mathcal{E}_{EMT}$, Eq. 1) with the primary sensor's relative position $f(x_i, x_{r,i}, \omega, \omega_r) = (x_i + g(x_i, \omega)) - (x_{r,i} + g(x_{r,i}, \omega_r))$.
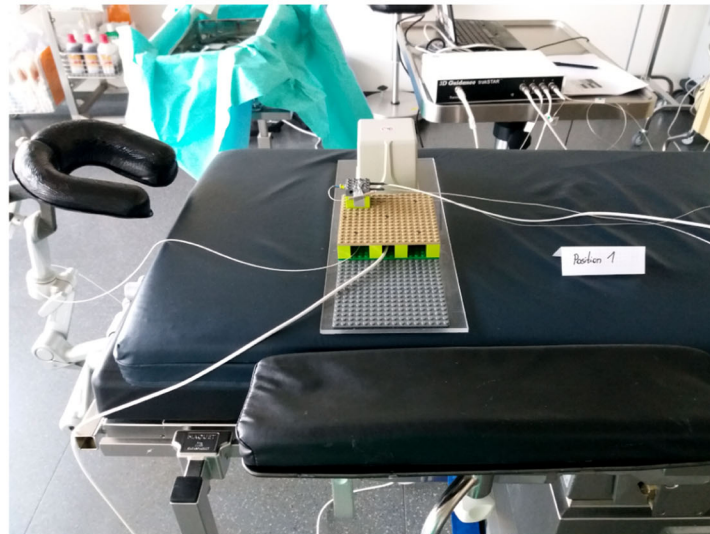
$$\mathcal{E}_{EMT} = \left\| f(x_2, x_{r,2}, \omega, \omega_r) - f(x_1, x_{r,1}, \omega, \omega_r) \right\|_2 - \|r_2 - r_1\|_2. \tag{1}$$

We train ω, ω_r using Levenberg–Marquardt honoring the phantom protocol error $\mathcal{E}_{EMT}$ and a α-weighed ℓ_2-regularization term in the loss:

Fig. 4 Clinical measurement environments

(a) x-ray source placed at short distance to setup (approx. 30 cm)

(b) x-ray source at minimum realistic distance to setup (approx. 15 cm)

(c) operating theater: operating table with head- and armrest

$$\mathcal{L} = \left| \mathcal{E}_{\text{EMT}}(x_{1/2}, x_{r,1/2}, \omega, \omega_r, r_{1/2}) \right|$$
$$+ \alpha \cdot (\ell_2(\omega) + \ell_2(\omega_r)). \tag{2}$$

The variables in Eq. (2) are: median measurements x_i of the primary sensor at position i, corresponding median position $x_{r,i}$ of the reference sensor and the phantom reference positions r_i. We ignore the rotation component of the reference sensor to simplify the loss function.

Fig. 5 Distributions of offset errors (uncalibrated vs. calibrated)

Experimental results and discussion

Since the uncertainties of the EMTS evaluation depend on the accuracy of the protocol, we start with an evaluation of the phantom-related uncertainty. To determine EMTS errors and uncertainties, we perform experiments in three environments: (1) a distortion-free laboratory environment, (2) in the vicinity of a c-arm and (3) in the operating theater (Fig. 4). Finally, we analyze how distortion compensation improves with the protocol's uncertainty. All evaluations focus on the position component of the tracked pose only.
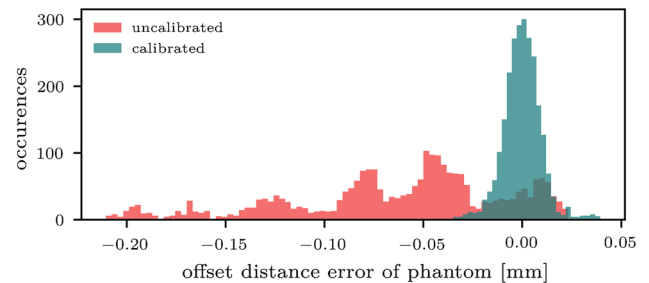
Phantom calibration and uncertainty estimation

We estimate the positioning uncertainty for two different phantom designs using (a) a single brick similar to state of the art; and (b) a block of multiple bricks. To discuss the effect of the calibration, we compare the distributions of offset distance errors of the multiple brick design in Fig. 5. The uncalibrated distribution emphasizes the multimodal nature of the incorrect positions and has significant errors of up to 0.2 mm, which is similar to the mean errors of EMTS. The
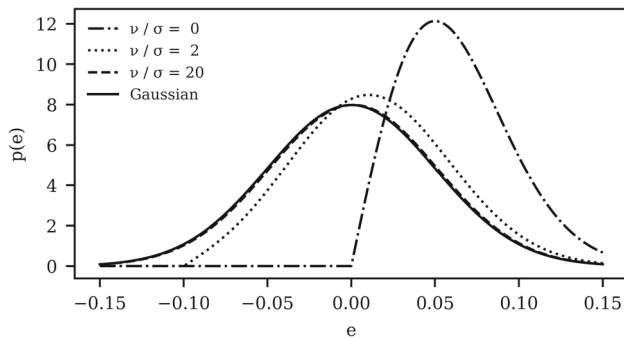
Fig. 6 Ideal mean-shifted Rice probability densities of distance error for $\nu \in \{0\,\text{mm}, 0.1\,\text{mm}, 1\,\text{mm}\}$ and $\sigma = 0.05$

calibrated scenario looks like a typical Gaussian distribution, which is easy to understand and work with, yet shows some outliers. With Gaussian-distributed 2d positions, distances among position pairs $((x_i, y_i), (x_j, y_j))$ follow the Rice distribution [14]: $R\left(\nu = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2},\ \sigma = \sqrt{2} \cdot \sqrt{\sigma_{LEGO}^2 + \sigma_{CMM}^2}\right)$. However, if the distance is significantly larger than the uncertainty of distance measurements, $R(\nu, \sigma)$ is well approximated by a Gaussian distribution [2]. With values for ν approximately between 8 and 120 mm (displacement distance), $\sigma \approx 0.01$ mm follows for the ratio $\frac{\nu}{\sigma} > 800$. Figure 6 illustrates that even for a generous upper bound $\sigma = 0.05$, the Gaussian approximation is valid for all distances on the LEGO® board.

The calibrated distribution passes the Kolmogorov–Smirnov test (*p* value = 0.0878) for a Gaussian distribution after removing 8 outlier CMM measurements (4%). We identify these outliers in a global analysis of component-wise CMM measurements. They might have been caused by inconsistencies in CMM measurements, e.g., inconsistent contact points and brick geometry.

The comparison of the evaluated phantom designs (Fig. 7) shows that even after calibration, the single-brick design exhibits root-mean-squared error (RMSE) of more than 0.15 mm. The contributing factors to the uncertainty of the calibration error (c.f. chain of reference standards, Sect. 3.2) are: (1) the coordinate measuring machine (see Sect. 4.4) and (2) the mean estimation error $\sigma/\sqrt{N-1}$. Together, these errors are indicated as black bars in Fig. 7. The bars are smaller for uncalibrated scenarios, because no mean estimation is performed, i.e., the improvement for single brick by calibration is not significant. However, the additional sta-

bility and rigidness given by building a larger sensor block (Fig. 3a) greatly reduce positioning uncertainty. Combining multi-brick with calibration reduces RMSE to approx. 0.01 mm ($u_{95,\text{calib}} \approx 2 \cdot \sqrt{x_{\text{RMSE}}^2 + \sigma_{\text{mean est.}}^2} \approx 0.021$ mm).

EMT evaluations

Our EMTS setup consists of a *3D Guidance* electromagnetic tracker by NDI (previously Ascension Technology Corporation), a mid-range field generator and several sensors, which we use for all experiments. The 3D Guidance system offers sensors in five different sizes ranging from 0.55 to 8 mm with model numbers indicating the size. We embed the largest and most accurate sensor (Model 800) rigidly into the phantom base using a 3D-printed fixture compatible to classic LEGO® designating it a reference sensor (see 2 in Fig. 3c). We tested several positions in the base and found no difference. Other sensors are used for the sensor block, as is typical for interventional scenarios, where the extra-corporal reference sensor can be much larger than sensors integrated into surgical instruments.

We use an acrylic board and a LEGO® stud board to fix the base w.r.t. the field generator. All samples are collected from within the optimal measurement volume as defined by the technical specifications.

For baseline measurements in a **laboratory environment**, we ensure minimal impact from magnetic field distortion by ensuring a significant distance of the evaluation setup to distortion sources, e.g., using a nonmetallic table.

In a **c-arm environment**, displacement measurements were performed in the vicinity of an X-ray device integrated into a cast steel gantry. The evaluation setup was placed in two different arrangements w.r.t. the c-arm (Fig. 4a, b). These positions produce the highest resolution X-rays. The scenario is realistic, if EMT is intended to be used at the same time as fluoroscopy.

In an **operating theater**, we place the evaluation setup on an electronically adjustable operating table with removable head- and armrests (Fig. 4c). This third environment provides two configurations: (1) operating table without head- and armrests and (2) with head- and armrests. For this environment, we focus on short-range offsets, reducing the acquisition procedure to six points.

The scatter plot (Fig. 8) shows the distribution of offset distance errors w.r.t. the offset distance in the **laboratory**

Fig. 7 Uncertainty of repeated phantom positioning: low root-mean-square errors (colored bars) imply consistent, repeatable sensor placement; black bars quantify uncertainty of this evaluation (CMM and mean estimation)
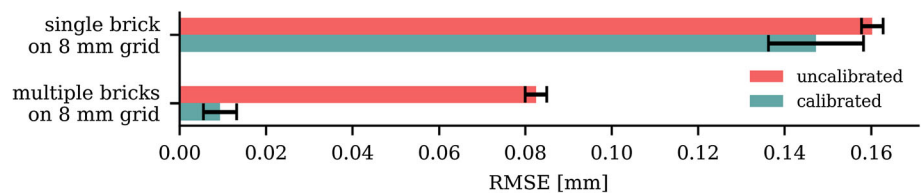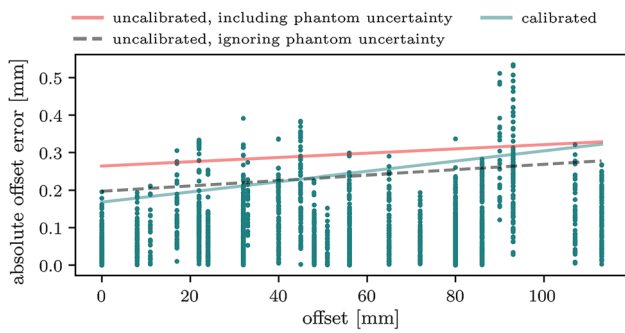
Fig. 8 Scatter plot of offset distance errors in the laboratory environment; line plots show a regression expressing 95% uncertainty estimation envelopes

environment. Uncertainties of individual measurements (calibrated phantom, dots) do not include the uncertainty derived from reference standards (e.g., phantom). However, regression lines include uncertainty from all sources for a calibrated (turquoise) and uncalibrated (red) multi-brick phantom. The dashed black line represents the believed EMTS uncertainty, if uncertainty contributions of reference

standards are ignored. The approximations are computed by combining the individually computed 95% confidence interval of the EMTS with the reference standard's u_{95}-uncertainties. This approximation represents an envelope, which covers 95% of measurements. The results confirm the linear relationship between offsets and errors [4] for small offsets.

Extending evaluations to typical **clinical environments** such as c-arm vicinity and operating theater, more sources of effective magnetic field distortion influence the tracking. As a consequence, average errors and uncertainty approximations develop a characteristic increase in the offset error as offset distances increase (Figs. 9 and 10). For the **operating theater**, we focus only on displacements of up to 45 mm and confirm observations.

In analogy to Wilson et al. [20], we derive and select the number of samples collected to exclude jitter effects, yielding that 100 samples are sufficient (Table 2).

Fig. 9 Offset errors in the c-arm environment; left: X-rays source close; right: at minimum distance to phantom and field generator
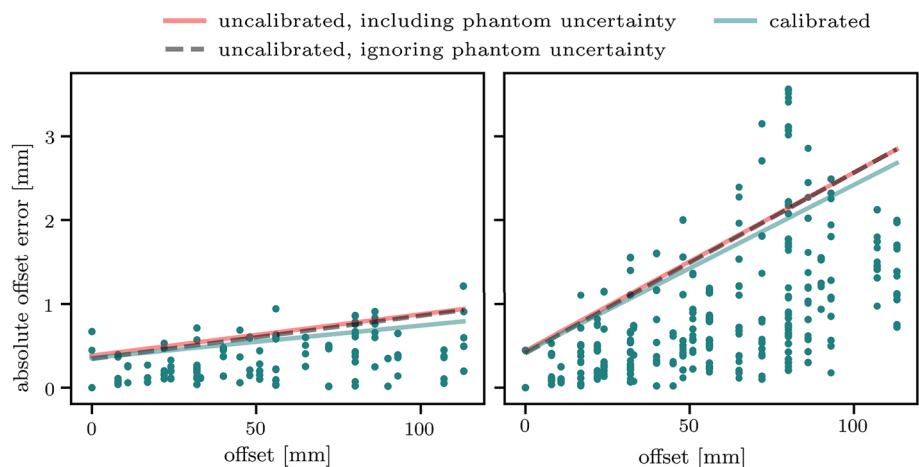

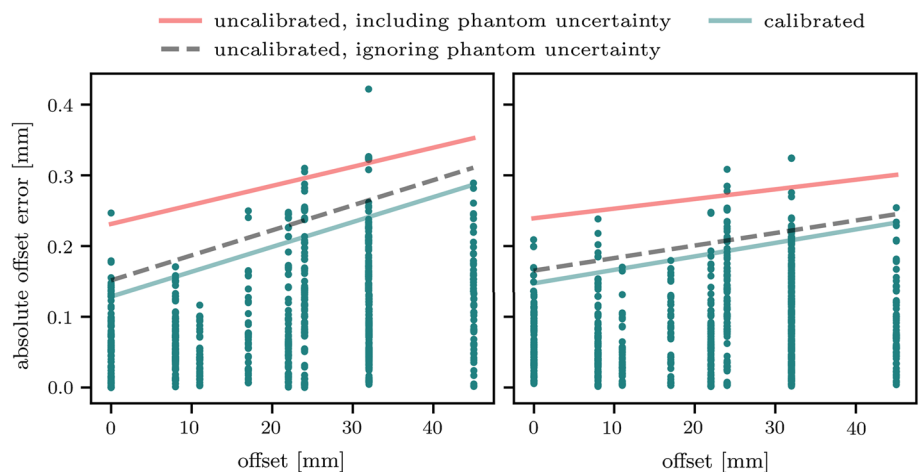**Fig. 10** Evaluations in operating theater environment; with (left) and without (right) arm- and headrest
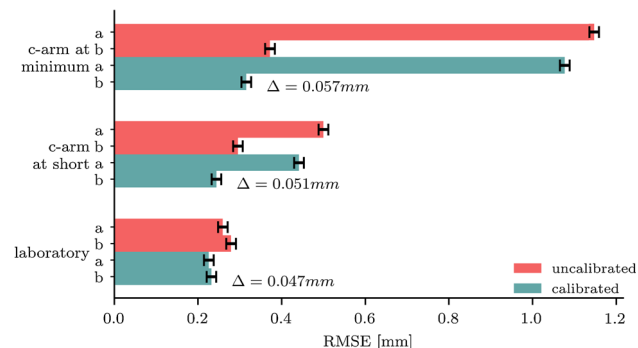
Table 2 Jitter of EMTS measurements (60 repetitions); min. STD ≈ 0

Samples	10	50	100	150	200	250	300
Avg. STD (mm)	0.0130	0.0257	0.0626	0.0609	0.0662	0.0639	0.0581
Max. STD (mm)	0.1855	0.1860	0.2078	0.2587	0.2634	0.2217	0.1799

Fig. 11 Impact of calibration on field distortion compensation; errors of uncompensated (**a**) and compensated (**b**) positions; Δ-values indicate improvements in calibrated over uncalibrated scenarios after compensation

Effect on error compensation

We employ a rudimentary machine learning method to correct systematic offsets of the tracking. The intended effect is to fix distortions of the magnetic field numerically in pose space. We acquire 12 position datasets at unique positions and orientations of the base board for each of the laboratory and the c-arm scenarios and perform an 8/4 training/test split. Independent trainings allow the comparison between calibrated and uncalibrated annotations (annotation quality). Data for testing on the other hand are always calibrated, as they are a higher-quality reference standard.

In Fig. 11, we compare the performance achieved by training on calibrated versus uncalibrated data, i.e., our protocol. The Δ-annotation quantifies the performance improvement. In light of the simplicity of the compensation scheme and the inherent EMTS accuracy (≈ 0.2 mm), improving errors by 0.05 mm is significant. For simple global regression, a bi-cubic polynomial and the regularization weight $\alpha = 0.2$ result in best performance (training time < 20 s on non-optimized single-core code). The small performance decrease in the distortionless laboratory environment is an expected side effect of overfitting on limited training data. However, errors of both c-arm experiments reach error levels comparable to laboratory error levels. The "minimum" scenario heavily impacted by magnetic distortions can benefit significantly from the compensation as evidenced by the reduction of approximately 70%.

Uncertainty of coordinate measuring machine

The contributors to the machine-dependent CMM uncertainty u_{CMM} are coordinate-dependent and derive from the CMM's specification (0.9 μm), the measuring tool (0.2 μm) and the CMM mean estimation error (1.4 μm). Finally, the underlying reference standard (CMM's calibration phantom, 1.0 μm) is taken into account. Combining these values for all axis yields, we determine that $u_{95,\text{CMM}} < 5$ μm is a safe overestimate.

Clinical relevance

Certainty of navigation has a significant impact on the surgical outcome of HP-MIS procedures. For instance in cochlear implantation, submillimeter drill navigation errors can lead to misplacement of implant electrodes [21] or permanent damage of risk structures, such as the facial nerve [10]. While in this work we show EMT currently does not meet the requirements for navigation in HP-MIS despite calibration and compensation, such protocols are required for improvements and clinical certification of tracking solutions for HP-MIS. The phantom's reduced inherent positioning uncertainty improves the analysis of errors and uncertainties of EMT measurements for submillimeter tracking. In the clinical scenario, absolute sensor positions can hardly be measured, as the sensor position w.r.t. the field generator's coordinate origin is unknown. The proposed EMTS assessment framework based on relative distances emulates the clinical scenario more realistically by eliminating the dependence on the tracker's coordinate origin. Significant performance increases for even a rudimentary learning-based compensation scheme illustrate the impact on clinical applications, where metal distortion is common.

Conclusion

We propose and evaluate a protocol for the assessment of EMTS performance in high-precision surgery. High-precision protocols not only improve the understanding of EMT errors, but are necessary to accurately quantify EMT performance for clinically viable HP-MIS. In the absence of LOS in HP-MIS, electromagnetic tracking has the potential to be the key technology for navigation. High-precision protocols, which include comprehensive uncertainty analysis,

should increase trust in the technology, fundamental toward being deployed in the operating room of the future. In the future, we plan to investigate error correction techniques based on advanced learning methods and larger training datasets. Hybrid tracking, the combination of EMT with more precise tracking techniques such as fluoroscopy [9], might retain the best of both worlds in such HP-MIS scenarios.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Research involving human participants and/or animals This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent This articles does not contain patient data.

References

1. Franz AM, Haidegger T, Birkfellner W, Cleary K, Peters TM, Maier-Hein L (2014) Electromagnetic tracking in medicine: a review of technology, validation, and applications. IEEE TMI 33(8):1702–1725
2. Gudbjartsson H, Patz S (1995) The rician distribution of noisy MRI data. Magn Reson Med 34(6):910–914
3. Haidegger T, Fenyvesi G, Sirokai B, Kelemen M, Nagy M, Takács B, Kovács L, Benyó B, Benyó Z (2011) Towards unified electromagnetic tracking system assessment-static errors. EMBC Proc 2011:1905–1908
4. Hummel J, Figl M, Birkfellner W, Bax MR, Shahidi R, Maurer CR, Bergmann H (2006) Evaluation of a new electromagnetic tracking system using a standardized assessment protocol. Phys Med Biol 51(10):N205–10
5. Hummel JB, Bax MR, Figl ML, Kang Y, Maurer C, Birkfellner WW, Bergmann H, Shahidi R (2005) Design and application of an assessment protocol for electromagnetic tracking systems. MedPhys 32(7):2371–2379
6. Joint Committee for Guides in Metrology (2008) Evaluation of measurement data—guide to the expression of uncertainty in measurement
7. Kindratenko VV (2000) A survey of electromagnetic position tracker calibration techniques. Virtual Real 5(3):169–182
8. Koivukangas T, Katisko JP, Koivukangas JP (2013) Technical accuracy of optical and the electromagnetic tracking systems. SpringerPlus 2(1):90–96
9. Kügler D, Stefanov A, Mukhopadhyay A (2018) i3posnet: i9nstrument pose estimation from X-ray. arxiv:802.09575
10. Labadie RF, Balachandran R, Mitchell J, Noble JH, Majdani O, Haynes D, Bennett M, Dawant BM, Fitzpatrick JM (2010) Clinical validation study of percutaneous cochlear access using patient customized micro-stereotactic frames. Otol Neurotol Off Publ Am Otol Soc Am Neurotol Soc Eur Acad Otol Neurotol 31(1):94
11. Lugez E, Sadjadi H, Pichora DR, Ellis RE, Akl SG, Fichtinger G (2015) Electromagnetic tracking in surgical and interventional environments: usability study. IJCARS 10(3):253–262
12. Maier-Hein L, Franz AM, Birkfellner W, Hummel J, Gergel I, Wegner I, Meinzer HP (2012) Standardized assessment of new electromagnetic field generators in an interventional radiology setting. MedPhys 39(6):3424–3434
13. Reichl T, Gardiazabal J, Navab N (2013) Electromagnetic servoing-a new tracking paradigm. IEEE TMI 32(8):1526–1535
14. Rice SO (1945) Mathematical analysis of random noise. In: Wax N (ed) Selected papers on noise and stochastic processes. Dover Publications Inc., New York, pp 133–294
15. Sadjadi H, Hashtrudi-Zaad K, Fichtinger G (2016) Simultaneous electromagnetic tracking and calibration for dynamic field distortion compensation. IEEE TBME 63(8):1771–1781
16. Schipper J, Aschendorff A, Arapakis I, Klenzner T, Teszler CB, Ridder GJ, Laszig R (2004) Navigation as a quality management tool in cochlear implant surgery. JLO 118(10):764–770
17. Sirokai B, Kiss M, Kovács L, Benyó B, Benyó Z, Haidegger T (2012) Best practice in electromagnetic tracking system assessment. In: II Proceedings of the joint workshop on new technologies for CRAS, vol 12, pp 1–4
18. Sładek JA (2016) Coordinate metrology: accuracy of systems and measurements, 1st edn. Springer, Berlin
19. Vakharia VN, Sparks R, O'Keeffe AG, Rodionov R, Miserocchi A, McEvoy A, Ourselin S, Duncan J (2017) Accuracy of intracranial electrode placement for stereoencephalography: a systematic review and meta-analysis. Epilepsia 58(6):921–932
20. Wilson E, Yaniv Z, Zhang H, Nafis C, Shen E, Shechter G, Wiles AD, Peters T, Lindisch D, Cleary K (2007) A hardware and software protocol for the evaluation of electromagnetic tracker accuracy in the clinical environment: a multi-center study. In: SPIE proceedings, SPIE, p 65092T
21. Wimmer W, Venail F, Williamson T, Akkari M, Gerber N, Weber S, Caversaccio M, Uziel A, Bell B (2014) Semiautomatic cochleostomy target and insertion trajectory planning for minimally invasive cochlear implantation. BioMed Res Int 2014:596498