33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Towards Unified Electromagnetic Tracking System Assessment— Static Errors

Tamás Haidegger, Gábor Fenyvesi, Beáta Sirokai, Márton Kelemen, Melinda Nagy, Barnabás Takács,
Levente Kovács, Balázs Benyó and Zoltán Benyó

Abstract— **Recent advances in Image-Guided Surgery allows physicians to incorporate up-to-date, high quality patient data in the surgical decision making, and sometimes to directly perform operations based on pre- or intra-operatively acquired patient images. Electromagnetic tracking is the fastest growing area within, where the position and orientation of tiny sensors can be determined with sub-millimeter accuracy in the field created by a generator. One of the major barriers to the wider spread of electromagnetic tracking solutions is their susceptibility to ferromagnetic materials and external electromagnetic sources. The research community has long been engaged with the topic to find engineering solutions to increase measurement reliability and accuracy. This article gives an overview of related experiments, and presents our recommendation towards a robust method to collect representative data about electromagnetic trackers.**

I. INTRODUCTION

Medical technology and surgical devices have shown an incredible development in the past three decades. One of the warmly welcomed trends is Minimally Invasive Surgery (MIS), which aims to reduce side effects and collateral damage to healthy tissue. However, MIS techniques require significant practice and superior hand–eye coordination, and in addition, the physicians cannot use their own senses. To compensate for these limitations, the concept of Image-Guided Surgery (IGS) and navigation were born, where surgeons use the patient's records for planning and execution. One of the key enabling technologies of IGS is intra-operative tracking, which made it possible to visualize the real-time tracking position and orientation information in the patient's 3D model.

II. TRACKING AND NAVIGATION IN THE MEDICAL ENVIRONMENT

High accuracy and fast sampling rate made optical tracking (OT) the most prevalent intra-operative IGS modalities, despite the limitation that it should be given full line-of-sight to the target objects [1]. This problem is eliminated with the use of Electromagnetic Tracking Systems (EMTS), where the transmitter generates a strong electromagnetic (EM) field and the sensor coils pick up a signal, since electric current is induced in them. Primary application of EMT has become MIS heart procedures. Calculation of the position and orientation of the sensor relative to a transmitter is based on the real-time electrical measurement and the physical model of the EM field. EMT is becoming widely popular, as it allows the tracking of an object inside of the body considering MIS procedures, where OT fails. However, ferromagnetic surgical instruments (laparoscopic forceps, metal trays, scalpels, etc.) distort the EM field, thus the tracking accuracy is significantly reduced. Due to recent advancement in system design, certain metallic materials do not cause significant error any more (e.g., titanium). However, there is still no general solution for metallic susceptibility of EMTS. In addition, errors jump in the case of fast motions, and in the presence of other electronic devices (e.g., endoscopes, monitors, light sources). To achieve similar precision to OT, it is essential to investigate the nature of errors of EMTS, to determine the typical distortions of the field. Proper system assessment allows for the development of effective calibration algorithms.

III. ELECTROMAGNETIC TRACKING SYSTEMS

The most commonly employed stand-alone medical grade EMTS (Fig. 1) are the Aurora from Northern Digital Inc. (NDI, Waterloo, ON), medSAFE from Ascension Technology Co. (Burlington, VT) and the FASTRAK from Polhemus (San Diego, CA). All of these systems have undergone significant changes in the past decade, the Aurora has seen four generations of transmitters (the latest is the Tabletop Field Generator), while Ascension's device is currently available with three different generators (including a Flat Transmitter).

Recently, EM capabilities have been integrated to most of the major surgical navigation systems (BrainLAB Kolibri, Medtronic AxiEM, GE InstaTrak, etc.) [2]. Despite the fact that every manufacturer admit the importance of objective evaluation of system performance, very limited data has been published on the experiments, and it is hard to confirm that data acquisition was performed in a repeatable way [3], [4].

IV. ERRORS IN EMT

Research groups have long been focusing on the issue of accuracy assessment of EMT [5], [6]. Although there are many solutions for determining the distortion of the EM field, no method has become generally accepted, or standardized.

The authors are with the Dept. of Control Engineering and Information Technology, Budapest University of Technology and Economics (BME), Magyar tudósok krt. 2., Budapest, Hungary {`haidegger, bbenyo, benyo`}`@iit.bme.hu, btakacs@digitalcustom.com`

The research was supported by the National Office for Research and Technology (NKTH), Hungarian National Scientific Research Foundation grant OTKA CK80316 and the Aktion Österreich–Ungarn foundation. It is connected to the scientific program of the " Development of quality-oriented and harmonized R+D+I strategy and functional model at BME" project, supported by the New Hungary Development Plan (Project ID: TÁMOP-4.2.1/B-09/1/KMR-2010-0002).

978-1-4244-4122-8/11/$26.00 ©2011 IEEE

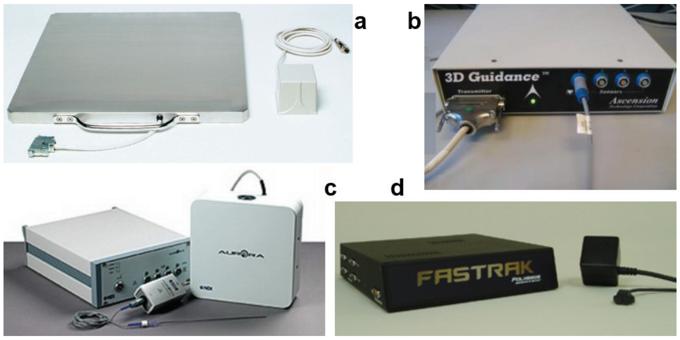

Fig. 1. The most commonly used EMT devices: a) the generators and b) the controller unit of the Ascension 3D Guidance medSAFE EMT, c) NDI Aurora, d) Polhemus Fastrak tracking system. (Courtesy of the manufacturers.)

The categorization of EMTS errors is not uniform in the literature, and reporting practices also vary. Position errors are typically described in terms of averaged *root mean square error* (RMSE) given as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \mathbf{x})^2}, \tag{1}$$

where N is the number of measurements, $\mathbf{x}$ is the desired point and $\mathbf{x}_i$ is the i^{th} measured point.

Our aim here is to describe a generic, clear and straightforward approach that can be understood and followed. For this reason, we choose to handle position and orientation errors separately. A valid approach is to divide errors into static and dynamic distortions [7]. Static errors are present as jitter and field distortions that can be characterized with RMSE or mean ± STD unambiguously. Confidence intervals and percentiles may hold further important information about error distribution.

A. Classification of erroneous tracking data

Various techniques exist to test position errors. For objectivity, a gold standard, such as an accuracy board may be used to collect data on the absolute measurement error of a system [8]. Some setups involve robotic devices, coordinate measuring machines, optical tracking systems [9], [10], or hybrid magneto–tracking systems [1].

It is crucial to test the effect of actual medical devices and the particular environment, including the operating table [3], [11], surgical tools [12] and other devices [13].

Previously published experiments can be grouped e.g., based on the level of automation, how fast and precise the data collection was, or simply the expenses of the procedure. Unfortunately, in most cases it is not clarified whether the experiment targeted the detection of static and dynamic errors, or position and orientation errors.

Table I shows various methods suitable to identify the different types of errors, and gives a summary of the main advantages and disadvantages of the above introduced methods. The collected information on the state-of-the-art in system assessment was indispensable to propose a unified method.

Further, dynamic errors depend on the speed of the sensor motion or the dynamic changes in the field, due to disturbing effects. Different experiments have been proposed to assess the dynamic error [7], [14]. While many groups assumed that the distribution of the errors is Gaussian, most commonly it is not true, as the acquired EM measurements are computed to position and orientation data through non-linear equations.

Modern tracking systems are typically not used in a standalone setup in research, but integrated into a more complex IGS system. In this case, various registration and calibration procedures are performed to link the coordinate systems of the separate parts. From the clinical point of view, the accuracy of treatment delivery is crucial, i.e., to meaningfully describe a system's *application accuracy*. It may be a highly non-linear function of the intrinsic and registration accuracies of the devices, therefore requiring special handling. Various error propagation techniques have been proposed in the literature to determine system errors as a function of the different integrated components [15].

V. RECOMMENDATIONS FOR STATIC EMTS ERROR ASSESSMENT

A. Method for static EMT measurement

Our approach for static measurements protocol is based on the standard manual collection, originally performed with a plexiglass calibration phantom, but we recommend the use of modular LEGO. The LEGO bricks are plastic, therefore they not distort the EM field, besides, as a single unit system, it is suitable to create a scalable data collection platform due to its extremely precise and repeatable fit of structures [22]. (It has an inherent fitting accuracy of 2 μm and development kits are widely available.) Our setup consists of a wooden table, a plastic palette, a rigid LEGO tower and a rigidly mounted sensor element (perpendicularly to the palette). This allows for the manual collection of static errors. The LEGO palette was divided into a 5×5 grid where its points are 88 mm away from each other.

To measure four different superior region of the allocated palette, we are using the LEGO tower. The distance between two stages is the height of two LEGO elements (19.2 mm). We chose the center of palette for origin, and collected 2500–3000 position and orientation data from each point in the same set-up. The measured data was evaluated statistically (computing the average along the 3 axes over the 25 points). This amount of data makes it able to smoothly reconstruct the original distributions. The STD of the positions were plotted in MATLAB and a higher-order surface was fitted for each level of results. This approach showed consistency for the measured configurations. The relatively low number of measurement points chosen enables faster completion, however, we verified the fitting of the error-surfaces with additional measurements in selected locations.

First, we used Ascension's 3D Guidance medSAFE system, which works with a DC field. More than 2,500 positions and orientations were recorded at various points in the workspace to investigate the distribution of the EM field. These measurements showed that the distribution at most of

TABLE I

ERROR CLASSIFICATION METHODS AND THEIR MOST IMPORTANT CHARACTERISTICS.

Method	**Static Error**	**Dynamic Error**	**Advantages**	**Disadvantages**
industrial robotic systems employed [7], [16]	X	X	high precision error evaluation	high cost, necessary calibration
LEGO robot [7], [9], [17], [18]	X	X	automate, low-cost	lower accuracy
redundancy based error detection [5]	X	X	simple, cheap, applicable to the OR	manual data collection
polycarbonate measurement plate [19]	X		very high acquisition accuracy	manual data collection
plexiglass [20]	X		simple yet effective, repeatable	manual data collection
plastic cube [21]	X		simple yet effective, repeatable	manual data collection
magneto-optical tracking systems [7], [1]	X	X	high precision error evaluation	need of an OTS, necessary calibration
manual movement of the dual-sensor [3]		X	good coverage of the workspace	only dynamic errors,unrepeatable

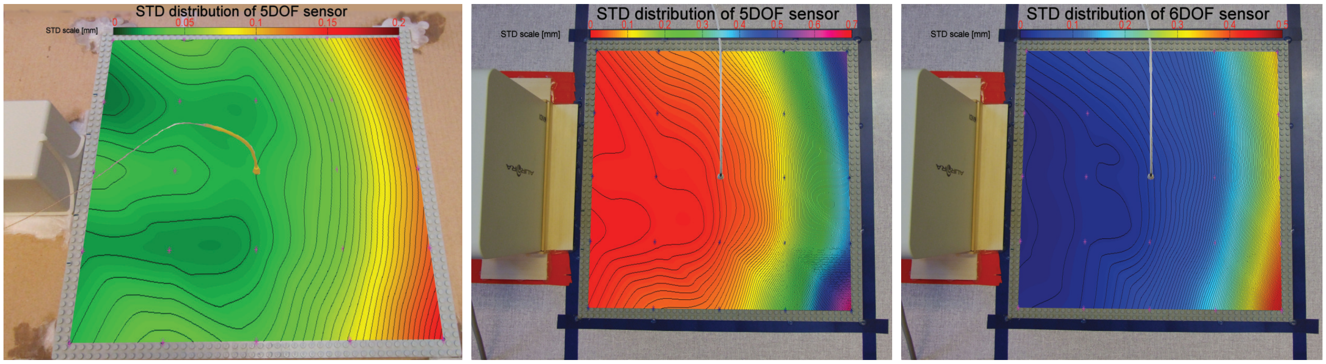

Fig. 2. Illustration of error measurements. a) Error STDs for the medSAFE 5 DOF sensor, interpolated over the whole workspace at the bottom level. b) Results with the Aurora 5 DOF sensor in the same arrangement. c) Aurora 6 DOF sensor results.

TABLE II

ERROR CLASSIFICATION METHODS AND THEIR MOST IMPORTANT CHARACTERISTICS.

	Level 0	Level 2	Level 4	Level 6	Level8
Aurora 5DOF	0.744±1.253	0.6505±1.1491	0.6153±1.2079	0.5607±1.4419	0.2747±1.3767
Aurora 6Dof	0.497±1.174	0.5432±1.2274	0.4854±1.1272	0.3996±1.0560	0.2937±1.2874
Ascension	0.383±1.495	0.3178±1.6875	0.6217±1.835	0.5113±1.9723	0.1694±1.4435

the points is close to Gaussian for 5 DOF sensors, however, it significantly changes close to the generator. Subsequently, we examined the static distribution of six points, uniformly spread apart in the workspace. Knowing the distance between points, average RMSE was found to be 0.6 mm (0.598 ± 0.047 mm), while orientation error was STD $0 \pm 0.95°$.

We found that in the origin—at ≈ 24 cm along the central axis from the base of the transmitter—the measurements had 0.046 mm (0 ± 0.021 mm) RSME for position and STD $0 \pm 0.43°$ for orientation. The 14 outer points on the grid, and an average 0 ± 0.027 mm position error and STD $0 \pm 0.351°$ orientation error was calculated (Fig. 2a). On the entire 25 points, the results were 0.57 ± 1.29 mm STD.

To facilitate the collection of large amount of data in a reproducible way, we designed and built a small rail car using a LEGO NXT set, which moves along a single axis. The vehicle driven by an outboard engine, actuated through a long string to minimize the distorting effects. The LEGO NXT robot programming is possible through the interface provided by the manufacturer. With the built-in PID controller, it is possible to position the car with a good repeatability (0.554 mm STD). The car moves along the axis and stops at preprogrammed positions to record a longer set of data. The rail needs to be repositioned afterwards along the two other axes. The first series of measurement results showed low distortion (0.03 mm STD) for the semi-automated data collection.

Next, the NDI Aurora was used with the same setup (5 DOF sensor), and we acquired measurements in all 125 measurement points over the workspace (Fig. 2b). Numeric results for all the 5 levels are presented in Table II. Average accuracy values derived to be 0.44 ± 1.18. The measurements were repeated with the 6 DOF sensor as well, where the results were 0.4 ± 1.7 (Fig. 2c).

B. External distortion measurements

Measurements were taken in the middle of the palette with three different laparoscopic tools in three orientations. We investigated how they affect the sensor, when put in its close proximity (Fig. 3). Without any disturbing objects, the RMSE was 0.05 mm (0.3 ± 0.024 mm) and STD $0.3 \pm 0.95°$ for position and orientation, respectively. First, introducing a disposable Covidien forceps the error (averaged for the three orientations) rose to 0.68 mm (0.65 ± 0.03 mm) and $1.1 \pm 0.86^\circ$. Next, with an artery clipper, the error derived to be 0.36 mm (0.35 ± 0.06 mm) and $0.9 \pm 1.06^\circ$, while with a da Vinci robot (Intuitive Surgical Inc.) Large Needle Driver, the error was 0.12 mm (0.1 ± 0.03 mm) and $0.12 \pm 0.65^\circ$. The results are preliminary, however, clearly show the serious impact of the tools. Tests will be carried out to observe the distorting effect of other medical tools, trays, scalpels, etc.

Fig. 3. Different laparoscopic tools used for distortion measurements.

VI. FUTURE WORK

In the next phase, measurements will be recorded with different orientations with the sparse coverage of the entire workspace of the EMT system. Further, we plan to develop a probe for dynamic error detection, and thorough investigations will be conducted to identify and develop a unified, robust approach for tool motions. In addition, we want to perform a fast and systematic environment assessment, to apply dynamic field distortion compensation. Finally, a standardizable mock operation setup should be developed, where application accuracy can be tested.

VII. CONCLUSION

Intra-operative electromagnetic tracking offers numerous advantages, however, its susceptibility to metallic distortions must be assessed and compensated. This paper reviewed the extensive literature on measurement protocols, based on what we started to develop a unified, repeatable and independent solution to evaluate different systems and setups. We tested our method with an Ascension medSAFE and an NDI Aurora tracker, and also performed basic measurements to understand distortions caused by common tools and devices in the medical environment. The ultimate solution should propose a practical way to determine inherent accuracies all over a device's workspace. Further, it should address how to conduct experiments with surgical tools and generic operating room equipment.

REFERENCES

[1] G. S. Fischer, "Electromagnetic Tracker Characterization and Optimal Tool Design—With Applications to ENT Surgery," MSc thesis in Mechanical Engineering, Johns Hopkins University, 2005.

[2] K. Schicho, M. Figl, M. Donat, W. Birkfellner, R. Seemann, A. Wagner, H. Bergmann, and R. Ewers, "Stability of miniature electromagnetic tracking systems," *Physics in Medicine and Biology*, vol. 50, pp. 2089–2098, 2005.

[3] C. Nafis, V. Jensen, and R. Von Jako, "Method for evaluating compatibility of commercial electromagnetic (EM) microsensor tracking systems with surgical and imaging tables," in *Proc. of SPIE Medical Imaging*, vol. 6918, 2008, pp. 1–15.

[4] E. Wilson, Z. Yaniv, D. Lindisch, and K. Cleary, "A buyer's guide to electromagnetic tracking systems for clinical applications," in *Proc. of SPIE Medical Imaging*, vol. 6918, 2008, pp. 1–12.

[5] D. D. Frantz, A. D. Wiles, S. E. Leis, and S. R. Kirsch, "Accuracy assessment protocols for electromagnetic tracking systems." *Physics in Medicine and Biology*, vol. 48, no. 14, pp. 2241–2251, 2003.

[6] J. Hummel, M. R. Bax, M. L. Figl, Y. Kang, J. Calvin Maurer, W. W. Birkfellner, H. Bergmann, and R. Shahidi, "Design and Application of an Assessment Protocol for Electromagnetic Tracking Systems," *Medical Physics*, vol. 32, no. 7, pp. 2371–2378, 2005.

[7] J. Much, "Error Classification and Propagation for EMT," MSc thesis in Computer Science, Technische Universitat Munchen, 2008.

[8] A. D. Milne, D. G. Chess, J. A. Johnson, and G. J. W. King, "Accuracy of an electromagnetic tracking device: A study of the optimal operating range and metal interference," *Journal of Biomechanics*, vol. 29, no. 6, pp. 791–793, 1996.

[9] X. Wu and R. H. Taylor, "A framework for calibration of electromagnetic surgical navigation system," in *Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2003, pp. 547–552.

[10] C. R. Mascott, "Comparison of magnetic tracking and optical tracking by simultaneous use of two independent frameless stereotactic systems." *Neurosurgery*, vol. 57, no. 4, pp. 295–301, 2005.

[11] S. R. Kirsch, C. Schilling, and G. Brunner, "Assessment of metallic distortions of an electromagnetic tracking system," *Proc. of SPIE Medical Imaging*, vol. 6141, pp. 143–151, 2006.

[12] F. Poulin and L.-P. Amiot, "Interference during the use of an electromagnetic tracking system under OR conditions," *Journal of Biomechanics*, vol. 35, no. 6, pp. 733–737, 2002.

[13] M. Hastenteufel, M. Vetter, H.-P. Meinzer, and I. Wolf, "Effect of 3D ultrasound probes on the accuracy of electromagnetic tracking systems." *Ultrasound in medicine & biology*, vol. 32, no. 9, pp. 1359–1368, 2006.

[14] K. McQuade, M. Finley, M. Harris-Love, and S. McCombe-Waller, "Dynamic Error Analysis of Ascension Flock of Birds$^{\text{TM}}$ Electromagnetic Tracking Device Using a Pendulum Model," *Journal of Applied Biomechanics*, vol. 18, no. 2, 2002.

[15] T. Haidegger, S. Győri, B. Benyó, and Z. Benyó, "Stochastic approach to error estimation for image-guided robotic systems," in *Proc. of the Annual Intl. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Buenos Aires, 2010, pp. 984–987.

[16] C. Nafis, V. Jensen, L. Beauregard, and P. Anderson, "Method for estimating dynamic EM tracking accuracy of surgical navigation tools," *Proc. of SPIE Medical Imaging*, vol. 6141, pp. 152–167, 2006.

[17] X. Wu and R. H. Taylor, "A Direction Space Interpolation Technique for Calibration of Electromagnetic Surgical Navigation Systems," in *Lecture Notes in Computer Science (LNCS), Annual Conf. of the Medical Image Computing and Computer Assisted Intervention Society (MICCAI)*, vol. 2879, 2003, pp. 215–222.

[18] M. Schneider and C. Stevens, "Development and testing of a new magnetic-tracking device for image guidance," *Proc. of SPIE Medical Imaging*, vol. 65090, pp. 1–11, 2007.

[19] J. Hummel, M. Figl, W. Birkfellner, M. R. Bax, R. Shahidi, J. C. R. Maurer, and H. Bergmann, "Evaluation of a new electromagnetic tracking system using a standardized assessment protocol," *Biomedical Engineering*, vol. 51, pp. 205–210, 2006.

[20] E. Wilson, "Accuracy Analysis of Electromagnetic Tracking within Medical Environments," Technical Report CAIMR TR-2006-2, 2006.

[21] E. Wilson, Z. Yaniv, H. Zhang, C. Nafis, E. Shen, G. Shechter, A. D. Wiles, T. Peters, D. Lindisch, and K. Cleary, "A hardware and software protocol for the evaluation of electromagnetic tracker accuracy in the clinical environment: a multi-center study," *Proc. of SPIE Medical Imaging*, vol. 6509, no. 1, 2007.

[22] F. G. Martin and R. Sargent, "Intelligent Machines Design Laboratory—LEGO Design," technical report.