Quantitative error analysis for computer assisted navigation: a feasibility study

Ö. Güler[*], **M. Perwög**,
4D Visualization Lab, Univ. ENT Clinic, Medical University Innsbruck, Anichstr. 35, 6020 Innsbruck, Austria

F. Kral,
4D Visualization Lab, Univ. ENT Clinic, Medical University Innsbruck, Anichstr. 35, 6020 Innsbruck, Austria

F. Schwarm[**], **Z. R. Bárdosi**,
4D Visualization Lab, Univ. ENT Clinic, Medical University Innsbruck, Anichstr. 35, 6020 Innsbruck, Austria

G. Göbel[***], and **W. Freysinger**
4D Visualization Lab, Univ. ENT Clinic, Medical University Innsbruck, Anichstr. 35, 6020 Innsbruck, Austria

[*]Childrens' National Medical Center, 111 Michigan Ave., N.W., Washington, D.C 20010, USA

[**]Department of Neurosurgery, Klinikum rechts der Isar, Technical University of Munich, Ismaninger Str. 22, 81675 Munich, Germany

[***]Department of Medical Statistics, Informatics and Health Economics (MSIG), Medical University Innsbruck, Anichstr. 35, 6020 Innsbruck, Austria

Abstract

Purpose—The benefit of computer-assisted navigation depends on the registration process, at which patient features are correlated to some preoperative imagery. The operator-induced uncertainty in localizing patient features – the User Localization Error (ULE) - is unknown and most likely dominating the application accuracy. This initial feasibility study aims at providing first data for ULE with a research navigation system.

Methods—Active optical navigation was done in CT-images of a plastic skull, an anatomic specimen (both with implanted fiducials) and a volunteer with anatomical landmarks exclusively. Each object was registered ten times with 3, 5, 7, and 9 registration points. Measurements were taken at 10 (anatomic specimen and volunteer) and 11 targets (plastic skull). The active NDI Polaris system was used under ideal working conditions (tracking accuracy 0.23 mm root mean square, RMS; probe tip calibration was 0.18 mm RMS. Variances of tracking along the principal directions were measured as 0.18 mm^2, 0.32 mm^2, and 0.42 mm^2. ULE was calculated from predicted application accuracy with isotropic and anisotropic models and from experimental variances, respectively.

Results—The ULE was determined from the variances as 0.45 mm (plastic skull), 0.60 mm (anatomic specimen), and 4.96 mm (volunteer). The predicted application accuracy did not yield consistent values for the ULE.

Conclusions—Quantitative data of application accuracy could be tested against prediction models with iso- and anisotropic noise models and revealed some discrepancies. This could potentially be due to the facts that navigation and one prediction model wrongly assume isotropic noise (tracking is anisotropic), while the anisotropic noise prediction model assumes an anisotropic registration strategy (registration is isotropic in typical navigation systems). The ULE data are presumably the first quantitative values for the precision of localizing anatomical landmarks and implanted fiducials. Submillimetric localization is possible for implanted screws; anatomic landmarks are not suitable for high-precision clinical navigation.

Keywords

application accuracy; navigation; human localization error; registration

Introduction

The acceptance of intraoperative guidance[1-4] rises and falls with its intraoperative added value. The registration step is crucial for this clinical tool and pair-point-matching[5-7] is used almost exclusively; recently it is being complemented by surface registration[8, 9] that seamlessly integrates into clinical routine due to the intraoperative ease of use. Knowledge of navigation application accuracy is the key for a reliable intraoperative use[10, 11], specifically for advanced applications like intraoperative augmented-reality guidance[12, 13] and robotic interventions[14-16]. Clinical data on application accuracy are inconsistent[17-21] and clinical validation studies of application accuracy beyond reporting registration RMS-values [17, 22-24] are scarce. Predictions of the application accuracy of Procrustes-type patient-to-image-registration[10, 25-27] are widely accepted. Fitzpatrick *et al.* have coined the terms Fiducial Localization Error, Fiducial Registration and Target Registration Error (**FLE, FRE** and **TRE**)[25]. **FLE** describes the error made when localizing loci for registration (i.e. fiducials) in image space <u>and</u> on the patient and can be treated as a single quantity[28-30]. **FRE** is the difference between registration fiducials mapped from tracker space into image space and the corresponding fiducials in the images; for normal-distributed data, $\mathcal{N}(0, \sigma)$, the norm of **FRE** is the RMS[31] as calculated by the navigation system. For all other loci, i.e. targets, the application error is determined via the **TRE** that can be predicted up to first order in a closed form[25]. We use boldface symbols for vectors and vectorial random variables throughout; <…. > denotes the statistic expectation value of a random variable.

From the literature[25, 28] it is known that for χ^2-distributed random variables $\langle \mathbf{FRE}^2 \rangle$

$$\langle \mathbf{FRE}^2 \rangle = K \varepsilon^2 \sigma^2 \quad (1)$$

K is the degree of freedom of the random variable (= 3 in our case), ε is the smallness parameter of the underlying the first-order perturbation theory, σ is the standard deviation of the $\mathcal{N}(0, \sigma)$ Normal-distribution of the random variable under consideration. Specifically, σ^2 is the variance of the difference between an actual point and the "ideal one", **FLE**^{2}, and relates to **FRE**^{2} as[25, 28]

$$\langle \mathbf{FRE}^2 \rangle = \left(1 - \frac{2}{N}\right) \langle \mathbf{FRE}^2 \rangle \quad (2)$$

N is the number of registration fiducials. <**TRE**^{2}(**r**)> at a position *r* can be calculated as

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

$<$**TRE**^{2}(**r**)$>$ is governed by the geometric configuration of fiducials - the sum term; the variance of the localization error $<$**FLE**^{2}$>$, and the position of the target, *r*, in the principal axes coordinate system of the fiducials. d_k is the distance of the target from principal axis *k*, and f_k is the RMS distance of the fiducials from principal axis *k*[25]. In the following $<$**TRE**^{2}$>$ is used for $<$**TRE**^{2}(**r**)$>$ for simplicity.

Recently Balachandran[32, 33] and Fitzpatrick[10] have introduced the Target Localization Error (**TLE**), the error associated with the physical localization of a locus on a patient (a marker, a screw, or an anatomical feature).This yields an overall estimate of the application error of clinical navigation in quadrature, the Total Target Error, **TTE**[10], of

$$\left\langle \mathbf{TTE}^2\right\rangle = \left\langle \mathbf{TRE}^2\right\rangle + \left\langle \mathbf{TLE}^2\right\rangle \quad (4)$$

Here $<$**TLE**^{2}$>$, $<$**FLE**^{2}$>$, and $<$**TRE**^{2}$>$ are assumed to be independent random variables. **TRE(r)** with anisotropic noise covariance[27] can be predicted as

$$\left(\Sigma_{\mathrm{TRE}}\left(r\right)\right)_{ij} \approx$$
$$\approx \varepsilon^2 \left\{\frac{\sigma_{ij}^2}{N} + \Sigma_{k\neq i}^{K}\Sigma_{m\neq j}^{K} \frac{r_k r_m\left(\Lambda_{kk}^2\delta_{km}\sigma_{ij}-\Lambda_{kk}^2\delta_{kj}\sigma_{im}-\Lambda_{ii}^2\delta_{im}\sigma_{kj}+\Lambda_{ii}^2\delta_{ij}\sigma_{km}\right)}{\left(\Lambda_{kk}^2+\Lambda_{ii}^2\right)\left(\Lambda_{mm}^2+\Lambda_{jj}^2\right)}\right\} \quad (5)$$

δ_{ij} is the Kronecker delta, ε the smallness parameter of the perturbation-theoretic approach, σ_{ij}^2 are the components of the **FLE** covariance matrix $\Sigma_{\mathbf{FLE}}$, r_i are the coordinates of a point **r** expressed in the principal axes coordinate system of the fiducials; Λ_{ii}^2 are the i-th singular values of the fiducial configuration in the principal axes. The RMS(TRE) can be measured and pre dicted[34] via

$$\left\langle \left(RMS_{\mathrm{TRE}}\left(\mathbf{r}\right)\right)^2\right\rangle =$$
$$= \varepsilon^2\left\{\frac{\sigma_{ij}^2}{N} + \sum_{i=1}^{K} + \Sigma_{j\neq i}^{K}\frac{r_j^2\left(\Lambda_{jj}^2\sigma_{ii}+\Lambda_{ii}^2\sigma_{jj}\right)}{\left(\Lambda_{ii}^2+\Lambda_{jj}^2\right)} ++ \sum_{i=1}^{K}\sum_{j\neq 1}^{K}\Sigma_{k\neq j}^{K}\frac{r_j r_k\Lambda_{ii}^2\sigma_{jk}}{\left(\Lambda_{ii}^2+\Lambda_{ii}^2\right)+\left(\Lambda_{kk}^2+\Lambda_{ii}^2\right)}\right\} = \quad (6)$$
$$= trace\left(\Sigma_{\mathrm{FLE}}\left(\mathbf{r}\right)\right)$$

This quantity is accessible via repetitions of measurements and includes the uncertainty of the tip of a navigated probe, p_r, $RMS_{\mathrm{TRE_{probe}}}$ (probe tip, p_r) relative to an uncertain Dynamic Reference Frame (DRF), and $RMS_{\mathrm{TRE_{DRF}}}$ (p_p)[34] in quadrature via

$$\left(RMS_{\mathrm{TRE_{DRF-probe}}}^2\left(p\right)\right)^2 = \left(RMS_{\mathrm{TRE_{DRF}}}^2\left(p_r\right)\right)^2 + \left(RMS_{\mathrm{TRE_{probe}}}^2\left(p_p\right)\right)^2 \quad (7)$$

None of the in vitro studies that we are aware of[15, 34-40] is providing data on the experimental User Localization Error (ULE) and none compares measurements with predictions of application accuracy[25, 34, 41-43]. The present experiments close this gap by mimicking a clinical setting and by exploiting all available information provided by an open-source navigation system. We performed measurements in our laboratory on a plastic skull, an anatomic specimen, and on a volunteer. As any clinical navigation system, the experimental navigation system employs optical tracking as if it had isotropic characteristics and performs registration and predicts application accuracy assuming zero-mean isotropic noise. This study is designed to include all factors affecting application accuracy and uses

the methodologies of Fitzpatrick *et al.*[25] and Wiles *et al.*[27] to estimate the rigid-body registration application accuracy. Our work might serve as a first feasibility study to quantify the human operator influence on application accuracy. Preliminary results were reported in[44, 45] where the TLE[10] was not considered and the experimental errors were not treated comprehensively. Large discrepancies between experimental and predicted application errors[25, 27] were thus found.

Approval of the local Ethics Committee was obtained for the measurements on the volunteer.

Materials and methods

Data generation and analysis

Navigation was done with open4Dnav[46], an IGSTK-based application with a state-machine architecture[47], full video capabilities[48], and optical tracking (active Polaris, first generation, NDI, Ontario, Canada). Standard patient-to-image registration with isotropic tracking[49] was used. Open4Dnav assumes isotropic tracking and imagery, as most clinical navigation systems do. open4Dnav logs all data for analysis with Matlab (version R2008b, Matlab Inc., Natick, MA, USA) and SPSS for Windows (release 15.0.1, SPSS Inc, Chicago, Ill., USA).

Imaging

CT data were acquired with a Siemens Sensation 16 CT (Siemens, Erlangen, Germany) and standard imaging protocols for paranasal sinuses / anterior skull base. Imaging parameters for plastic skull were: convolution kernel H60s, 120 kV, 74 mA, 1 mm slice thickness, and for the anatomic specimen: convolution kernel H30s, 120 kV, 175 mA, 0.6 mm slice thickness. The volunteer was scanned on a Siemens Somatom Plus 4 Volume Zoom, reconstruction filter H30s, 140 kV, 150 mA, 1.25 mm slice thickness.

Setup

The experimental setup is shown in Figure 1. Objects and patient DRF were rigidly mounted on an operating table. The DRF and the face of the navigated probe with the light emitting diodes (LEDs) were oriented perpendicular to the $-z$-axis of the optical tracker. All measurements were performed in the silo-type working volume specified by NDI. The positioning in this zone, c. f. Figure 2 for a schematic, was done with the help of a small helper application.

Optical tracker validation was done by measuring a calibrated length[50-52] at predefined locations in the measuring volume, see Figure 3. Two active optical rigid bodies (DRFs) were mounted on a metal bar, see Figure 4. Their distance, 62.75 mm, was calibrated with a micro calliper. This assembly was positioned with a hydraulic mechanical arm at 18 positions around and 5 positions in the center of the Polaris tracker working volume[53], see Figure 3, with the DRFs facing the tracker perpendicularly. 500 transformations of each DRF were sampled at each position from which the DRF-DRF distances were obtained as specified in NDI's accuracy assessment kit (http://www.ndigital.com/medical/documents/polaris/NDIAccuracyAssessmentKit.pdf) and Wiles *et al.*[53].

Tracking

Thermal effects of the Polaris tracker[54] were avoided by a three hours warm-up period prior to the measurements[54, 55]. Top and frontal views of the navigated probes with dimensions taken from our decommissioned ISG-Elekta Viewing Wand[56, 57] navigation system are shown in Figure 5. For calibration and verification the probe was firmly placed in a bore on the side of the DRF. These relative measurements do include the errors associated with

geometry[58], relative probe-to-DRF orientation[59], probe calibration, and tracking. Probe calibration (probe-DRF-measurements) was performed on the locations shown in Figure 3. The probe was calibrated with the IGSTK pivot-calibration routine on base of ten repetitions of calibrations with 1800 sampled transformations each. Orientational effects[34, 58, 60] of probe and DRF were minimized by manually aligning their active faces perpendicular to the $-z$-axis of the Polaris tracker throughout the experiments.

Tracker covariances

An assembly of DRF and DRF-probe, see above, was mounted on a hydraulic arm in the center of the working volume. Six-degrees-of-freedom (DOF) poses of the difference poses of DRF and DRF-probe were recorded from which the positional data sub-matrix of the covariance matrix was created. Measurements were done with a) the probe inserted into its socket on the DRF and b) with the tip located ± 15 cm off the DRF (but pointing to the DRF) to simulate the maximum distance one would expect during the measurements. The latter probe positions were used for predicting the application accuracy with anisotropic noise. Three covariance matrices were determined from a set of 1000 measurements each: a stationary DRF in the center of the working volume, Σ_{DRF}, the probe in the bore of the DRF, $\Sigma_{DRF-probe}$, and the tip of the navigated probe in the DRF ($\Sigma_{DRF-probe\_tip}$). Covariance matrices Σ_{object} for object= {plastic skull, anatomic specimen, or volunteer} for fiducials in image space, $\Sigma_{object}^{FLE_{image}}$, were obtained from the experimental data in image space; $\Sigma_{object}^{FLE_{pat}}$ was obtained for fiducials and targets on the patient in DRF space. $\Sigma_{object}^{FLE_{image}}$ and $\Sigma_{object}^{FLE_{pat}}$ were determined from registration and measurement data with 3, 5, 7, and 9 fiducials, 10 targets, with ten registrations, and four sets of fiducials, yielding a sample size of 880 for each object. 8 Covariances for M measurements / registrations (M= 10 in the experiments) for fiducials and targets were combined as

$$\Sigma_{combined} = \Sigma_{i=1}^{M} \frac{\Sigma_i}{M^2}. \quad (8)$$

Point definitions for registration and measurements

Experiments were done with 3, 5, 7, and 9 registration points, and with 11 (plastic skull) and 10 measurement points for anatomic specimen and volunteer, respectively. In a "preoperative" setting, all authors agreed on suitable registration points and optimized them in the sense of uniqueness and accessibility. This definition was saved. These registration fiducials were localized with a computer mouse ten times in image space, stored, averaged and used as the image-space fiducials for all registrations. All registration features (screws and anatomical landmarks) are shown as green dots in Figure 6. a – c; red dots in the figures show targets. Targets were defined in image space by the ENT specialist once; target image coordinates were stored as the reference for the measurements. All fiducials and targets of plastic skull and anatomic specimen were Ti screws; the volunteer had anatomic landmarks exclusively.

Registration and measurement procedures

One person placed the navigated probe on the fiducials, another person ran open4Dnav; a third one monitored the probe orientation relative to DRF and tracker. Probe positions relative to the DRF were stored when the "surgeon" decided that the probe was "optimally" placed on a fiducial or a target. The RMS, or **FRE**[25], of the registration was calculated by open4Dnav and was recorded for every registration.

For every registration j, the tracked probe was placed on registration loci i to obtain an experimental measure for $\mathbf{FRE}_{ji}$ from the vectorial difference of defined and actually displayed probe positions in image space. For each registration j targets k were measured to yield $\mathbf{TTE}_{jk}$. For each fiducial i and each target k the measurements of the ten registrations were averaged over the registrations j to approximate the statistical expectation value of $<\mathbf{FRE}_i>$ or $<\mathbf{TTE}_k>$ at fiducial i and target k by the mean. $<\mathbf{FRE}^2_i>$ and $<\mathbf{TTE}^2_k>$ at fiducial i and target k were obtained from the variances of $\mathbf{FRE}_{ji}$ and $\mathbf{TTE}_{jk}$, j= 1, … , 10.

The error model

$\mathbf{FLE}$ in eqns. (3) and (6) has to be replaced by the Total Fiducial Localization Error, $\mathbf{TFLE}$, which sums up all the errors pertaining to registration:

- errors in defining registration loci in the imagery, $\mathbf{FLE}_{image}$,

- probe calibration and tracker errors, $\mathbf{FLE}_{probe\_calib}$ and $\mathbf{FLE}_{tracker}$, respectively,

- user-induced errors when physically placing a probe on a patient feature: the User Localization Error, $\mathbf{ULE}$.

$\mathbf{TFLE}$ covers **all** sources of uncertainty in the present experimental setup and $<\mathbf{TFLE}^2>$ can be decomposed in quadrature[10] as

$$
\begin{aligned}
\left\langle \mathbf{TFLE}^2 \right\rangle &= \left\langle \mathbf{FLE}^2_{image} \right\rangle + \left\langle \mathbf{FLE}^2_{tracker} \right\rangle + \left\langle \mathbf{FLE}^2_{probe_{calib}} \right\rangle + \left\langle \mathbf{ULE}^2 \right\rangle =: \\
&=: \left\langle \mathbf{TLE}^2 \right\rangle + \left\langle \mathbf{FLE}^2_{tracker} \right\rangle + \left\langle \mathbf{FLE}^2_{probe_{calib}} \right\rangle =: \\
&=: \left\langle \mathbf{FLE}^2_{pat} \right\rangle + \left\langle \mathbf{FLE}^2_{image} \right\rangle
\end{aligned} \quad (9)
$$

$\left\langle \mathbf{TFLE}^2 \right\rangle$ is the sum of $\left\langle \mathbf{TFLE}^2 \right\rangle$ and the variances associated with tracking, $\left\langle \mathbf{FLE}^2_{tracker} \right\rangle$ and $\left\langle \mathbf{FLE}^2_{probe_{calib}} \right\rangle$, in line with the definition of TLE[10]. $\left\langle \mathbf{FLE}^2_{pat} \right\rangle$ is measured at patient fiducials / targets in DRF coordinates and contains **all** errors in patient space.

$\mathbf{TTE}$ is measured by placing the navigated probe on physical target loci (markers, screws, or anatomical landmarks) that were <u>not</u> used for registration purposes. $\mathbf{TTE}$ is the vectorial difference between "patient" targets transformed in image space and their predefined positions. This measures all errors inherent to the experiment. $<\mathbf{TLE}^2>$ can be obtained from eqns. (3) and (4):

$$
\left\langle \mathbf{TLE}^2 \right\rangle = \left\langle \mathbf{TTE}^2 \right\rangle - \left\langle \mathbf{TRE}^2 \right\rangle = \left\langle \mathbf{TTE}^2 \right\rangle - \frac{1}{N}\left(1 + \frac{1}{3}\sum_{k=1}^{N}\frac{d_k^2}{f_k^2}\right)\left\langle \mathbf{TFLE}^2 \right\rangle \quad (10)
$$

Inserting $<\mathbf{TFLE}^2>$ eq. (9), first definition, yields

$$
\left\langle \mathbf{TLE}^2 \right\rangle = \left\langle \mathbf{TTE}^2 \right\rangle - \frac{1}{N}\left(1 + \frac{1}{3}\Sigma_{k=1}^{N}\frac{d_k^2}{f_k^2}\right)\left(\left\langle \mathbf{FLE}^2_{image\_def} \right\rangle + \left\langle \mathbf{FLE}^2_{tracker} \right\rangle + \left\langle \mathbf{FLE}^2_{probe_{calib}} \right\rangle + \left\langle \mathbf{ULE}^2 \right\rangle\right) \quad (11)
$$

Solving for $<\mathbf{ULE}^2>$ and recalling eq. (9), first definition, yields

$$
\begin{aligned}
&\left\langle \mathbf{ULE}^2 \right\rangle \left\{1 + \frac{1}{N}\left(1 + \frac{1}{3}\Sigma_{k=1}^{N}\frac{d_k^2}{f_k^2}\right)\right\} + \\
&+ \frac{1}{N}\left(1 + \frac{1}{3}\Sigma_{k=1}^{N}\frac{d_k^2}{f_k^2}\right)\left\{\left\langle \mathbf{FLE}^2_{image} \right\rangle + \left\langle \mathbf{FLE}^2_{tracker} \right\rangle + \left\langle \mathbf{FLE}^2_{probe_{calib}} \right\rangle\right\} = \\
&= \left\langle \mathbf{TTE}^2 \right\rangle - \left\langle \mathbf{FRE}^2_{image} \right\rangle
\end{aligned} \quad (12)
$$

Simplifications and rearranging gives

$$\left\langle \mathbf{ULE}^2 \right\rangle \ = \frac{1}{1+\frac{1}{N}\left(1+\frac{1}{3}\Sigma_{k=1}^{N}\frac{d_k^2}{f_k^2}\right)}\left[\left\langle \mathbf{TTE}^2 \right\rangle - \right.$$
$$\left. -\frac{1}{N}\left(1+\frac{1}{3}\Sigma_{k=1}^{N}\frac{d_k^2}{f_k^2}\right)\left(\left\langle \mathbf{FLE}_{tracker}^2 \right\rangle + \left\langle \mathbf{FLE}_{probe_{calib}}^2 \right\rangle\right) - \left\langle \mathbf{FLE}_{\text{image\_def}}^2 \right\rangle\right] \quad (13)$$

The geometry factor for every target is calculated with the available data of the fiducials in image space and is constant for each registration and target. Eq. (9) gives, without any assumptions of prediction models,

$$\left\langle \mathbf{ULE}^2 \right\rangle = \left\langle \mathbf{TFLE}^2 \right\rangle - \left\langle \mathbf{FLE}_{image}^2 \right\rangle - \left\langle \mathbf{FLE}_{tracker}^2 \right\rangle - \left\langle \mathbf{FLE}_{\text{probe\_calib}}^2 \right\rangle \quad (14)$$

$\left\langle \mathbf{FLE}_{tracker}^2 \right\rangle, \left\langle \mathbf{FLE}_{\text{probe\_calib}}^2 \right\rangle >$ and $\left\langle \mathbf{FLE}_{image}^2 \right\rangle$ were measured. $<\mathbf{TFLE}^2>$ was determined from the unbiased estimate of the variance at target k, averaged over all registrations j and all targets k.

Σ_{TFLE}, the overall **TFLE** covariance matrix is required for the anisotropic **TRE** model[27] and is a composition of **FLE** covariance matrices in image and patient spaces. It is necessary to propagate $\Sigma_{tracker} + \Sigma_{probe\_calib} + \Sigma_{ULE}$ from DRF space into image space. In the special case of a navigation system the Jacobian of the rigid transformation turns out to be especially simple, the rotation from the patient-to-image registration[61], to yield

$$\Sigma_{TFLE} = \Sigma_{\text{FLE\_image}} + R * \left(\Sigma_{tracker} + \Sigma_{\text{probe\_calib}} + \Sigma_{ULE}\right) * R^T \quad (15)$$

Σ_{TFLE} implicitly contains the uncertainties of the tracked probe tip, eq. (7). As a consistency test, our implementation of eq. (6) with the covariance matrix Σ_{TFLE} of the generalized FLE, the Total Fiducial Localization Error, TFLE, was successfully validated against eq. (3) with the experimental values for eq. (7).

Results

Validation of the infrastructure

All measurements are given in millimeters as mean ± standard deviation (std), where applicable. Probe geometry verification with the Polaris tracker yielded 0.22 ± 0.03 mm (0.22 mm RMS) and 0.23 ± 0.04 mm (0.24 mm RMS) at the center and at the border of the working volume, respectively. The probe was pivot calibrated relative to its predefined mechanical dimensions in (and on the border of) the working volume as 0.18 ± 0.06 mm (0.19 mm RMS) and as 0.20 ± 0.12 mm (0.23 mm RMS), respectively. The experimental TFLE for plastic skull, anatomic specimen and volunteer was found as 0.74 mm, 0.95 mm, and 6.76 mm.

Normal distribution of measured coordinates of fiducials and targets was assessed with error probability of 0.05, see Table 1. Most of the data for the plastic skull and the anatomic specimen are normal distributed, deviations mainly occur in image z-axis. For the volunteer only one coordinate of one fiducial was not normal-distributed in image z-coordinate. Data for $\left\langle \mathbf{FLE}_{image}^2 \right\rangle, \left\langle \mathbf{FLE}_{tracker}^2 \right\rangle, \left\langle \mathbf{FLE}_{\text{probe\_calib}}^2 \right\rangle$, Table 4, and $<\mathbf{ULE}^2>$ were found to be normal distributed and independent with Matlab's Shapiro Wilk and Wilcoxon rank-sum tests, respectively. A significant correlation of $<\mathbf{TTE}^2>$ with $<\mathbf{TLE}^2>$ was found for all objects. Anatomic specimen, plastic skull, and volunteer had Pearson's ρ = 0.87, 0.2 9, and 0.738, respectively, see Figures 8 a-c.

Covariances

Σ_{DRF}, and $\Sigma_{DRF-probe\_tip}$ for eq. (6)[27] are given in Table 2. They are anisotropic along the main diagonal; the eigenvalues of Σ_{DRF} (x: $3.4*10^{-4}$, y: $1.62*10^{-4}$, z: $1.664*10^{-3}$) and of $\Sigma_{probe_{tip}-DRF}$ (x: $1.57*10^{-4}$, y: $4.624*10^{-3}$, z: $5.20*10^{-2}$) yield upper limits for the variances in x-, y- and z-directions. The DRF covariance has maximum variance in the tracker z–coordinate, well in line with literature, e. g. Ma *et al.*[41]. The tip of the navigated probe in the DRF frame has maximum uncertainty in DRF x-coordinates, c. f. Figure 7 a; the probe is almost aligned with the tracker z-coordinate axis. In the center of the working volume minor deviations between $\Sigma_{probe-DRF}$ and $\Sigma_{probe_{tip}>-DRF}$ occur.

$RMS\left(\Sigma_{DRF-probe\_tip}\right) = \sqrt{trace\left(\Sigma_{DRF-probe\_tip}\right)}$, the error for tracking a probe tip. In and at the border of the optimal working volume it was found as 0.01 mm and 0.24 mm, respectively. Off the optimum working center Σ_{DRF} changes slightly, but $\Sigma_{DRF-probe\_tip}$ changes significantly when approaching the border of the measurement volume, in line with the literature[27, 58].

<u>Covariances</u> $\Sigma_{object}^{FLE_{image}}$ and $\Sigma_{object}^{FLE_{pat}}$ were calculated from recorded coordinates of fiducial and target loci in image and tracker coordinates, respectively, see Table 3. The traces of diagonalized $\Sigma_{object}^{FLE_{image}}$ and $\Sigma_{object}^{FLE_{pat}}$ (object: plastic skull, anatomic specimen or volunteer) are upper limits for $<\mathbf{FLE}^2_{image}>$ and $<\mathbf{FLE}^2_{pat}>$. The ratios of eigenvalues in x-, y- and z-directions of $\Sigma_{plastic\_skull}^{FLE_{image}}$, $\Sigma_{anatomic\_specimen}^{FLE_{image}}$, and $\Sigma_{volunteer}^{FLE_{image}}$ are approximately 2:2:1, 1:1:3, and 1:1:3, respectively. $\Sigma_{plastic\ \ skull}^{FLE_{pat}}$ and $\Sigma_{volunteer}^{FLE_{image}}$ are almost aligned with the z-axis of the tracker, $\Sigma_{anatomic\_specimen}^{FLE_{pat}}$ has the major contribution along the DRF y-axis (for interpretation of axes see Figure 7). The ratios of eigenvalues of $\Sigma_{object}^{FLE_{pat}}$ in x-, y- and z-directions in image space are approximately 3:1:1, 5:5:1, and 2:1:4, respectively. $\mathbf{FLE}_{pat}$ covariances are given in Table 3 b as the mean over all targets and fiducials for each object;

$RMS\left(\mathbf{FLE}_{pat}\right) = \sqrt{trace\left(\Sigma_{object}^{TLE_{DRF}}\right)} =$ 0.61 mm, 0.79 mm, and 4.98 mm, for plastic skull, anatomic specimen and volunteer, respectively.

$<\mathbf{ULE}^2>$ from eq. (13) (values from eq. (14) in parentheses) for fiducials were 0.52^2 (0.45^2) mm^2 and 0.07^2 (0.60^2) mm^2. $<\mathbf{ULE}^2>$ for the volunteer was 1.57^2 (4.96^2) mm^2 with anatomical landmarks, c. f. Table 5.

Discussion

General remarks

Fiducials and targets are either screws or anatomical landmarks. This is different from a real clinical setting, but it allows detailed analyses of the application accuracy in computer-assisted navigation. We have used similar data sets – 1 mm for plastic skull, 1.25 mm for the volunteer – and 0.6 mm for the anatomical specimen. In view of the spatial resolution of the optical tracker (~0.3 mm) these data can safely be assumed to be equal. So, effects of image resolution and imaging are excluded.

Ten repeated registrations and measurements of $\mathbf{TFLE}^2_j$ and $\mathbf{TTE}^2_k$ via the experimental standard deviations at fiducials j and targets k (10 repetitions of measurements at 10 targets and 10 repetitions of registrations and measurements at 3, 5, 7 and 9 fiducials yield via

2[probe localizations during registration and measurement]*(3+5+7+9)[number of fiducials]*10[repetitions of registrations] + 4[sets of fiducials]*10[registration repetitions]*10[targets] = 480 + 400 data points for fiducials and targets of one object.

This is likely to be an adequate approximation to the statistical expectations for $<\mathbf{TFLE}^2>$ and $<\mathbf{TTE}^2>$ achievable in an experimental setting. Each target has sample size 40 which is deemed appropriate for an experiment, further supported by standard errors of mean ($< 2*10^{-3}$) and standard deviation ($< 3*10^{-3}$). Experimental sample covariances are unbiased estimators obtained by the 10 repetitions of registrations / measurements. More robust estimations of covariances would need a prohibitively large body of data and experimental time. The data are mostly normally distributed (see Table 1). As a generous first order approximation all data points were treated as normal distributed data in order to ease calculations and to allow the prediction of application accuracy.

Hardware aspects

Polaris accuracy measurements show a long-term stable operation with precision comparable to published[50, 62] and NDI data[53, 63]. The tracker was operated in thermal equilibrium[54]. We could not measure the trueness of the Polaris position measurements calibrated against a coordinate-measuring machine (CMM); these high-precision measuring tools are not available at our university.

Optical position measurement devices do have significant anisotropy, namely in the *z*-coordinate along the depth of view. Our data, c.f. Table 2 do show a pronounced anisotropy in the *z*-direction for the tracker and all derived quantities like the navigated probe tip, the DRF, and the measured coordinates. The covariance matrices are roughly diagonal, where the entries on the main diagonal are larger by at least one order of magnitude. This supports proper tool alignment in the experiments, see Figure 7.

ULE

ULE calculations are based on equations (13) and (14) and on the measured variances. One would expect a correct prediction of the **ULE** via the eq. (13). However, a notable difference between both approaches, c.f. Table 5, appears. The values for $<\mathbf{ULE}^2>$ from eq. (13) are smaller than those from eq. (14). Both approaches yield similar results for the plastic skull and the volunteer, but not for the anatomic specimen. Here, eq. (13) yields negative values for 28 of 40 points, resulting in an extremely low $<\mathbf{ULE}^2>$ of 0.07^2 mm^2 vs. 0.60^2 mm^2 from eq. (14). For the volunteer the results for $<\mathbf{ULE}^2>$ scale by almost a factor of ten! $<\mathbf{FLE}^2_{image\ object}>$, c. f. Table 4, ($0.26^2$ mm^2, 0.54^2 mm^2, 0.92^2 mm^2) are comparable to the sum of the eigenvalues of $\Sigma^{FLE_{image}}_{object}$ (for object: plastic skull, anatomic specimen, volunteer 0.11 mm^2, 0.43 mm^2, and 1.10 mm^2, respectively) which provide upper limits for the RMS. Within the limits of the experiments the results are in good agreement. **FLE**$_{pat}$ covariances, Table 3 b, as the mean over all targets and fiducials for each object (plastic skull, anatomic specimen and volunteer) yield an $\mathrm{RMS}\left(\mathbf{FLE}_{pat}\right) = \sqrt{trace\left(\Sigma^{TLE_{DRF}}_{object}\right)}$ of 0.61 mm, 0.79 mm, and 4.98 mm, respectively. This shows that implanted fiducial screws can be localized with submillimetric precision, whereas anatomical landmarks can be localized by a human operator in the millimetric range only. In all covariance matrices for localizing fiducials and targets the entries on the main diagonal are larger by at least one to two orders of magnitude than the other entries. In other words, the coordinate systems of probe and DRF are almost aligned with the tracker coordinate system.

Presumably all errors have been taken into account and thus **ULE** is the pure user error of physically placing the probe into the implanted fiducial screws and on anatomical landmarks. <u>Isotropic</u> noise models are known to <u>over-estimate</u> the x- and y-variances, while it underestimates z-variance[27]. However, in a real experimental setting these differences will go unnoticed: $<\mathbf{TFLE}^2>$, eq. (9), dominates the entries of the covariance matrices, Table 3, that are accessible via the RMS. Our data do not show this difference, see Figures 9 and 10. The predictions of $<\mathbf{TTE}^2>$ are based on the measurements of the $<\mathbf{TLE}^2> = <\mathbf{ULE}^2> + <\mathbf{FLE}^2_{image}>$; a large $<\mathbf{TLE}^2>$ – eventually due to unknown experimental errors – would result in large predicted $<\mathbf{TTE}^2>$ with either noise model. If the correct $<\mathbf{TLE}^2>$ was smaller than the one we find, the $<\mathbf{TTE}^2>$ predictions would decrease, ultimately leading to corresponding experimental results and predictions for $<\mathbf{TTE}^2>$. Note that **TLE** and **TRE**, eq. (13), are not statistically independent: the **ULE** appears in the **TRE** and the **TLE** terms. Thus this prediction is not valid.

TFLE contains all measuring experimental uncertainties: **ULE**, $\mathbf{FLE}_{tracker}$, $\mathbf{FLE}_{probe\_calib}$, and $\mathbf{FLE}_{image}$. **ULE** will be the relevant error source dominating eqns. (4) and (6). Both isotropic and anisotropic noise models for predicting the application accuracy of clinical navigation systems yield almost equal results. Concerning **ULE** prediction, neither noise model seems to be adequate. Both show significant correlations between the $<\mathbf{TLE}^2>$ and $<\mathbf{TRE}^2>$, c. f. Figure 8. Eq. (13) does by far underestimate the $<\mathbf{ULE}^2>$ as compared to eq. (14), which only assumes isotropic zero-mean Gaussian noise. It would be difficult to extract $<\mathbf{ULE}^2>$ from the non-linear eqns. (5) and (6) due to the sensitivity of the solution to experimental error and numerical instabilities. Eq. (14) yields plausible results for $<\mathbf{ULE}^2>$. Plastic skull and anatomic specimen have the same type of fiducial markers (screws), so similar **ULE**s should be likely: the tip of the probe almost exactly mates the notch of the screws. The **ULE**s are in a realistic range. In contrast to that, it can hardly be explained why eq. (13) does yield so different results for the experiments with implanted screws. The overall $<\mathbf{ULE}^2>$, calculated for all registrations, fiducials and targets, is almost zero for the anatomic specimen, which is due to the fact that the $\mathbf{ULE}^2_i$ was negative (!) in most cases (28 out of 40 registration experiments).

Aspects of predicting application accuracy

The discrepancies for **ULE** and **TTE** data suggested analyzing the experimental **TTE** with **TRE** prediction[25]; Figs. 11 and 12 show qualitatively good correspondences between experiments and predictions. This is further supported by a statistical analysis with a Wilcoxon test, for significance level 0.05, where the Null hypothesis ("the experimental **TTE** can be described by isotropic (anisotropic) **TTE** predictions") can **always** be rejected. Alternatively, when comparing pure isotropic (anisotropic) **TRE** predictions with the experimental **TTE**s, the same test statistic ("the experimental **TTE** can be described by isotropic (anisotropic) **TRE** predictions") can be rejected in only 25 % of the cases for the plastic skull and the volunteer for either noise model. This suggests that $<\mathbf{TRE}^2>$ prediction with isotropic / anisotropic noise might be adequate to model experimental navigation; this is further supported by a statistical analysis of whether $<\mathbf{TRE}^2>$ or $<\mathbf{TTE}^2>$ is a better description for the experiments: $<\mathbf{TRE}^2>$ is significantly better suited to predict our experiments (two-sided t-test, $\alpha = 0.05$) than the $<\mathbf{TTE}^2>$ in 75 % of the registration experiments with the plastic skull and the volunteer for isotropic noise model, Fig. 11. The anisotropic noise model was adequate for 3, 5, and 7 fiducials for the plastic skull and for 9 fiducials for the anatomic specimen, Fig. 12.

This is not fully convincing and might be attributed to the sample size of the experimental data. To test for this, the power of the study was determined. β, the error of the second kind, was small for the comparison of experimental data with isotropic TTE-predictions. In all

other cases the power of the study was low. A much larger sample size is needed to find decisive experimental evidence for one or the other predication model; for this end the data acquisition process should be redesigned. Reduction of data scatter and robust estimates of the covariance matrices need to be achieved while experimental data acquisition is still feasible.

Potential bias of data

By inspecting Figs. 9 and 10 one could argue that there is a significant bias in the data, notably for specific points. On the other hand, pure **TRE** predictions with the measured **TFLE** can adequately model the experiments with the plastic skull and the volunteer for a surprisingly large fraction. This may serve as an *a posteriori* confirmation of the approach presented and furthermore supports the absence of bias in the data and the correctness of the **TFLE.** Moreover, the present experiment is designed to measure variances of the random variables involved in the navigation / registration process for different objects to account for the influence of operator-induced error.

Clinical aspects

It might be interesting to study this problem where anisotropic data from tracking and imagery are registered anisotropically[64], with a more elaborate approach, e.g. a weighted Procrustes registration. In view of clinical reality of navigation, however, the good – because safe – approach for informing surgeons about the application accuracy is to provide upper limits, i.e. to inform the surgeons about the maximum error / deviation of navigation to be expected intraoperatively. This is the one and only approach satisfying patient safety. The large variability of the ULE suggests that anatomical landmarks be avoided when high-precision navigation with pair-point-registration is required clinically. The use of surface registration and of other iterative approaches like the Extented Kalman Filter[61] might avoid this problem. A potential benefit for clinical navigation might be that e.g. repeat measurements/localizations of the same anatomical landmarks could help reducing the **ULE** of these localizations on a patient. Moreover, all necessary parameters (**TFLE**) can be determined during the setup procedure of a clinical navigation system. This, however, pertains to rigid-body registration only. In summary, we quantified the fact that anatomical landmarks are not suitable for clinical high-precision navigation.

Conclusions

We have used implanted Ti screws and anatomical landmarks as fiducials and targets for experiments with a standard navigation system; all error sources like probe calibration, tracker error, image-space fiducial localization error and registration error have been taken into consideration. On base of our experimental data it was possible to determine the pure user-localization error, $<\mathbf{ULE}^2>$ with which features on a "patient", i.e. a plastic skull, anatomic specimen, and a volunteer, can be localized. The absolute values of ULE were found to be 0.61 mm, 0.79 mm, and 4.98 mm, for plastic skull, anatomic specimen, and volunteer, respectively. If anatomical landmarks are used for registration, **ULE** in patient and image spaces will be prominent; if implanted screws are used for registration, **ULE**, will be less dominant (most likely sub-millimetric) and can potentially be ignored; if a specific design of fiducials and registration probes[65] is used, **ULE** can safely be ignored.

To the best of our knowledge, these are first quantitative data for the operator-induced error in computer-assisted intraoperative navigation that employ a quantitative verification of application accuracy against predictions thereof.

To conclude, clinical navigation systems could use the **ULE** from registration data as a valuable indicator whether registration fiducials have been localized with a sufficient precision; thus it could complement intraoperative **TRE** prediction. Intraoperative **ULE** monitoring in combination with fiducial optimization[66] might prove useful for optimizing the application accuracy of fiducial based rigid body registration for clinical navigation.

Acknowledgments

References

1. Mösges R, Schlöndorff G. A new imaging method for intraoperative therapy control in skull-base surgery. Neurosurg. Rev. 1988; 11:245–247. [PubMed: 3251168]

2. Bucholz RD, Ho HW, Rubin JP. Variables affecting the accuracy of stereotactic localization using computerized tomography. J. Neurosurg. 1993; 79:667–673. [PubMed: 8410245]

3. Watanabe E, Watanabe T, Manaka S, Manayagi Y, Takakura K. Threedimensional digitizer (Neuronavigator): New equipment for computed-tomography guided stereotaxic surgery. Surg. Neurol. 1987; 27:543–547. [PubMed: 3554569]

4. Zinreich SJ, et al. Frameless stereotaxic integration of CT imaging data: accuracy and initial applications. Radiology. 1993; 188:735–742. [PubMed: 8351341]

5. Arun KS, Huang TS, Blostein SD. Least-Squares Fitting of Two 3-D Point Sets. IEEE Trans. Pattern Analysis and Machine Intelligence. 1987; 9:698–700.

6. Maurer, CRJ.; Fitzpatrick, JM. A review of medical image registration. In: Maciunas, RJ., editor. Interactive Image-Guided Neurosurgery. AANS; Park Ridge, Ill., USA: 1993. p. 17-44.

7. Maintz JBA, Viergever MA. A survey of medical image registration. Medical Image Analysis. 1998; 2:1–36. [PubMed: 10638851]

8. Besl PJ, Mckay ND. A Method for Registration of 3-D Shapes. IEEE Trans. Pattern Analysis and Machine Intelligence. 1992; 14:239–256.

9. Zhang Z. Iterative point matching for registrationof free-form curves and surfaces. Int. J. Comp. Vis. 1994; 12:119–152.

10. Fitzpatrick JM. The role of registration in accurate surgical guidance. Proceedings of the Institution of Mechanical Engineers Part H-Journal of Engineering in Medicine. 2010; 224:607–622.

11. Mascott CR, Sol JC, Bousquet P, Lagarrigue J, Lazorthes Y, Lauwers-Cances V. Quantification of true in vivo (application) accuracy in cranial image-guided surgery: influence of mode of patient registration. Neurosurgery. 2006; 59:ONS146–ONS156. [PubMed: 16888546]

12. Kockro RA, et al. Planning and simulation of neurosurgery in a virtual reality environment. Neurosurgery. 2000; 46:118–135. [PubMed: 10626943]

13. Dixon BJ, Daly MJ, Chan H, Vescan A, Witterick IJ, Irish JC. Augmented image guidance improves skull base navigation and reduces task workload in trainees: a preclinical trial. Laryngoscope. 2011; 121:2060–2064. [PubMed: 21898439]

14. Wilbert J, et al. Semi-robotic 6 degree of freedom positioning for intracranial high precision radiotherapy; first phantom and clinical results. Radiat. Oncol. 2010; 5:42. [PubMed: 20504338]

15. Labadie RF, Fenion M, Cevikalp H, Harris S, Galloway-RL J, Fitzpatrick JM. Image-guided otologic surgery. International Congress Series. 2003; 1256:627–632.

16. Coulson CJ, Reid AP, Proops DW, Brett PN. ENT challenges at the small scale. International Journal of Medical Robotics and Computer Assisted Surgery. 2007; 3:91–96. [PubMed: 17619240]

17. Shamir RR, Joskowicz R, Spektor S, Shoshan Y. Localization and registration accuracy in image-guided neurosurgery: a clinical study. Int. J. CARS. 2009; 4:45–52.

18. Berry J, O'Malley BW Jr. Humphries S, Staecker H. Making image guidance work: understanding control of accuracy. ANN. OTOL. RHINOL. LARYNGOL. 2003; 112:689–692. [PubMed: 12940666]

19. Nicolau S, Pennec X, Soler L, Ayache N. An accuracy certified augmented reality system for therapy guidance. LNCS. 2004; 3023:79–91.

20. Pillai P, Sammet S, Ammirati M. Application accuracy of computed tomography-based, image-guided navigation of temporal bone. Neurosurgery. 2008; 63:326–332. [PubMed: 18981839]

21. Woerdeman PA, Willems PW, Noordmans HJ, Tulleken CA, van der Sprenkel JW. Application accuracy in frameless image-guided neurosurgery: a comparison study of three patient-to-image registration methods. J. Neurosurg. 2007; 106:1012–1016. [PubMed: 17564173]

22. West J, Fitzpatrick JM. Point-based rigid registration: clinical validation of theory. Proc SPIE. 2000; 3979:353–359.

23. Fried MP, Kleefield J, Gopal H, Reardon E, Ho BT, Kuhn FA. Image-guided endoscopic surgery: results of accuracy and performance in a multicenter clinical study using an electromagnetic tracking system. Laryngoscope. 1997; 107:594–601. [PubMed: 9149159]

24. Sipos EP, Tebo SA, Zinreich SJ, Brem H. In vivo accuracy testing and clinical experience with the ISG Viewing Wand. Neurosurgery. 1996; 39:194–204. [PubMed: 8805161]

25. Fitzpatrick JM, West JB, Maurer CR Jr. Predicting error in rigid-body point-based registration. IEEE Trans. Med. Imaging. 1998; 17:694–702. [PubMed: 9874293]

26. Moghari MH, Ma B, Abolmaesumi P. A theoretical comparison of different target registration error estimators. LNCS. 2008; 5142:1032–1040.

27. Wiles AD, Likholyot A, Frantz DD, Peters TM. A statistical model for point-based target registration error with anisotropic fiducial localizer error. IEEE Trans. Med. Imaging. 2008; 27:378–390. [PubMed: 18334433]

28. Sibson R. Studies in the robustness of multidimensional scaling: Procrustes Statistic. J. R. Statist. Soc. B. 1978; 40:234–238.

29. Sibson R. Studies in the robustness of multidimensional scaling: pertubational analysis of classical scaling. J. R. Statist. Soc. B. 1979; 40:324–328.

30. Langron SP, Collins AJ. Perturbation-Theory for Generalized Procrustes Analysis. Journal of the Royal Statistical Society Series B-Methodological. 1985; 47:277–284.

31. HSU, DY. Spatial error analysis. IEEE Press; New York, USA: 1998.

32. Balachandran R, Fitzpatrick JM, Labadie RF. Accuracy of image-guided surgical systems at the lateral skull base as clinically assessed using bone-anchored hearing aid posts as surgical targets. Otol. Neurotol. 2008; 29:1050–1055. [PubMed: 18836389]

33. Balachandran R, Labadie RF, Fitzpatrick JM. Clinical determination of target registration error of an image-guided otologic surgical system using patients with bone-anchored hearing aids. Proc SPIE MI. 2007; 6509:650930–650936.

34. Wiles AD, Peters TM. Improved statistical TRE model when using a reference frame. LNCS. 2007; 4791:442–449.

35. Dorward NL, Alberti O, Palmer JD, Kitchen ND, Thomas DG. Accuracy of true frameless stereotaxy: in vivo measurement and laboratory phantom studies. Technical note. J. Neurosurg. 1999; 90:160–168. [PubMed: 10413173]

36. Schmerber S, Chassat F. Accuracy evaluation of a CAS system: laboratory protocol and results with 6D localizers, and clinical experiences in otorhinolaryngology. Comput Aided Surg. 2001; 6:1–13. [PubMed: 11335954]

37. Labadie RF, et al. In vitro assessment of image-guided otologic surgery: submillimeter accuracy within the region of the temporal bone. Otolaryngol. Head Neck Surg. 2005; 132:435–442. [PubMed: 15746858]

38. Vogele M, Freysinger W, Bale R, Gunkel AR, Thumfart WF. [Use of the ISG viewing wand on the temporal bone. A model study]. HNO. 1997; 45:74–80. [PubMed: 9173073]

39. Kral F, Riechelmann H, Freysinger W. Navigated Surgery at the Lateral Skull Base and Registration and Preoperative Imagery. Arch Otolaryngol. 2011; 137:144–150.

40. Vrionis FD, Foley KT, Robertson JH, Shea JJ. Use of cranial surface anatomic fiducials for interactive image-guided navigation in the temporal bone: a cadaveric study. Neurosurgery. 1997; 40:755–763. [PubMed: 9092849]

41. Ma B, Moghari MH, Ellis RE, Abolmaesumi P. Estimation of optimal fiducial target registration error in the presence of heteroscedastic noise. IEEE Trans Med. Imaging. 2010; 29:708–723. [PubMed: 20199909]

42. Moghari MH, Abolmaesumi P. Distribution of fiducial registration error in rigid-body point-based registration. IEEE Trans Med. Imaging. 2009; 28:1791–1801. [PubMed: 19884067]

43. Moghari MH, Abolmaesumi P. Distribution of target registration error for anisotropic and inhomogeneous fiducial localization error. IEEE Trans Med. Imaging. 2009; 28:799–813. [PubMed: 19423435]

44. F, Schwarm; Güler, Ö.; Kral, F.; Diakov, GM.; Reka, A.; Freysinger, W. Characterization of Open4Dnav, an IGSTK-based 3D-navigation system for FESS. Int. J. CARS. 2008; 3:S 248.

45. Bickel M, Güler Ö, Kral F, Freysinger W. Exploring the validity of predicted TRE in navigation. Proc SPIE MI. 2010; 7625:261–265.

46. Bickel M, Güler Ö, Kral F, Freysinger W. Evaluation of the application accuracy of 3D-Navigation through measurements and prediction. Proc IFBME. 2009; 25/VI:349–351.

47. Ibanez, L., et al. IGSTK - Image-Guided Surgery Toolkit. Washington, DC, USA: 2009. Release 4.2 ed. www.igstk.org

48. Güler Ö, Yaniv Z, Cleary K, Freysinger W. New Video Component for the Image Guided Surgery Toolkit IGSTK. Proc.IFBME, IFBME Proceedings. 2009; 25 / VI:359–390.

49. Horn BKP. Closed-form solution of absolute orientation using unit quaternions. J. Opt. Soc. Am. A. 1987; 4:629–642.

50. Chassat F, Lavallee S. Experimental Protocol of Accuracy Evaluation of 6-D Localizers for Computer-Integrated Surgery: Application to Four Optical Localizers. LNCS. 1998; 1496:277–284.

51. Atuegwu NC, Galloway RL. Volumetric characterization of the Aurora magnetic tracker system for image-guided transorbital endoscopic procedures. Phys. Med. Biol. 2008; 53:4355–4368. [PubMed: 18660560]

52. VDI-VDE and GMDA. VDI/VDE Handbook Measuring Technology II. Vol. 2634. VDI, Düsseldorf, Germany; Berlin, Germany: 2002. Optical 3D measuring systems. Imaging systems with point-by-point probing; p. 1-10.

53. Wiles AD, Thompson DG, Frantz DD. Accuracy assessment and interpretation for optical tracking systems. Proc SPIE. 2004; 5367:1–12.

54. Elfring R, de la FM, Radermacher K. Assessment of optical localizer accuracy for computer aided surgery systems. Comput. Aided Surg. 2010; 15:1–12. [PubMed: 20233129]

55. Güler, Ö.; Bardosi, ZR.; Ertugrul, M.; Di Franco, M.; Freysinger, W. Extending the tracking device support in the Image-Guided Surgery Toolkit (IGSTK): CamBar B2, EasyTrack 500, and Active Polaris. Kitware, Inc.; 2011. (version http://hdl.handle.net/10380/3288) [online]," Available: http://www.insight-journal.org/?journal=31

56. Gunkel AR, Vogele M, Martin A, Bale RJ, Thumfart WF, Freysinger W. Computer-Aided Surgery in the Petrous Bone. Laryngoscope. 1999; 109:1793–1799. [PubMed: 10569409]

57. Freysinger W, et al. Computer Assisted Interstitial Brachytherapy. LNCS. 1998; 1496:352–357.

58. West JB, Maurer CR Jr. Designing optically tracked instruments for image-guided surgery. IEEE Trans. Med. Imaging. 2004; 23:533–545. [PubMed: 15147007]

59. Ma B, Mogharm MH, Ellis RE, Abolmaesumi P. On fiducial target registration error in the presence of anisotropic noise. LNCS. 2007; 4792:628–635.

60. Simpson AL, Ma B, Ellis RE, Stewart AJ, Miga MJ. Uncertainty propagation and analysis of image-guided surgery. Proc SPIE. 2011; 7964:79640H-1–79640H-7.

61. Pennec X, Thirion J-P. A framework for uncertainty and validation of 3-D registration methods based on points and frames. Int. J. Comp. Vis. 1997; 25:203–229.

62. Khadem R, et al. Comparative tracking error analysis of five different optical tracking systems. Comput-Aided-Surg. 2000; 5:98–107. [PubMed: 10862132]

63. Frantz DD, Kirsch SR, Wiles AD. Specifying 3D tracking system accuracy. *Proc BVM*, Bildverarbeitung in der Medizin. 2004:234–238.

64. Danilchenko A, Fitzpatrick JM. General Approach to First-Order Error Prediction in Rigid Point Registration. IEEE Transactions on Medical Imaging. 2011; 30:679–693. [PubMed: 21075718]

65. Maurer CR Jr. Fitzpatrick JM, Wang MY, Galloway RL Jr. Maciunas RJ. Registration of Head Volume Images Using Implantable Fiducial Markers. IEEE Trans. Med. Imaging. 1997; 16:447–462. [PubMed: 9263002]

66. Shamir RR, Joskowicz L, Shoshan Y. Fiducial optimization for minimal target registration error in image-guided neurosurgery. IEEE Trans. Med. Imaging. 2012; 31:725–737. [PubMed: 22156977]

67. Bale RJ, Martin A, Vogele M, Freysinger W, Springer P, Giacomuzzi SM. The VBH Mouthpiece - A Registration Device for Frameless Stereotaxic Surgery. Radiology. 1997; 205:1107.
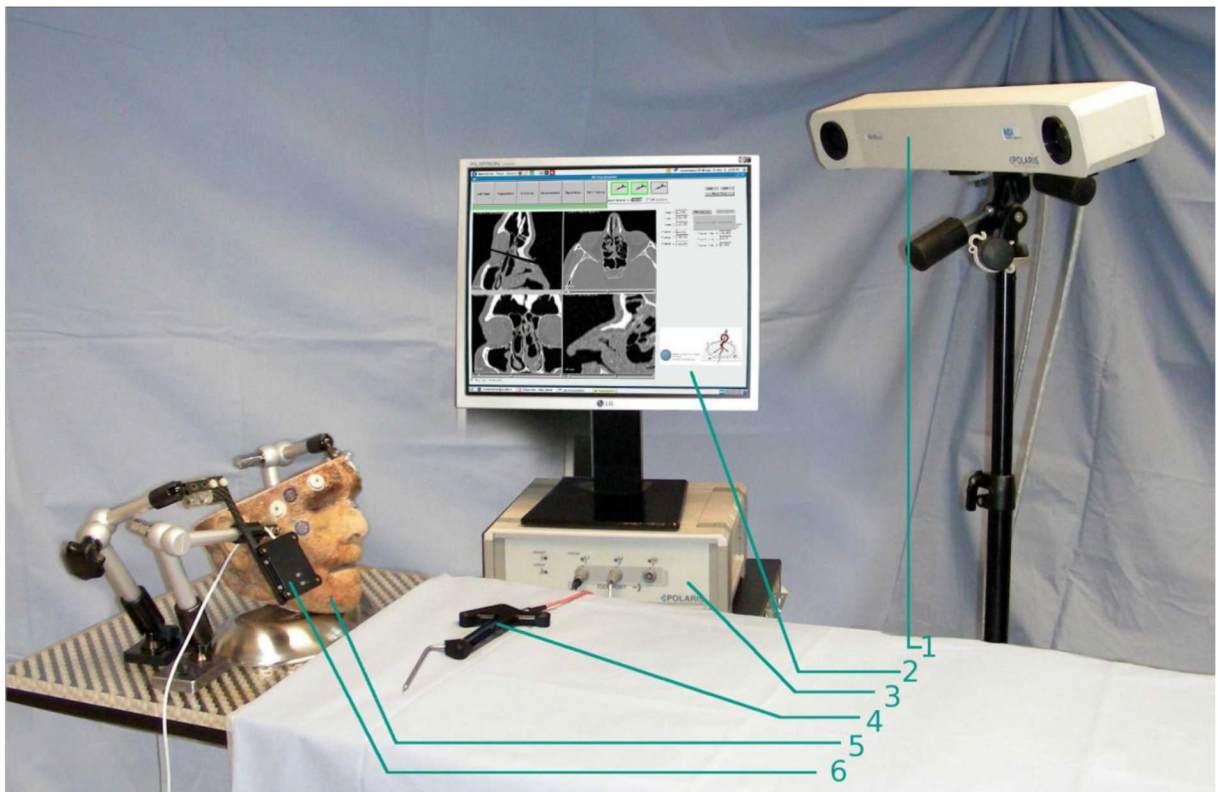
Figure 1.
Setup used for the experiments. The objects rest on a standard operating table (Brumaba, Germany) on a wood-plexiglass combination to hold hydraulic immobilization arms. The volunteer was resting in a comfortable position directly on the operating table to which he was immobilized with a tape running across the forehead. The active NDI Polaris camera (1) is placed in the optimal working distance from the object. The navigation system's monitor (2) and tracker control unit (3) are placed opposite to the surgeon. The probe used for all experiments (4) is lying on the table. In the example shown, the anatomic specimen (5) is held by two hydraulic arms and the patient tracker (a NDI rigid body, 6), is held separately. Thus a rigid mechanical setup could be achieved.
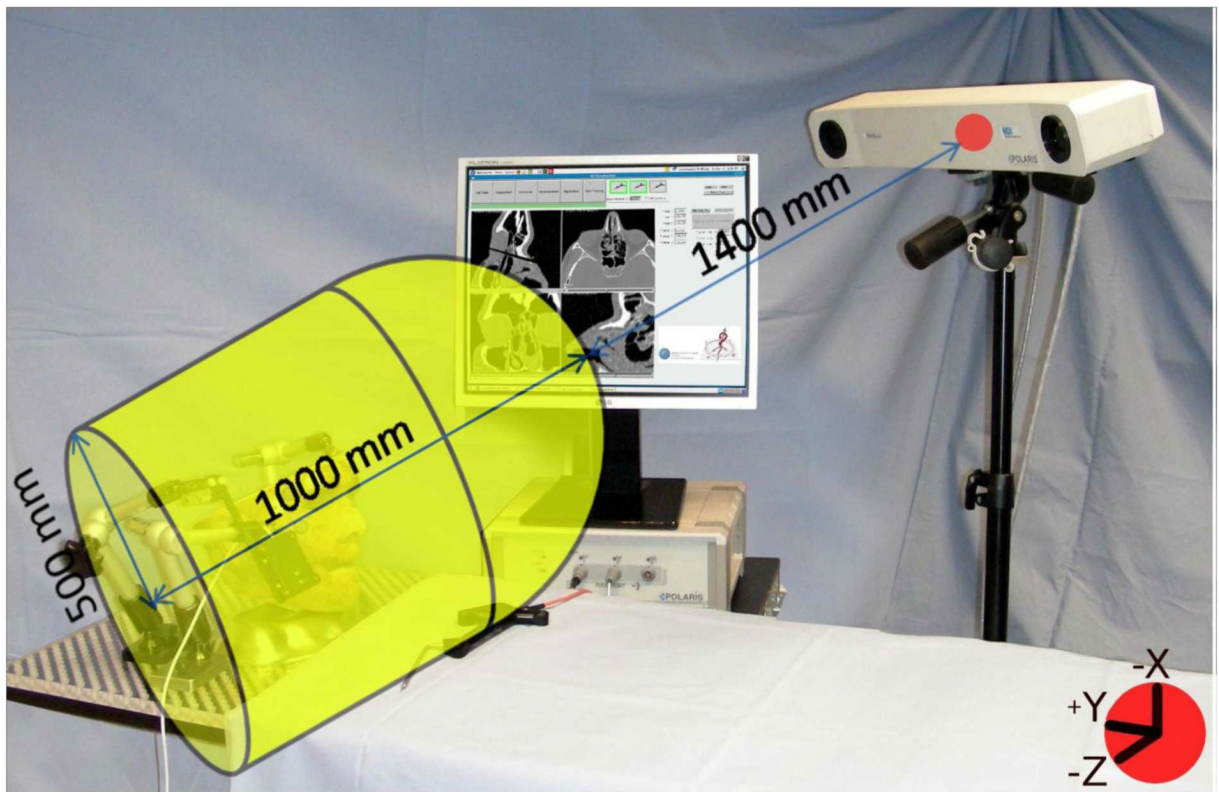
Figure 2.
Setup with overlay of optimal working zone. The active Polaris tracker was placed 1400 mm
away from the zone of maximum precision, a silo-type volume made up by a cylinder of
1000 mm height and diameter, covered by a semi-sphere of radius of 500 mm. All numbers
in the figures given in millimeters. Object placement within the ideal measurement zone was
verified with a custom application used for centering patient tracker and tracked probe as
seen by the tracker within this volume specified by the manufacturer. The red dot on the
camera marks the origin of the camera coordinate system. Down to the right the coordinate
axis are shown
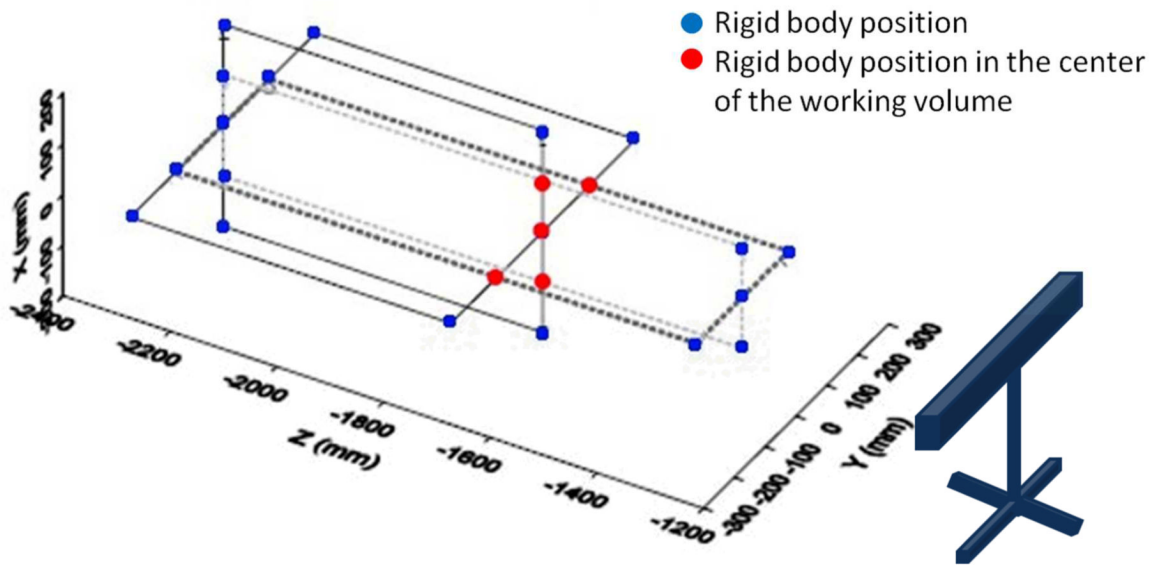
Figure 3.
Schematic drawing of the positions selected for the DRF-DRF tracker calibration within the measurement volume. The blue dots were selected on the outer border; red dots show positions at the border of the optimal measurement volume.
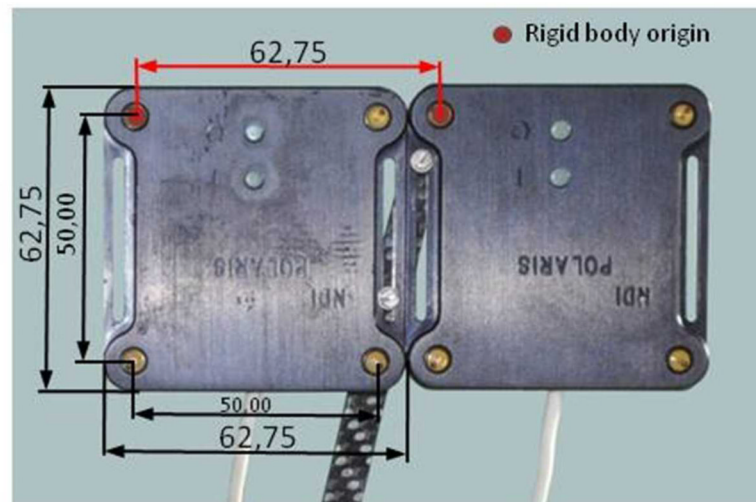
Figure 4.
Photograph of the DRF-DRF assembly used for tracker calibration. The dark structure protruding downwards is a carbon holder. The origins of the DRF coordinate systems are marked with red dots; distances were obtained with a micro-calliper and all dimensions are given in millimeters.
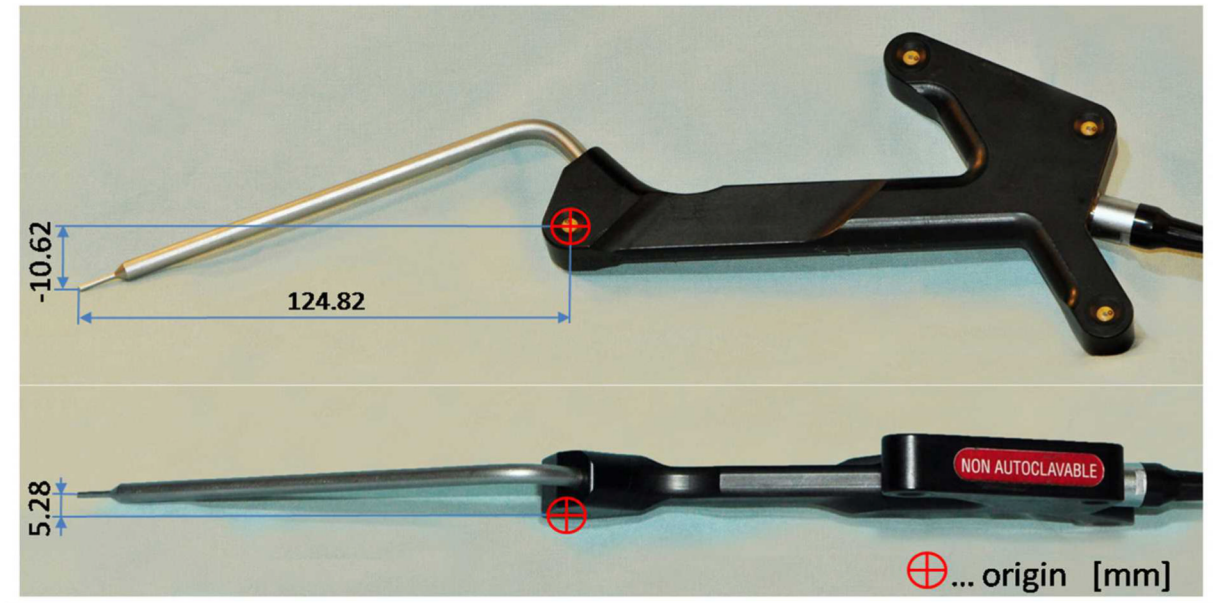
Figure 5.
Active tracked probe used for the experiments, front and top view with dimensions (in millimeters). The origin is located on the most distal LED (the crossed circle).

A plastic skull
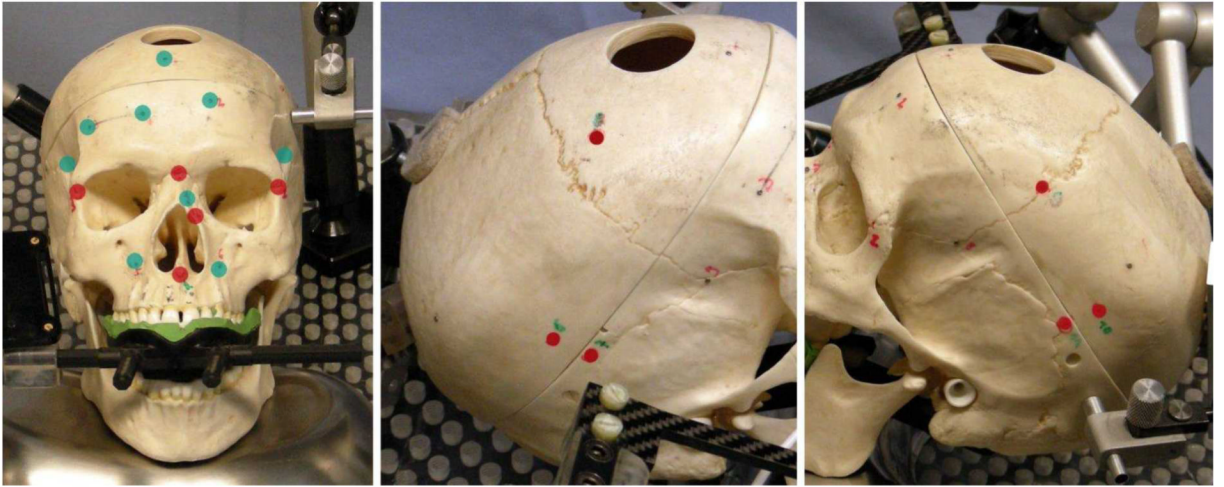
B: anatomic specimen
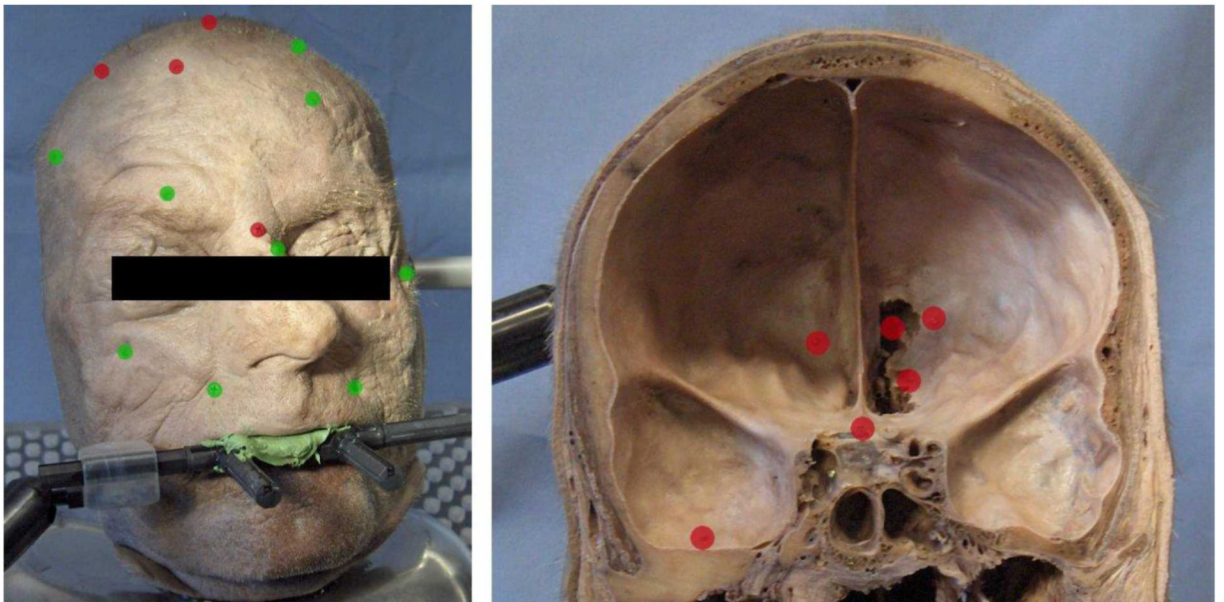
C. Volunteer
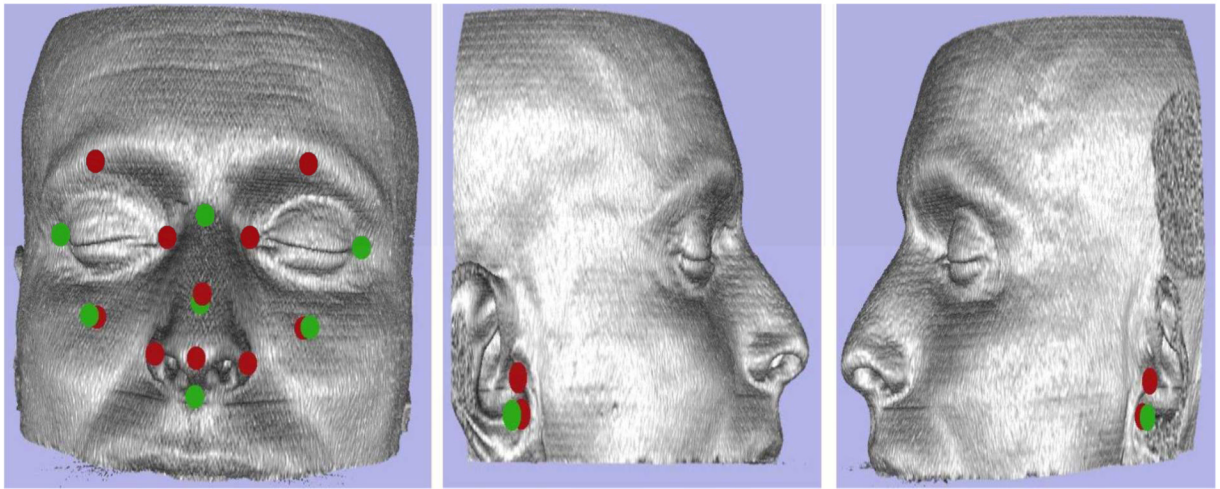
Figure 6.
A: Plastic skull with landmarks used for registration (green) and targets (red) on which the system accuracy was tested. The skull is placed on a base plate to hold the mechanical immobilization on base of the VBH headholder's mouthpiece[67]. B: Anatomic specimen with landmarks used for registration (green) and targets (red) on which the system accuracy was tested. The specimen was cut to allow accessing various anatomical structures. C: Volunteer (3D model) with landmarks used for registration (green) and targets (red) on which the system accuracy was tested. Surface reconstruction of the CT-data, thresholded to skin.
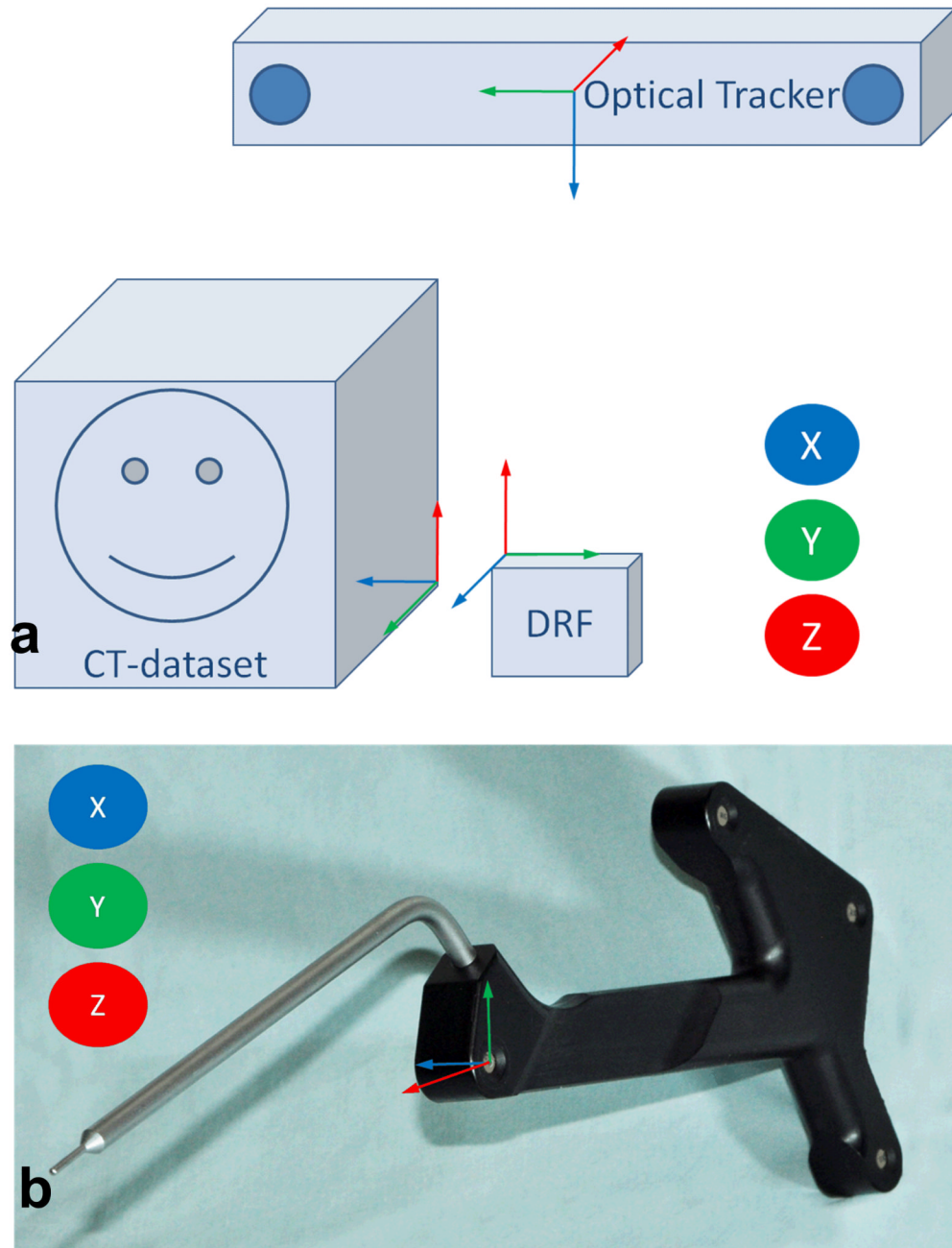
Figure 7.
a) Definitions of the coordinate systems for the experiments: image, DRF and tracker.
b) The probe coordinate system associated with the navigated probe.
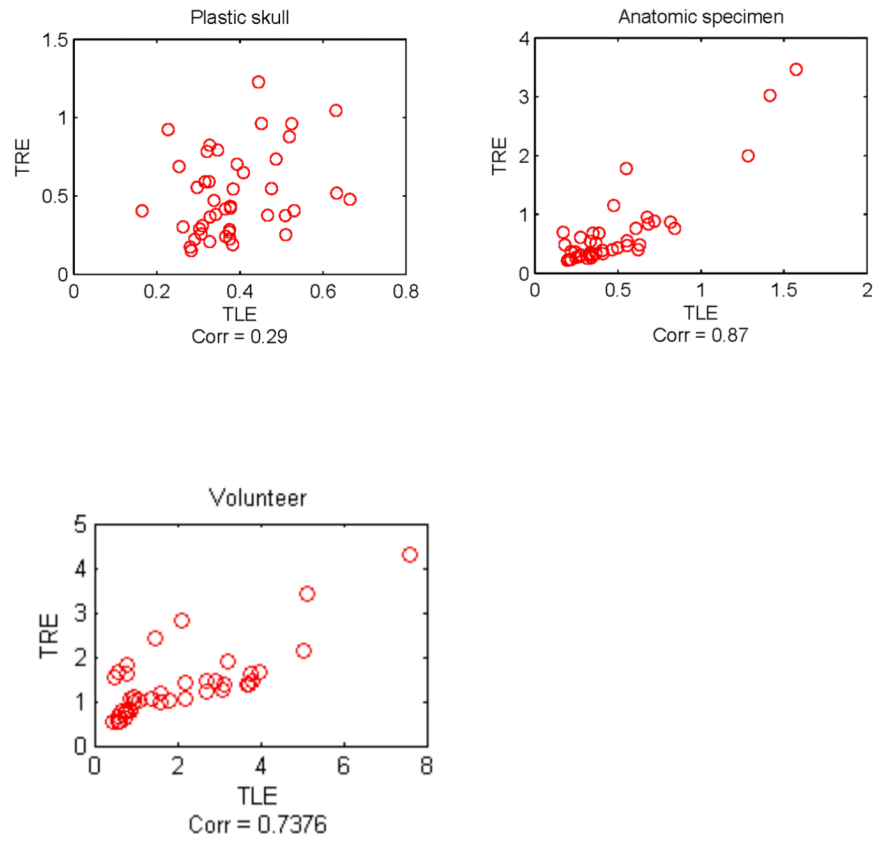
Figure 8.
Correlation plots for plastic skull, anatomic specimen and volunteer for TLE with TRE from the experimental data.
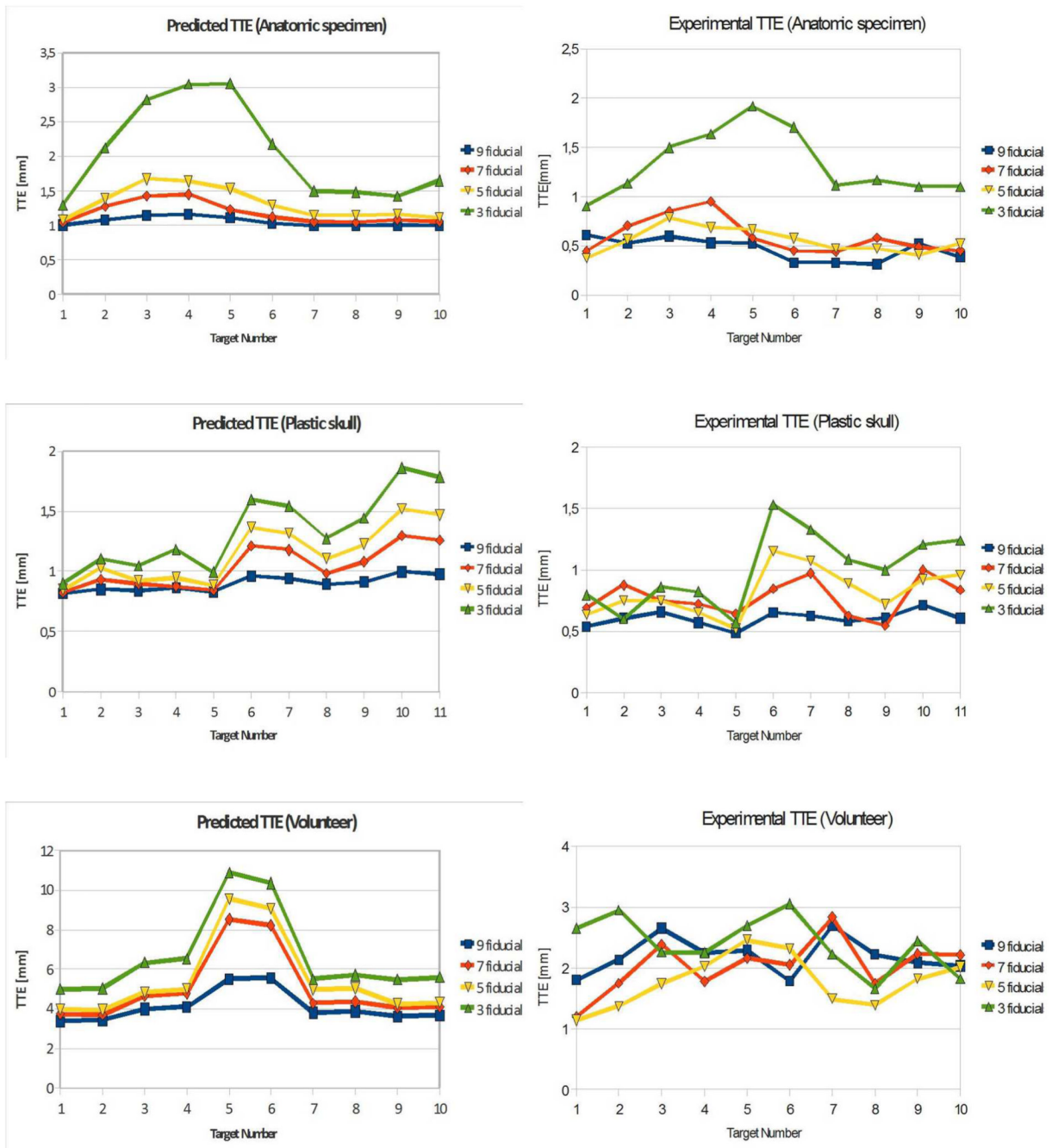
Figure 9.
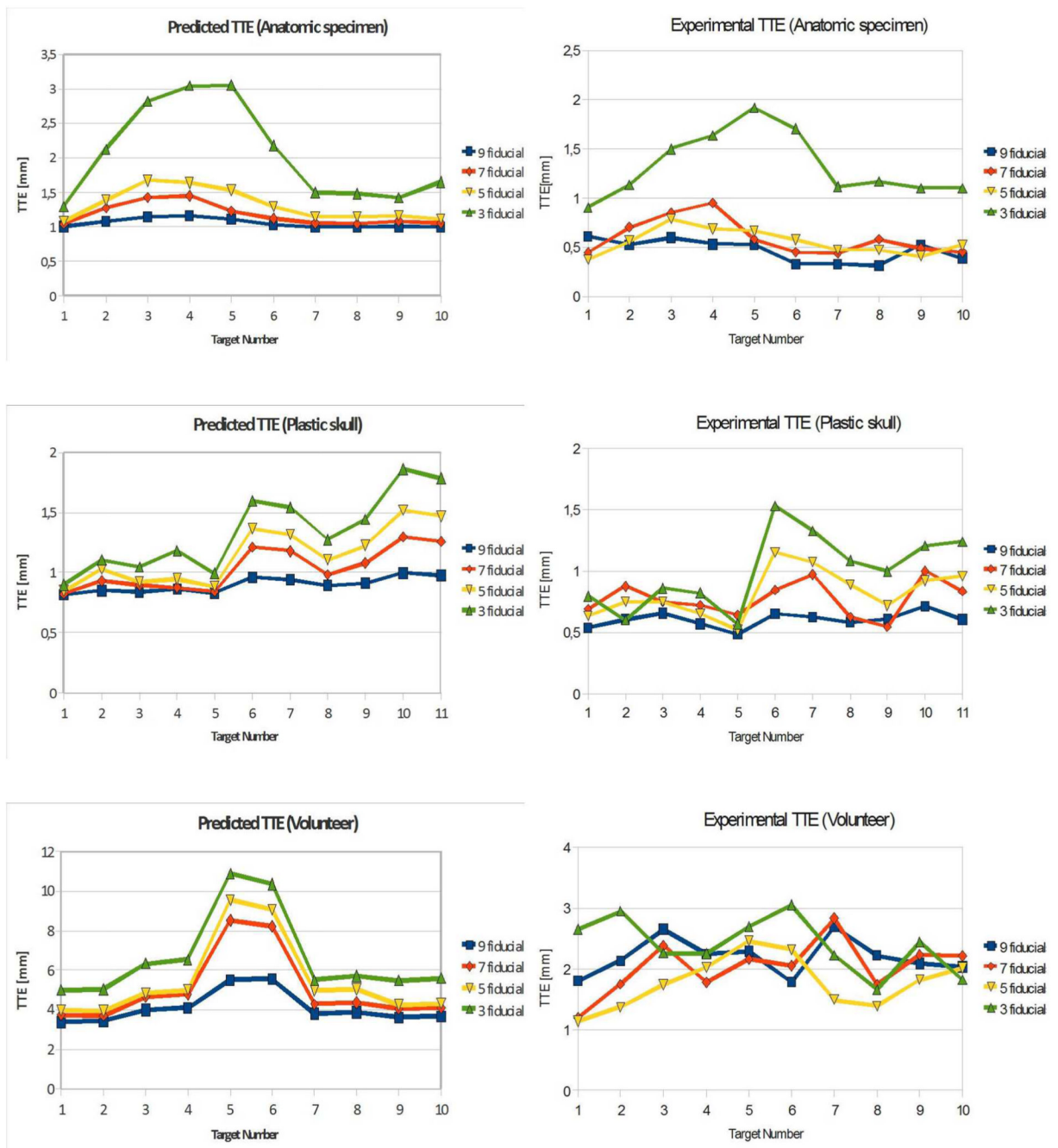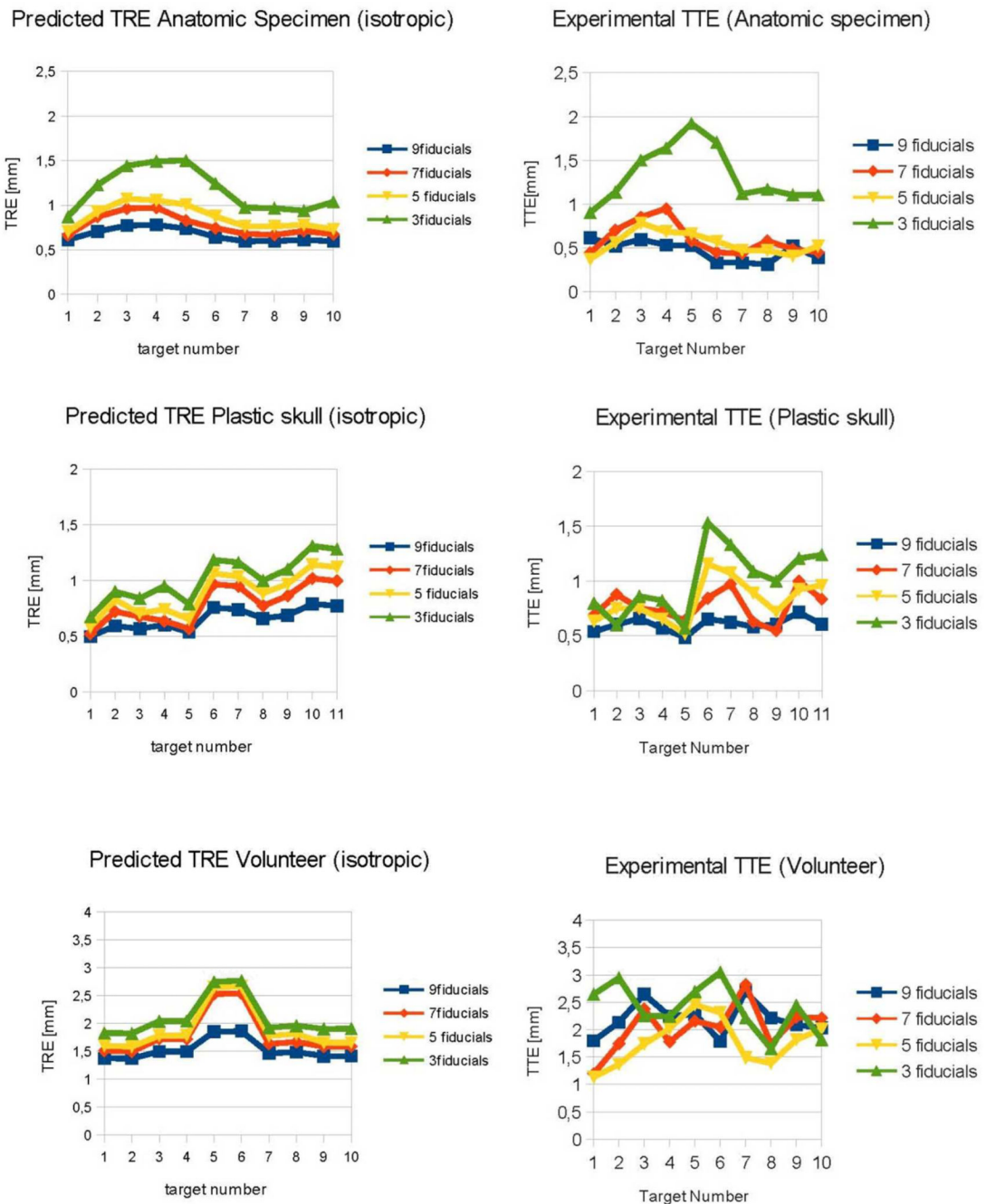Experimental **TTE**s for the three specimens, right column, and predictions of the **TTE** with the isotropic model[25], left column.

Figure 10.
Experimental **TTE**s for the three specimens, right column, and predictions of the **TTE** with the anisotropic model[27], left column.
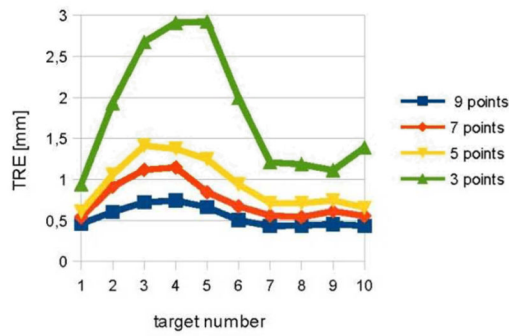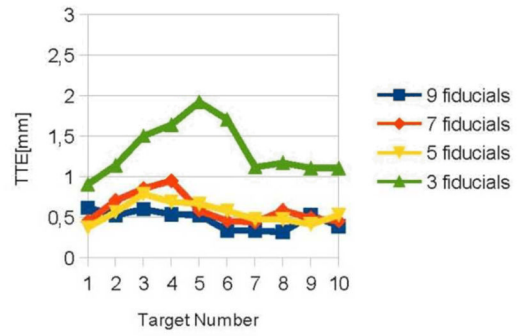
Figure 11.
Experimental **TTE**s for the three specimens, right column, and predictions of the **TRE** with the isotropic model[25], left column.
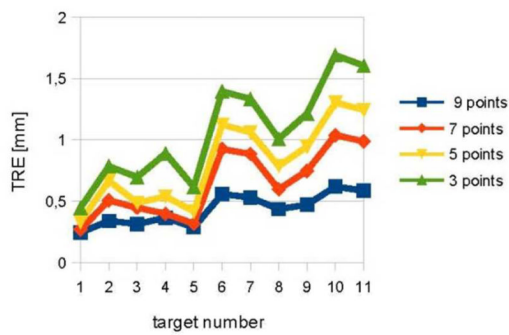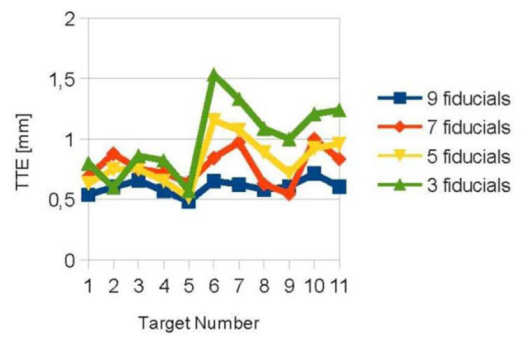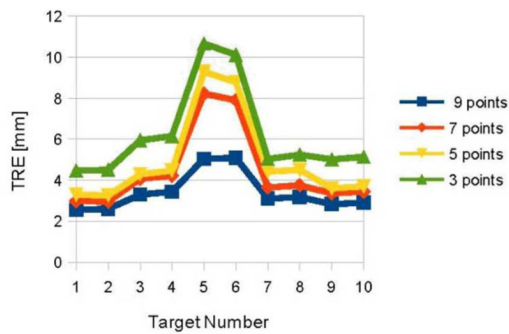
Figure 12.
Experimental **TTE**s for the three specimens, right column, and predictions of the **TRE** with the anisotropic model[27], left column.
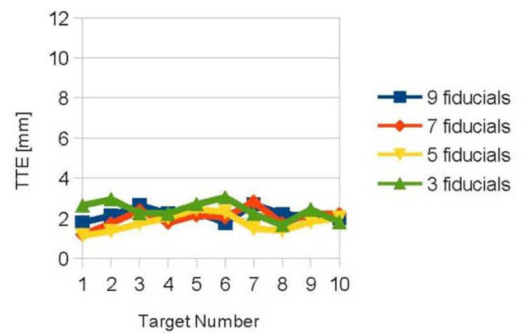
Table 1

Test for normal distribution of measurements for all measured objects, fiducials and targets, used in the experiments. Testing was done with the Shapiro-Wilk for $\alpha = 0.05$. Deviations from normal distribution of data are given in image coordinates.

	Target No.	Fiducial No.	Direction
Plastic skull	6		Z
	8		Z
		2	X
		3	X
		4	Z
Anatomic specimen	8		Z
	3		X
	8		Z
Volunteer		8	Z

Table 2

The measured covariance matrices of a DRF, the probe-DRF assembly, and the probe-tip-DRF calibration measurements as required for calculation of Σ_{TRE}.

a) placement of the DRF, probe-DRF and probe-tip-DRF in the center of the working volume. For the probe-DRF measurements the probe was placed in the bore of the DRF assembly to be as close as possible to the origin of the DRF, defined in one LED on it. Data in [mm²].	
$\Sigma_{DRF} = 10^{-5}\begin{pmatrix} 40.56 & -6.26 & 32.0 \\ -6.26 & 4.09 & -6.08 \\ 32.0 & -6.085 & 165.87 \end{pmatrix}$	tracker coordinates
$\Sigma_{probe-DRF} = 10^{-3}\begin{pmatrix} 8.68 & -0.14 & 0.05 \\ -0.14 & 0.29 & -0.04 \\ 0.05 & -0.04 & 0.12 \end{pmatrix}$	DRF coordinates
$\Sigma_{probe\_tip-DRF} = 10^{-3}\begin{pmatrix} 8.99 & -0.12 & 0.06 \\ -0.12 & 0.33 & -0.05 \\ 0.06 & -0.05 & 0.12 \end{pmatrix}$	DRF coordinates
b) DRF, probe-DRF and probe-tip-DRF each 150 mm left / right off the optimal center, but still in the optimal working volume. Data are in [mm²].	
$\Sigma_{DRF} = 10^{-5}\begin{pmatrix} 16.23 & 0.84 & -0.64 \\ 0.84 & 3.65 & -5.51 \\ -0.64 & -5.51 & 166.19 \end{pmatrix}$	tracker coordinates
$\Sigma_{probe-DRF} = 10^{-3}\begin{pmatrix} 52.11 & -12.96 & 1.07 \\ -12.96 & 6.83 & 0.11 \\ 1.07 & 0.11 & 0.23 \end{pmatrix}$	DRF coordinates
$\Sigma_{probe_{tip}-DRF} = 10^{-3}\begin{pmatrix} 48.85 & -11.81 & 1.21 \\ -11.81 & 7.71 & 0.22 \\ 1.21 & 0.22 & 0.24 \end{pmatrix}$	DRF coordinates

Table 3

Covariance matrices for defining features in image and tracker space, respectively, with the computer mouse and tracked probe for plastic skull and anatomic specimen from experimental data.

a) For fiducials:	
$\Sigma_{plastic\_skull}^{FLE_{image}} = 10^{-3} \begin{pmatrix} 40.91 & 6.85 & 1.85 \\ 6.85 & 41.81 & -0.5 \\ 1.85 & -0.5 & 22.80 \end{pmatrix}$	in image coordinates
$\Sigma_{anatomic\_specimen}^{FLE_{image}} = 10^{-3} \begin{pmatrix} 71.93 & 1.80 & 3.25 \\ 1.80 & 83.36 & -25.95 \\ 3.25 & -25.95 & 271.51 \end{pmatrix}$	in image coordinates
$\Sigma_{volunteer}^{FLE_{image}} = 10^{-3} \begin{pmatrix} 191.14 & -36.42 & 23.54 \\ -36.42 & 253.40 & 94.06 \\ 23.54 & 94.06 & 647.84 \end{pmatrix}$	in image coordinates
$\Sigma_{plastic\_skull}^{FLE_{pat}} = 10^{-3} \begin{pmatrix} 206.14 & 17.25 & 2.85 \\ 17.25 & 64.31 & -0.7 \\ 2.85 & -0.7 & 84.44 \end{pmatrix}$	in DRF coordinates
$\Sigma_{anatomic\_specimen}^{FLE_{pat}} = 10^{-3} \begin{pmatrix} 460.62 & 75.04 & -4.32 \\ 75.04 & 584.31 & -118.14 \\ -4.32 & -118.14 & 100.59 \end{pmatrix}$	in DRF coordinates
$\Sigma_{volunteer}^{FLE_{pat}} = 10^{-0} \begin{pmatrix} 11.3503 & -5.4361 & -14.7960 \\ 5.4361 & 5.7581 & 8.6004 \\ -14.7960 & 8.6004 & 23.3476 \end{pmatrix}$	in DRF-coordinates
b) For targets	
$\Sigma_{plastic\_skull}^{TLE_{DRF}} = 10^{-3} \begin{pmatrix} 146.50 & 19.23 & 12.10 \\ 19.23 & 139.67 & -27.26 \\ 12.10 & -27.26 & 86.82 \end{pmatrix}$	in DRF coordinates
$\Sigma_{plastic\_skull}^{TLE_{DRF}} = 10^{-3} \begin{pmatrix} 312.01 & 45.37 & 98.98 \\ 45.37 & 201.657 & 57.54 \\ 98.98 & 57.54 & 110.60 \end{pmatrix}$	in DRF coordinates
$\Sigma_{volunteer}^{TLE_{DRF}} = 10^{-0} \begin{pmatrix} 6.1140 & -3.0477 & -8.3359 \\ -3.0477 & 3.3940 & 5.2338 \\ -8.3359 & 5.2338 & 15.3394 \end{pmatrix}$	in DRF coordinates

Table 4

Measured values for $<\mathbf{FLE}_{image}^2>$, $<\mathbf{FLE}_{tracker}^2>$, $<\mathbf{FLE}_{probe\_calib}^2>$, and $<\mathbf{TFLE}^2>$ from the experiments with 9 registration points for plastic skull, anatomic specimen and volunteer, all values in millimeters. $<\mathbf{FLE}_{tracker}^2>$ and $<\mathbf{FLE}_{probe\_calib}^2>$ were obtained from the tracker measurements and probe calibration, respectively. Values are given in mm^2.

	$<\mathbf{FLE}_{image}^2>$	$<\mathbf{FLE}_{tracker}^2>$	$<\mathbf{FLE}_{probe\_calib}^2>$	$<\mathbf{TFLE}^2>$	$<\mathbf{TTE}^2>$
Plastic skull	0.26^2	0.22^2	0.18^2	0.73^2	1.11^2
Anatomic specimen	0.54^2	0.22^2	0.18^2	0.95^2	1.08^2
Volunteer	0.92^2	0.22^2	0.18^2	6.76^2	7.21^2

Table 5

$<$**ULE**$^2>$ from eqns. (13) and (14) for all sets of registration points (3, 5, 7, and 9) and objects (anatomic specimen, plastic skull, and volunteer) studied; the average over all fiducials is given. All values are given in mm^2. *Some loci of the plastic skull are anatomic, but resemble screws: holes originating from earlier implanted screws. ** Evaluation was possible with three loci only; all the others yielded complex numbers. See text.

	$<$ULE$^2>$ [mm^2] from eq. (13)	$<$ULE$^2>$ [mm^2] from eq. (14)
Plastic skull*	0.52^2	0.45^2
Anatomic specimen**	0.07^2	0.60^2
Volunteer	1.57^2	4.96^2