

# New Approaches to Online Estimation of Electromagnetic Tracking Errors for Laparoscopic Ultrasonography

Marco Feuerstein<sup>1</sup>, Tobias Reichl<sup>2,3</sup>, Jakob Vogel<sup>3</sup>, Joerg Traub<sup>3</sup>, and Nassir Navab<sup>3</sup>

<sup>1</sup> Department of Media Science, Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Aichi, Japan

<sup>2</sup> CSIRO BioMedIA Lab, Australian e-Health Research Centre, Brisbane, Australia

<sup>3</sup> Computer-Aided Medical Procedures (CAMP), Technische Universität München, Munich, Germany  
fmarco@suenaga.m.is.nagoya-u.ac.jp

**Abstract.** In abdominal surgery, a laparoscopic ultrasound transducer is commonly used to detect lesions such as metastases. The determination and visualization of position and orientation of its flexible tip in relation to the patient or other surgical instruments can be of great support for surgeons using the transducer intraoperatively. This difficult subject has recently been paid attention to by the scientific community. Electromagnetic tracking systems can be applied to track the flexible tip. However, current limitations of electromagnetic tracking are its accuracy and its sensibility, i.e. the magnetic field can be distorted by ferromagnetic material. This paper presents two novel methods for electromagnetic tracking error estimation. Based on optical tracking of the laparoscope as well as magneto-optic and visual tracking of the transducer, these methods automatically detect in 85 % of all cases, whether tracking is erroneous or not, and reduce tracking errors by up to 2.5 mm.

*Keywords:* Electromagnetic Tracking, Optical Tracking, Hybrid Tracking, Image-Guided Surgery, Augmented Reality, Laparoscopic Surgery

## Introduction

Laparoscopic ultrasonography (LUS) nowadays plays an increasing role in abdominal surgery. Its main application areas include the examination of liver, biliary tract, and pancreas. Unfortunately LUS is often difficult to perform, especially for novice surgeons. Therefore several groups tried to support surgeons by providing navigated LUS: The position and orientation (“pose”) of the ultrasound transducer is estimated using robot or optical tracking (OT) [1], electromagnetic tracking (EMT) [2–5], or magneto-optic tracking, i.e. the combination of OT and EMT [6]. Optical tracking, known to be one of the most accurate tracking technologies, always requires an unobstructed line of sight to OT bodies. This can not be established using flexible instruments in minimally invasive surgery. Mechanical tracking, i.e. the use of robotic control and thus additional hardware, is only justified in the presence of a robotic guidance system such as the da Vinci telemanipulator. The only currently available method to track flexible instruments is EMT.

\* Author Posting. (c) 'Informa UK Ltd.', 2008.

This is the author's version of the work. It is posted here by permission of 'Informa UK Ltd.' for personal use, not for redistribution.

The definitive version was published in Computer Aided Surgery, Volume 13 Issue 5, September 2008. doi:10.1080/10929080802310002 (<http://dx.doi.org/10.1080/10929080802310002>)

When using EMT clinically, considerable problems are fast sensor movements and the distortion of the EMT field, both leading to erroneous tracking data. Electrically powered or metallic objects inside or close to the working volume can cause this distortion. For example, operating room equipment like surgical tools, instruments, the operating table, or imaging devices such as a C-arm and CT scanner can lead to tracking errors of several millimeters or centimeters [7, 8]. Many calibration techniques were presented to correct measurement errors due to stationary ferromagnetic objects [9]. To utilize such a calibration technique, the user or a robot needs to record many well distributed distorted measurements inside the EMT volume along with their corresponding undistorted values. These measurements are used for a field distortion function based on polynomials or look-up tables, which reduces the static errors caused by a specific stationary operating room setup. This time consuming and laborious calibration process needs to be repeated for every new intervention, as the operating room setup and hence the distortion field changes between interventions. Additionally, instrument movements or the relocation of the EMT transmitter can cause dynamic changes of the field distortion. A previously computed distortion correction function will not be able to deal with such dynamic changes. Therefore, research groups started to develop solutions for detecting dynamic changes of the field distortion [10, 11]. Their common idea is to integrate two (rigidly connected) EMT sensors into a pointer or instrument in order to get redundant measurements. If the obtained measurements do not reflect the fixed distance between the two sensors, tracking errors can be identified and a plausibility value can be generated.

This paper builds upon our previous work [12–14] and describes two new approaches to detect and partially correct EMT errors online, i.e. intraoperatively without a pre-computed distortion correction function. Both methods are applied to a flexible laparoscopic transducer, the pose of which is determined by a magneto-optic tracking system. A rigorous accuracy evaluation of both online EMT error estimation methods was conducted. As an exemplary application of the improvement of the tracking quality, the B-scan images of the transducer are overlaid on the live images of an optically tracked laparoscope in real time. This may provide surgeons with a better understanding of the spatial relationship between the two imaging modalities and guide them with the required accuracy and reliability.

## System Setup

Our hardware setup consists of the following components: A SONOLINE Omnia US system by Siemens Medical Solutions equipped with a flexible laparoscopic linear array transducer (LAP8-4, 5 MHz, 10 mm diameter), a laparoscopic camera with a forward-oblique 30° HOPKINS telescope by Storz, a workstation PC including two frame grabbers (for capturing the transducer and camera video in real time), and our hybrid magneto-optic tracking system. Its optical component comprises 4 ARTtrack2 cameras and a book-size PC running the DTrack tracking software. The electromagnetic component is a 3D Guidance unit by Ascension equipped with a mid-range transmitter and insulated 1.3 mm sensors, which have a total diameter of 1.7 mm including the vinyl tubing. Time synchronization of all (video and tracking)

data streams, visualization, and the user interface is implemented within our medical augmented reality software platform CAMPAR [15].

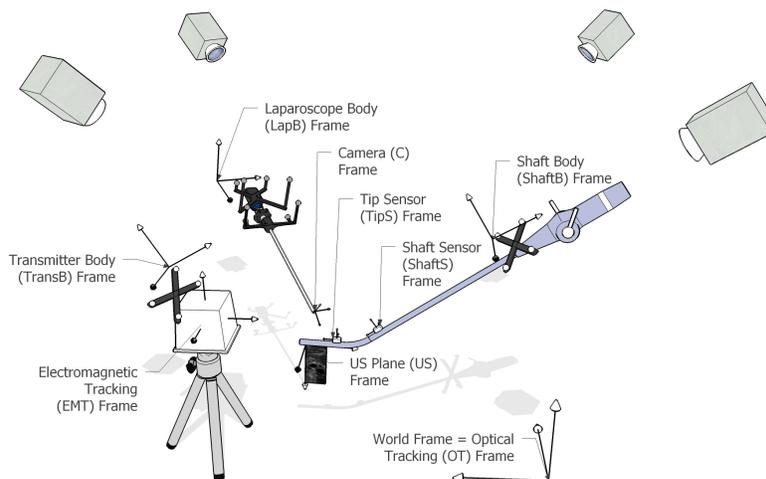
## Methods

In the following section, a “body” always refers to an optical tracking (OT) body consisting of usually four or more retroreflective spheres, while “sensor” always refers to a wired electromagnetic tracking (EMT) sensor. Both types are used to acquire 6DOF pose information in their respective tracking coordinate frames.

In addition to an OT body, which is attached to the transducer handle (below referred to as “shaft body”), two EMT sensors are attached to the transducer: One to the flexible tip (“tip sensor”), the other one to the rigid shaft (“shaft sensor”), as close to each other as possible. Another OT body is mounted on the EMT transmitter (“transmitter body”). This setup allows us to co-calibrate EMT and OT and to obtain redundant tracking information of the rigid part of the transducer shaft, which is important to detect EMT errors. Finally, two OT bodies are attached to the laparoscopic camera, one to the head (“laparoscope body”) and another one to the telescope to adjust for telescope rotations.

## System Calibration

Spatial and temporal system calibration is performed offline in a distortion-free environment. All coordinate frames, which we calibrate to each other, are shown in figure 1.



**Fig. 1.** Coordinate frames associated with our hybrid magneto-optic tracking setup to localize the laparoscope and the tip of the laparoscopic ultrasound.

**Hand-eye Calibration** To compute the Euclidean transformation  ${}^{ShaftB}T_{ShaftS}$  between the shaft body and the shaft sensor frames, the transducer is moved to a set

of  $n$  stations with distinct rotation axes, so  $n$  pairs of poses can be recorded in both the OT and EMT coordinate frames. Two stacked matrices  $\mathbf{A}_{OT_{4m \times 4}}$  and  $\mathbf{B}_{EMT_{4m \times 4}}$  are then generated from all  $m = n(n-1)/2$  possible unidirectional motions between these pose pairs. Each stacked matrix therefore consists of  $m$  homogeneous transformation matrices. The stacked matrices are related to each other by the equation system  $\mathbf{A}_{OT}^{ShaftB} \mathbf{T}_{ShaftS} = {}^{ShaftB} \mathbf{T}_{ShaftS} \mathbf{B}_{EMT}$ , which is commonly written as  $\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{B}$  and solved by hand-eye calibration [16].

The same transducer stations can also be used to estimate the rigid hand-eye transformation  ${}^{EMT} \mathbf{T}_{TransB}$  between the EMT transmitter coordinate frame and its rigidly mounted OT body.

In a final optimization step, the two hand-eye calibration matrices  ${}^{ShaftB} \mathbf{T}_{ShaftS}$  and  ${}^{EMT} \mathbf{T}_{TransB}$  are optimized for all recorded poses using the Levenberg-Marquardt algorithm. The matrix  $\mathbf{T}_\delta \in \mathbb{R}^{4 \times 4}$  resulting from the transformation chain “shaft sensor to shaft body to OT to transmitter body to EMT to shaft sensor frame”, which ideally would be an identity matrix in an error-free setup, represents the accumulated transformation errors:

$$\mathbf{T}_\delta = \begin{bmatrix} \mathbf{R}_\delta & \mathbf{t}_\delta \\ 0 & 1 \end{bmatrix} := {}^{ShaftS} \mathbf{T}_{EMT} \cdot {}^{EMT} \mathbf{T}_{TransB} \cdot {}^{TransB} \mathbf{T}_{OT} \cdot {}^{OT} \mathbf{T}_{ShaftB} \cdot {}^{ShaftB} \mathbf{T}_{ShaftS} \quad (1)$$

From the rotational part  $\mathbf{R}_\delta \in \mathbb{R}^{3 \times 3}$  and the translational part  $\mathbf{t}_\delta \in \mathbb{R}^3$  of  $\mathbf{T}_\delta$  we can compute a cost function  $\delta$  that weights translational errors in millimeters to rotational errors in degrees 1:3

$$\delta = \delta_{translational} + 3 \cdot \delta_{rotational} = \|\mathbf{t}_\delta\| + 3 \cdot \frac{180}{\pi} \cdot \arccos\left(\frac{\text{trace}(\mathbf{R}_\delta) - 1}{2}\right) \quad (2)$$

where the rotational error  $\delta_{rotational}$  is the rotation angle of  $\mathbf{R}_\delta$  decomposed into axis-angle parameters.

The 1:3 ratio reflects the root mean squared (RMS) error ratio provided independently by the two tracking system manufacturers: The RMS measurement errors of the OT system are stated as 0.4 mm (position) and 0.12° (orientation), the static RMS errors of the EMT system as 1.4 mm and 0.5°. (See also <http://www.ar-tracking.de> and <http://www.ascension-tech.com> for the specifications of the typical accuracy of both tracking systems.)

Actually we optimize both sought transformations  ${}^{ShaftB} \mathbf{T}_{ShaftS}$  and  ${}^{EMT} \mathbf{T}_{TransB}$  using the results of a closed-form solution [16] as starting values. Although to our knowledge no work has been done evaluating a further optimization, we strongly believe that minimizing the error component relevant to our particular approach improves our performance. The Levenberg-Marquardt algorithm works in real time, as we have a very good initialization. With this initialization, any other iterative optimizer would have given very similar results.

The maximum error  $\delta_{thresh}$  of all recorded poses determined after optimization is chosen as a measure of distrust for the overall performance of the hand-eye calibration (cf. section *Online Error Detection*). Alternatively, an optimal  $\delta_{thresh}$  can be selected

for specific requirements according to computed receiver operating characteristics (cf. section *Receiver Operating Characteristics*).

**Laparoscopic Camera Calibration** For camera calibration of the forward-oblique laparoscope, the projection geometry including distortion coefficients [17] and the transformation from laparoscope body coordinates to camera center coordinates are estimated [18]. A more detailed description can also be found in our previous work [19].

**Transducer Tip Axis Calibration** We calibrate the axis of the transducer tip by rotating a calibration cylinder, which contains an additional EMT sensor at one end, around the transducer tip. Overall, two full 360 degree rotations are performed, the second one with the calibration cylinder flipped. The measurements during the two rotations of the additional sensor describe two circular patterns, which can be used to estimate the transducer tip axis. The axis is stored as a “base point”  $\mathbf{b}_{TipS}$ , which is the point on the axis closest to the origin of the tip sensor coordinate frame, and a directional unit vector  $\mathbf{d}_{TipS}$ , which points towards the tip, both given in the local coordinate frame of the tip sensor mounted on the flexible transducer tip [13].

**Temporal Calibration** In order to provide a smooth visualization without lag, all data is given a time stamp and brought into the same time frame. While the OT PC and our workstation are synchronized via the Network Time Protocol (NTP) to the same reference time, the ultrasound and EMT systems require a more advanced synchronization. As these systems do not automatically provide reliable time stamps corresponding to the actual data acquisition time, a time stamp is generated when their data arrives at the workstation. Therefore a fixed offset is subtracted from this time stamp to compensate for any lag introduced while traveling to the workstation. To determine this offset, the magneto-optically tracked transducer is moved up and down and the translation along the principal motion axes is compared, as proposed by Treece et al. [20].

### Online Error Detection

Intraoperatively, for all measurements of the pose of the shaft sensor the deviation from ideal tracking is computed using equation 1. If the corresponding distrust value  $\delta$  (cf. equation 2) is bigger than our previously determined threshold  $\delta_{thresh}$ , the surgical staff is automatically warned. Such errors are often caused by dynamic or static field distortions. Additionally, as the tip sensor is in close proximity to the rigid one, its measurements will most likely be affected by these distortions as well, and this can be used for a redundancy-based approach to error correction.

### Online Error Correction

**Redundancy-Based Correction** In order to also approximate a correction of erroneous measurements of the tip sensor, one approach (also referred to as tracking

redundancy-based approach) is to apply the deviation between the previously hand-eye calibrated (“calib”) and the measured (“meas”) transformation of the shaft sensor to the measured tip sensor transformation, all relative to the fixed OT (world) reference frame:

$${}^{OT}\mathbf{R}_{TipS(corr)} = {}^{OT}\mathbf{R}_{ShaftS(meas)}^T \cdot {}^{OT}\mathbf{R}_{ShaftS(calib)} \cdot {}^{OT}\mathbf{R}_{TipS(meas)} \quad (3)$$

$${}^{OT}\mathbf{t}_{TipS(corr)} = -{}^{OT}\mathbf{t}_{ShaftS(meas)} + {}^{OT}\mathbf{t}_{ShaftS(calib)} + {}^{OT}\mathbf{t}_{TipS(meas)} \quad (4)$$

**Vision-Based Correction** Following common surgical procedures, the LUS probe tip already has to be constantly monitored to prevent inadvertent injury of the patient, so laparoscopic images of it are readily available.

As the intrinsic and extrinsic camera parameters of the laparoscope and hence the spatial location of the image plane are known, another approach to improve the tracking accuracy of the tip sensor is to automatically detect the transducer tip in the images of the laparoscope camera and align the detected transducer tip axis with the tracked axis.

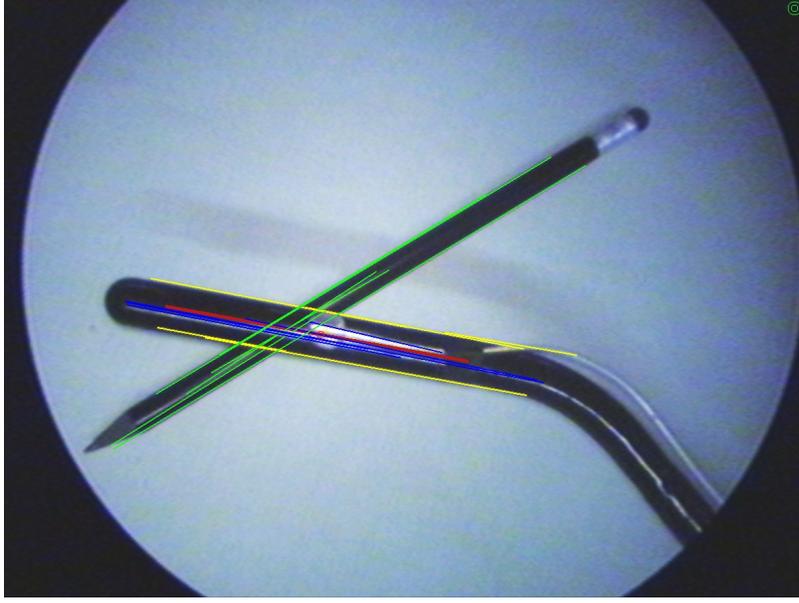
For an automatic detection of the ultrasound tip, we follow approaches that already showed promising results under conditions close to real laparoscopic surgery: similar to Climent and Marés [21] and Voros et al. [22] we use an edge detection filter and a Hough transformation [23] to extract edges from laparoscopic images. We also use additional information to select candidate lines belonging to the transducer edges. Voros et al. [22] and Doignon et al. [24] use information about the insertion points of laparoscopic instruments, which stay relatively fixed during an intervention and thus can be used to support the segmentation of instruments. However, in our case the laparoscopic ultrasound transducer might be bent so that its edges are no more aligned with the insertion point. Instead of information on the insertion point, we use the tracking data of the tip sensor.

*Line Detection:* The results of our transducer tip axis segmentation algorithm are illustrated in figure 2.

To find the 2D image coordinates of the transducer tip axis, the Open Source Computer Vision Library (<http://www.intel.com/technology/computing/opencv/>) is used to automatically segment the transducer tip axis in the undistorted laparoscope images in real time. First, the Canny edge detection algorithm [25] is applied to provide a binary image of edges, which is fed into a Hough transform [23] to give a set of lines in the camera image. We obtain the end points of each line.

To find the two lines corresponding to the two edges of the transducer tip, the whole set of segmented lines is first back-projected into 3D space, i.e. each end point  $\mathbf{x}_C$  given in image coordinates (pixels) is projected back to  $\mathbf{X}_C$  given in camera coordinates (millimeters):

$$\mathbf{X}_C = \begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} := \begin{bmatrix} X_C \\ Y_C \\ 1 \end{bmatrix} = \mathbf{K}^{-1} \begin{bmatrix} \mathbf{x}_C \\ 1 \end{bmatrix} \quad (5)$$



**Fig. 2.** Screenshot of axis segmentation. Lines classified as belonging to the transducer tip edges are automatically colored yellow, lines belonging to the transducer (but not to the edges) are colored blue, the corrected transducer axis is thick red. Lines belonging to the pencil are rejected (colored green), because they do not match the measured transducer axis rotation.

where  $\mathbf{K}$  is the  $3 \times 3$  camera calibration matrix containing the principal point and focal length [17] and  $\mathbf{X}_C$  lies on the image plane, i.e.  $Z_C = 1$ . Together with the camera center, each line represented by its two end points  $\mathbf{X}_{C1}$  and  $\mathbf{X}_{C2}$  defines a plane, which can be completely described by its unit normal  $\mathbf{n} = \frac{\mathbf{X}_{C1} \times \mathbf{X}_{C2}}{\|\mathbf{X}_{C1} \times \mathbf{X}_{C2}\|}$ .

As illustrated in figure 3, all planes are now compared to the measurements of the transducer tip axis (which is defined by  $\mathbf{b}_{TipS}$  and  $\mathbf{d}_{TipS}$  in tip sensor coordinates; cf. section *Transducer Tip Axis Calibration*), acquired by EMT and transformed into camera coordinates:

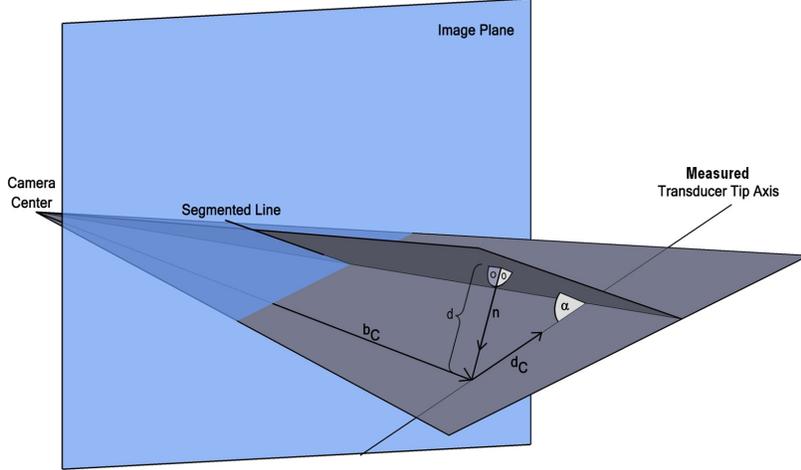
$$\begin{aligned} \mathbf{b}_C &= {}^C\mathbf{T}_{TipS} \mathbf{b}_{TipS} \\ \mathbf{d}_C &= {}^C\mathbf{T}_{TipS} \mathbf{d}_{TipS} \end{aligned} \quad (6)$$

where

$${}^C\mathbf{T}_{TipS} = {}^C\mathbf{T}_{LapB} {}^{LapB}\mathbf{T}_{OT} {}^{OT}\mathbf{T}_{TransB} {}^{TransB}\mathbf{T}_{EMT} {}^{EMT}\mathbf{T}_{TipS} \quad (7)$$

${}^C\mathbf{T}_{LapB}$  is the transformation from the laparoscope OT body to the camera center,  ${}^{LapB}\mathbf{T}_{OT}$  is the transformation from the world (OT) coordinate system to the laparoscope OT body, and  ${}^{OT}\mathbf{T}_{TransB}$  is the transformation from the EMT transmitter OT body to the world (OT) coordinate system.

In order to obtain a unified representation of all planes, we adjust each of the respective normals  $\mathbf{n}$  to point into the same direction as the vector defined by the cross product of  $\mathbf{d}_C$  and  $\mathbf{b}_C$  (cf. figure 3), i.e. if  $\mathbf{n} \cdot (\mathbf{d}_C \times \mathbf{b}_C) < 0$ , we will negate  $\mathbf{n}$ .



**Fig. 3.** Back-projection of a segmented line and its comparison to the transducer tip axis measured by EMT.

The angle  $\alpha$  between the measured transducer tip axis and each plane can be determined by

$$\alpha = \arcsin(\mathbf{n} \cdot \mathbf{d}_C) \quad (8)$$

The distance  $d$  between the base point of the measured transducer tip axis and the plane is described by

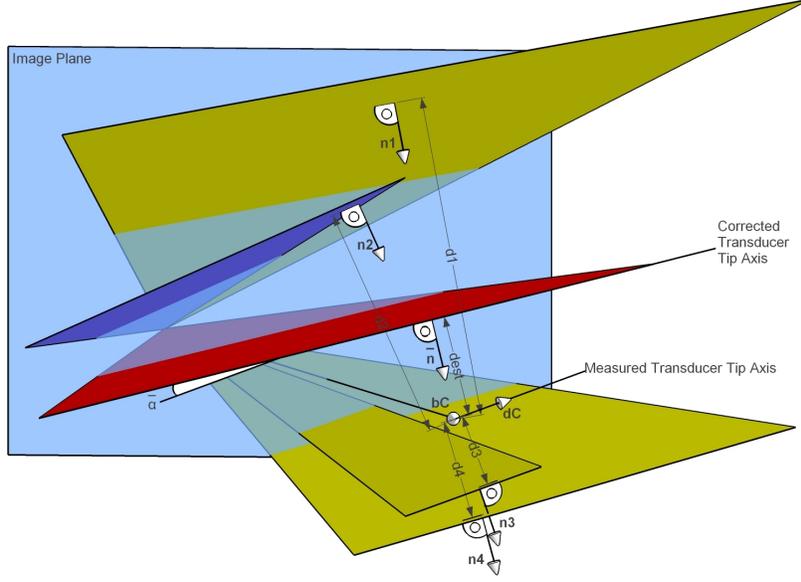
$$d = \mathbf{n} \cdot \mathbf{b}_C \quad (9)$$

Depending on whether  $d$  is positive, negative, or zero, the base point  $\mathbf{b}_C$  of the measured transducer tip axis will be above (on the half-space, the normal is pointing to), beneath, or on the plane.

For each line,  $|\alpha|$  and  $|d|$  are compared to the thresholds  $\alpha_{thresh}$  and  $d_{thresh}$ , respectively. If both parameters are below the corresponding threshold, it can be assumed that the current line corresponds to an edge of the transducer tip. To take care of providing enough tolerance to compensate for erroneous EMT measurements, the thresholds are chosen in a way that on the one hand segmented lines are selected which are definitely part of the transducer tip, and on the other hand lines are rejected which are actually part of the image background. We empirically determined  $\alpha_{thresh} = 5$  (degrees) and  $d_{thresh} = 30$  (millimeters). These values gave a good balance between stability against distortions and the potential for additional error correction.

*Correction Transformation:* To compute the final correction transformation, we analyze the previously evaluated set of lines belonging to the edges of the transducer tip, as illustrated in figure 4. Iterating over all these lines, the distance  $d$  between the plane described by the back-projection of each line and the base point of the measured transducer tip axis is computed, and in both directions the greatest distance is stored. Because the sign of  $d$  is different for both directions, this is equivalent to storing the maximum negative distance  $d_{neg_{max}}$  and the maximum positive distance  $d_{pos_{max}}$ . Ideally, the absolute difference between these distances  $|d_{pos_{max}} - d_{neg_{max}}|$  is equal to the diameter of the ultrasound transducer tip, which is 10 millimeters. If this absolute difference stays within certain limits, say  $10 \pm 2$  mm, it can be assumed

with high probability that lines were extracted which belong to both edges of the transducer.



**Fig. 4.** Back-projection of four segmented lines, which generates four planes and their corresponding normals  $n_1$ ,  $n_2$ ,  $n_3$ , and  $n_4$  (plane coloring according to figure 2). For each plane  $j$ , its distance  $d_j$  to  $\mathbf{b}_C$  and angle  $\alpha_j$  to  $\mathbf{d}_C$  are computed. We suppose that all four planes satisfy  $|\alpha_j| < 5^\circ$  and  $|d_j| < 30$  mm. Because  $d_1$  is the maximum positive distance from  $\mathbf{b}_C$ , and  $d_4$  the maximum negative distance from  $\mathbf{b}_C$ , we can set  $d_{pos_{max}} = d_1$  and  $d_{neg_{max}} = d_4$ . If  $|d_{pos_{max}} - d_{neg_{max}}| = |d_1 - d_4|$  stays within  $10 \pm 2$  mm, we continue our computations. For all distances  $i$  unequal  $d_{pos_{max}}$  and  $d_{neg_{max}}$ , we check, whether  $d_{pos_{max}} - d_i < 2$  mm or  $d_i - d_{neg_{max}} < 2$  mm. All planes not satisfying this criterion are excluded from further computations, as for plane 2 (blue). Using the remaining three yellow planes, a mean plane (red) defined by  $\bar{\mathbf{n}}$  and  $\bar{\alpha}$  is determined. The transducer tip axis can finally be translated along  $\bar{\mathbf{n}}$  for  $d_{est} = 0.5(d_1 + d_4)$  mm and rotated by  $\bar{\alpha}$  around  $\mathbf{r} = \bar{\mathbf{n}} \times \mathbf{d}_C$ .

Because of e.g. reflections on the transducer tip surface there might be artifacts affecting lines across the tip, so we want to exclude lines that are not closely aligned with its probable edges. Thus, for the computation of the correction transformation we use information about lines within a certain maximum distance, say 2 mm, from the outermost lines. This is the case for all lines  $i$ , where either  $d_{pos_{max}} - d_i < 2$  mm or  $d_i - d_{neg_{max}} < 2$ . From these  $i = 1 \dots n$  lines we compute the mean plane normal  $\bar{\mathbf{n}} = \frac{\sum_{i=1}^n \mathbf{n}_i}{\|\sum_{i=1}^n \mathbf{n}_i\|}$  and the mean angle  $\bar{\alpha} = \frac{\sum_{i=1}^n \alpha_i}{n}$  between transducer tip axis and plane.

The distance  $d_{est}$  between segmented transducer axis and electromagnetically measured transducer axis can be estimated as the average of the minimum and maximum distances  $d_{est} = 0.5(d_{pos_{max}} + d_{neg_{max}})$ .

When translating the measured transducer tip axis along the mean plane normal  $\bar{\mathbf{n}}$  by the estimated distance  $d_{est}$ , the axis origin will be in the middle of the segmented transducer tip. Next, the tip axis needs to be rotated into the plane. Since the rotation axis  $\mathbf{r}$  has to be orthogonal to the plane normal as well as to the measured tip axis direction, it can be computed as  $\mathbf{r} = \bar{\mathbf{n}} \times \mathbf{d}_C$ . Together with the mean angle  $\bar{\alpha}$  between measured tip axis and plane, a homogeneous correction transformation

can be estimated: The translation component along the mean plane normal can be calculated as  $d_{est} \cdot \bar{\mathbf{n}}$  and the rotation component can be computed from  $\mathbf{r}$  and  $\bar{\mathbf{a}}$  using Rodrigues’ rotation formula [26]. This transformation maps the electromagnetically measured tip axis to a pose, from where it can be projected onto the image plane in such a way that it is exactly aligned with the segmented axis of the transducer tip.

Note that for the computation of the correction transformation we only used the original electromagnetic tracking measurements of the tip sensor and the laparoscope video as input into our algorithm for adjusting the axis alignment of the transducer tip. No initial correction is performed before, e.g. by applying the redundancy-based method.

## Experimental Evaluation Results

To reliably achieve meaningful results, all EMT measurements were acquired in a tracking volume of 20–36 cm for  $x$ , and  $\pm 15$  cm for  $y$  and  $z$ , as verified by Ascension for their microBIRD system [27] to yield the most accurate measurements for the sensors and transmitter in use. We did not experience any issues at the borders of this volume, so we strongly expect our method to be applicable to larger volumes, especially because our artificially introduced distortions should by far outweigh any inaccuracies from exceeding the optimal tracking volume.

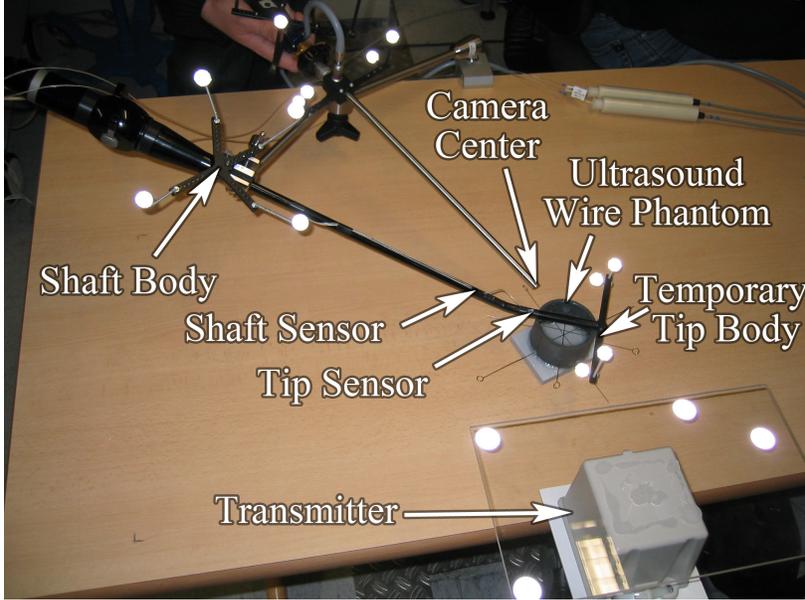
### Error Detection

In order to estimate the laparoscope augmentation errors automatically, an additional OT body (“tip body”) was temporarily attached to the transducer tip and co-calibrated to the tip sensor by another hand-eye calibration (cf. section on system calibration and figure 5). One marker of the tip body was chosen as a reference and automatically segmented whenever visible in the laparoscopic video. We compared its center coordinates to the projection of its respective OT coordinates onto the image plane. The corresponding EMT measurements as well as their approximated corrections were projected using the previously determined hand-eye calibration transformations.

Evaluation data was recorded using a laparoscope-to-marker distance  $d_{lap.to.m}$  of 5 to 10 cm, which is a typical intraoperative working distance. The current distance can be recovered from OT data and the camera calibration parameters, and thus from the position  $\mathbf{m}_C$  of the marker in respect to the camera coordinate system.  $\mathbf{m}_C$  can be computed similarly to equation 7 by:

$$\begin{bmatrix} \mathbf{m}_C \\ 1 \end{bmatrix} = {}^C\mathbf{T}_{LapB} \cdot {}^{LapB}\mathbf{T}_{OT} \cdot \begin{bmatrix} \mathbf{m}_{OT} \\ 1 \end{bmatrix} \quad (10)$$

where  ${}^C\mathbf{T}_{LapB}$  is the transformation from the laparoscope OT body to the camera center,  ${}^{LapB}\mathbf{T}_{OT}$  is the transformation from the world (OT) coordinate system to the laparoscope OT body, and  $\mathbf{m}_{OT}$  is the current position of the marker in respect to the world (OT) coordinate system. The current laparoscope-to-marker distance  $d_{lap.to.m}$  can now be determined from  $\mathbf{m}_C$  by simply computing its norm:  $d_{lap.to.m} = \|\mathbf{m}_C\|$ .



**Fig. 5.** Setup for error evaluation.

We also used this information to scale pixel units to mm. The distance in pixels of the projected OT marker center to the camera center can be computed from the in-plane distance to the principal point as follows:

$$\hat{d}_{lap.to.m} = \left\| \begin{pmatrix} \hat{d}_x & \hat{d}_y & \hat{f} \end{pmatrix}^T \right\| \quad (11)$$

where  $\hat{d}_x$  and  $\hat{d}_y$  are the distances of the projected OT marker center from the principal point in  $x$  and  $y$  direction, respectively, and  $\hat{f}$  is the focal length in pixels. As our laparoscope camera has nearly square pixels,  $\hat{f}$  can be defined as the mean between the focal lengths given in  $x$  and in  $y$  direction. The current ratio between millimeter units (referring to the spatial position of the marker) and pixel units (referring to its position in the image plane) can then be computed as  $d_{lap.to.m}/\hat{d}_{lap.to.m}$ .

For each of six evaluation series, the transducer was fixed at a different pose and the laparoscope was used to measure the projected distances from five differing poses, each in an undistorted and a distorted environment. To distort the EMT field, two alternatives were evaluated: A metal plate was placed on the table to simulate primarily static distortions caused for instance by an operating table. For dynamic distortions, a steel rod of 10 mm diameter was brought close to the transducer to simulate a surgical instrument, changing its proximity and angle to the transducer in five measurements.

In order to evaluate our distrust function statistically, we computed the distrust level (cf. equation 2) for each of the poses. An offset between the segmented marker and the EMT projections of more than 2 mm was regarded as erroneous measurement. In this case, we expect a distrust level  $\delta$  of more than  $\delta_{thresh}$  (during hand-eye calibration,  $\delta_{thresh}$  was empirically determined to be 20). We defined the following cases for our evaluation:

- A *true positive* is a measurement, in which the EMT error was above 2 mm with a distrust level of above 20 – the detector rejected an erroneous reading correctly.
- A *true negative* is a measurement, in which the EMT error was below 2 mm with a distrust level below 20 – we correctly accepted the original EMT data.
- A *false positive* (type 1 error) is a measurement, in which the EMT error was below 2 mm, but the distrust level above 20 – we have not been able to detect a correct value and rejected it without necessity.
- A *false negative* (type 2 error) is a measurement, in which the EMT error was above 2 mm, but the distrust level below 20 – the record was accepted although the real error was large.

The results are listed in table 1. In about 85 % of all cases, we correctly detected the true situation (*true positives* and *true negatives*).

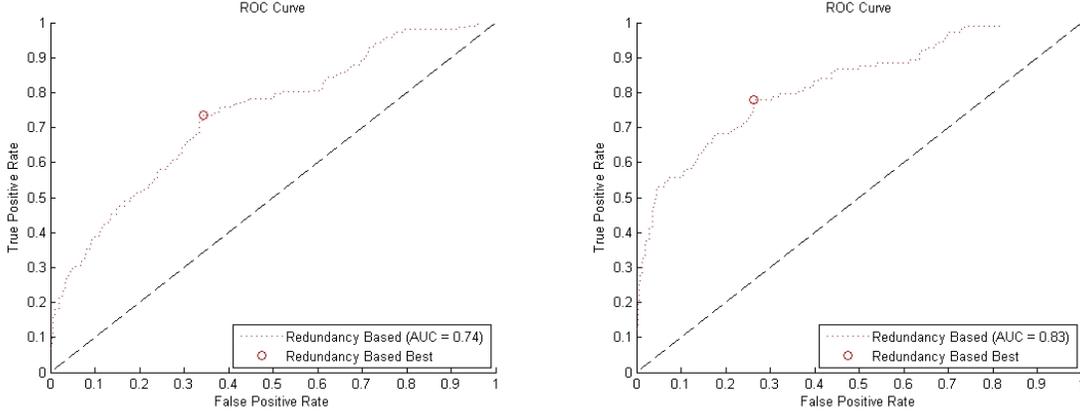
**Table 1.** Distortion detection rate by our distrust level without distortion, with static, and with dynamic field distortion.

distortion		true	false
without:	positive	40.0%	10.0%
	negative	30.0%	20.0%
static:	positive	100.0%	0.0%
	negative	0.0%	0.0%
dynamic:	positive	73.8%	13.8%
	negative	12.4%	0.0%
average:	positive	71.3%	7.9%
	negative	14.1%	6.7%

**Receiver Operating Characteristics** Additionally, for our set of distorted measurements we computed several receiver operating characteristics (ROC) for predicting errors between 2.5 mm and 10 mm. ROCs have the benefit of not considering a single, possibly manually chosen threshold, as done in the evaluation above, but of considering the whole range of possible thresholds. They therefore extend the analysis of our error prediction performance.

Our error prediction algorithms compute a distrust level for each measurement, and using this distrust level the measurement can be considered either correct or erroneous in comparison to a selected threshold  $\delta_{thresh}$ . The computation used to determine the distrust level remains fixed, but according to the selected threshold various rates of false positive or false negative decisions can be achieved. An ROC curve visualizes the interdependence between low false positive and low false negative rates, displaying all possible trade-off results based upon all possible thresholds. The performance of our redundancy-based method to classify measurements of the flexible tip sensor position is shown in figure 6.

A key metric for evaluating ROCs commonly used in statistics is the Youden index [28], which is defined as follows:



(a) Prediction of an error of 5.0 mm or greater.

(b) Prediction of an error of 7.5 mm or greater.

**Fig. 6.** ROC curves for error prediction.

$$J = \frac{ad - bc}{(a + b)(c + d)} = \underbrace{\frac{a}{a + b}}_{\text{Sensitivity}} + \underbrace{\frac{d}{c + d}}_{\text{Specificity}} - 1 \quad (12)$$

where  $a$  is the fraction of true positives,  $b$  the fraction of false negatives,  $c$  the fraction of false positives, and  $d$  the fraction of true negatives. The possible range of values is from zero to one inclusively. In figures 6(a) and 6(b) we also marked those values with the maximum Youden index, because those can be considered to yield the best trade-off. Depending on the application other values might be favored, i.e. if low false positive rates are favorable over low false negative rates or vice versa.

In addition to the figures, we also present the key values for each ROC in table 2: for both algorithms we give the the area under the ROC curve (AUC) and the maximum Youden index  $J_{max}$ . For  $J_{max}$  we also give the corresponding threshold value  $\delta_{thresh}$ , sensitivity (= true positive rate, TPR) and specificity (SPC), and both the smallest false positive value  $FP_{min}$  and the greatest false negative value  $FN_{max}$  are given.  $FP_{min}$  and  $FN_{max}$  are the most extreme case where our method would have led to a wrong classification.

The redundancy-based prediction of an error of 5 mm or greater would have achieved a sensitivity of 74% and a specificity of 66% in the best case, i.e. it would have correctly predicted 74% of all errors of 5 mm or greater and correctly predicted 66% of errors below 5 mm. For the prediction of an error of 7.5 mm or greater, sensitivity and specificity of our method would have been 78% and 74%. For the remaining values including those for 2.5 mm and 10 mm, see table 2.

## Error Correction

For assessing the overlay accuracy in both the undistorted and distorted case, the ultrasound transducer was fixed in various poses and the laparoscope was used to observe the transducer tip from various angles and distances.

**Table 2.** Receiver operating characteristic key figures for prediction of distortion errors of at least 2.5, 5, 7.5, and 10 mm

	AUC	$J_{max}$	$\delta_{thresh}$	TPR	SPC	$FP_{min}$	$FN_{max}$
2.5 mm	0.65	0.24	6.26	0.91	0.33	0.84	6.44
5.0 mm	0.74	0.39	11.34	0.74	0.66	0.84	13.69
7.5 mm	0.83	0.52	13.98	0.78	0.74	0.85	13.69
10.0 mm	0.86	0.67	27.75	0.72	0.94	0.88	14.29

In the course of the experiments the transducer tip was steered to different angles and the laparoscope was also rotated around its own axis. For distorting the magnetic field we used the steel rod with 10 mm diameter again.

At each measurement, the uncorrected position of the flexible tip sensor, the tracking redundancy-based corrected position of the flexible tip sensor, and the vision-based corrected position of the flexible tip sensor were transformed using the transformation  ${}^{TipB}\mathbf{T}_{TipS}$  from the flexible tip sensor to the flexible tip OT body. The resulting positions were then projected into the image plane, the spatial location of which was known from camera calibration. Also, the measured three-dimensional position of the flexible tip body (centered in one of the markers) was projected into the image plane. The distance in millimeters within the image plane to the segmented midpoint of the OT marker was computed and taken as a measure for the overlay accuracy.

Only measurements within a distance of 5 to 20 cm between the OT marker and the camera center were accepted. This rather high distance of up to 20 cm is required to observe both the transducer tip (for axis segmentation) and the marker at the same time. Keeping such a high distance is however very unlikely during surgery, wherefore the laparoscope was calibrated to be most accurate for a maximum working distance of about 10 cm only. This reduces the overlay accuracy when the laparoscope is further away than 10 cm, so theoretically the results obtained here could be further improved.

To assess the overlay accuracy of the two error correction methods we took 207 undistorted and 935 distorted measurements. We projected the target point onto the image plane using only the OT (projected OT), only the EMT (projected EMT), the combination of OT and EMT (Redundancy-Based Correction), and the correction using the image information (Vision-Based Correction) and we computed its distance to the centroid of a marker segmented in the image.

For the results in both undistorted and distorted cases see table 3, where we computed the minimum (Min), maximum (Max), mean, and root-mean squared (RMS) errors and their standard deviation (SD). As illustrated, the simple tracking redundancy-based error correction as well as the vision-based error correction approach yielded improvements compared to the uncorrected flexible tip sensor. In both undistorted and distorted environments, the vision-based method is superior to the redundancy-based method.

## Exemplary Application: Ultrasound Augmentation

**Ultrasound Calibration** For the determination of the pixel scaling of the ultrasound B-scan plane and its transformation to the tip sensor frame, a single-wall

**Table 3.** Overlay errors in an undistorted and a distorted field

	Min	Mean	SD	RMS	Max
<hr/> <hr/> Undistorted (mm)					
Projected OT	0.11	2.67	1.38	3.00	7.57
Projected EMT	0.17	3.57	2.49	4.35	11.81
Redundancy-Based Correction	0.38	3.73	1.99	4.23	11.19
Vision-Based Correction	0.20	2.91	1.75	3.39	9.51
<hr/> <hr/> Distorted (mm)					
Projected OT	0.05	1.81	1.02	2.08	6.10
Projected EMT	0.11	10.03	7.81	12.71	39.84
Redundancy-Based Correction	0.07	8.55	7.63	11.45	36.66
Vision-Based Correction	0.17	7.77	6.60	10.19	38.43

calibration is performed [20]. Instead of scanning the planar bottom of a water bath, we scan a nylon membrane stretched over a planar frame, as proposed by Langø [29].

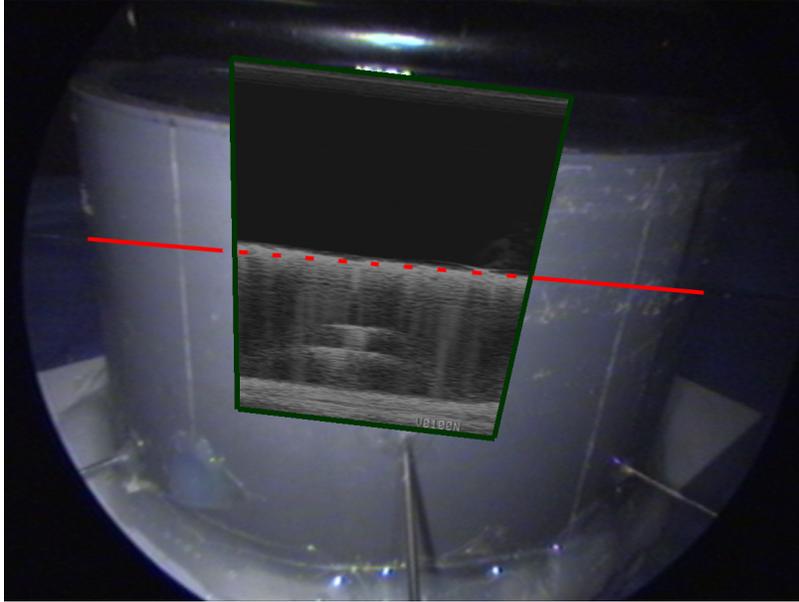
After acquiring 40 tip sensor poses and their corresponding lines that were automatically detected in the B-scan images, the calibration matrix was computed using the Levenberg-Marquardt optimizer. To determine the ultrasound calibration accuracy, a single EMT sensor with tip coordinates given in the EMT frame was submerged into the water bath. Its tip was segmented manually in 5 regions of the B-scan plane, which was repeated for 4 poses of the transducer differing from the ones used during calibration. The tip coordinates were transformed into the B-scan plane coordinates and compared to the segmented tip coordinates (scaled to mm). An RMS error of 1.69 mm with standard deviation of 0.51 mm and maximum error of 2.39 mm was obtained.

**Visualization of the Detected EMT Error** To visually inspect the overlay of the B-scan plane on the laparoscopic live video, we constructed a cylindrical phantom containing straight wires which extend through the walls of the phantom. It was filled with water of known temperature. Adjusting the pixel scaling factors to an adequate speed of sound, the B-scan plane was augmented, allowing the camera to view a wire on the augmented plane and its extension outside the phantom walls. A typical augmented laparoscope image can be seen in figure 7.

Whenever the occurrence of an error is determined, it is shown by drawing a red frame around the ultrasound plane. Otherwise the frame is drawn in green. An attempt to correct the error can be shown in yellow. The supplementary video demonstration (<http://campar.in.tum.de/files/publications/feuerste2007miccai.video.avi>) summarizes the results of all experiments and allows the observer to qualitatively evaluate the performance of automatic distortion estimation.

## Discussion

The flat tablet transmitter recently presented by Ascension may be an alternative to overcome field distortions, caused e.g. by the operating table. However, due to its lower excitation, in the same setup it performed worse than the mid-range transmitter for ultrasound calibration, resulting in errors of about 4-8 mm. Bigger sensors could



**Fig. 7.** Ultrasound plane augmented on the laparoscope video – red line added manually to show the extension of the straight wire, which matches its ultrasound image.

be used to improve the accuracy, but this would probably require bigger trocars. Using 1.3 mm sensors, the total diameter of the laparoscopic transducer is only 11.8 mm (including sterile cover), so it would still fit a regular 12 mm trocar.

Conditions in laparoscopic gastrointestinal surgery differ from those in e.g. neurosurgery or orthopedic surgery. Generally, it is sufficient to distinguish between structures of approximately 5 mm. For instance, canalicular structures such as bile ducts and vessels can be considered relevant and critical, if they are at least 5 mm in size, and lymph nodes are suspected to be tumorous, if they are at least 10 mm in diameter. Therefore, an error of about 2-3 mm, as obtained for the distortion free environment, is acceptable for clinical conditions. Errors of more than 2 mm are successfully detected by our system in most cases.

As an alternative to the error correction methods presented here, all possible transducer tip movements can be modeled relative to the shaft body to correct the tip sensor measurements. This model-based error correction method is, however, not the focus of this paper and is addressed in another work of our group. Its results will be described in a future publication.

In comparison to a model-based correction, the advantages of the vision-based correction method proposed here are that there are no assumptions necessary about the internal configuration of the transducer or any calibration thereof. But then, for accurate overlay purposes a calibration of the extrinsic and intrinsic camera parameters of the laparoscope should be readily available. The imposed constraints of transducer tip size or shape (cylindrical with a diameter of 8-12 mm) should be applicable to a wide range of transducer models without further modification.

## Conclusion

We presented new methods to detect EMT tracking errors online and partially correct these errors by a magneto-optic tracking setup. We improve the state of art [1, 6] for augmenting laparoscopic ultrasound images directly on the laparoscopic live images to give surgeons a better understanding of the spatial relationship between ultrasound and camera images. The laparoscopic ultrasound transducer tip is flexible. Therefore our method could be applied to a larger set of applications. We are using two attached sensors and hence are able to additionally provide a distrust level of the current EMT measurements. Therefore the system is able to automatically update and warn the surgical staff of possible inaccuracies.

## References

1. Leven, J., Burschka, D., Kumar, R., Zhang, G., Blumenkranz, S., Dai, X.D., Awad, M., Hager, G.D., Marohn, M., Choti, M., Hasser, C., Taylor, R.H.: Davinci canvas: A telerobotic surgical system with integrated, robot-assisted, laparoscopic ultrasound capability. In: Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI). Volume 3749 of Lecture Notes in Computer Science., Springer-Verlag (2005) 811–818
2. Ellsmere, J., Stoll, J., Wells, W., Kikinis, R., Vosburgh, K., Kane, R., Brooks, D., Rattner, D.: A new visualization technique for laparoscopic ultrasonography. *Surgery* **136**(1) (2004) 84–92
3. Harms, J., Feussner, H., Baumgartner, M., Schneider, A., Donhauser, M., Wessels, G.: Three-dimensional navigated laparoscopic ultrasonography. *Surgical Endoscopy* **15** (2001) 1459–1462
4. Krücker, J., Viswanathan, A., Borgert, J., Glossop, N., Yanga, Y., Wood, B.J.: An electro-magnetically tracked laparoscopic ultrasound for multi-modality minimally invasive surgery. In: Computer Assisted Radiology and Surgery, Berlin, Germany (2005) 746–751
5. Kleemann, M., Hildebrand, P., Birth, M., Bruch, H.P.: Laparoscopic ultrasound navigation in liver surgery: technical aspects and accuracy. *Surgical Endoscopy* **20**(5) (2006) 726–729
6. Nakamoto, M., Sato, Y., Miyamoto, M., Nakamjima, Y., Konishi, K., Shimada, M., Hashizume, M., Tamura, S.: 3d ultrasound system using a magneto-optic hybrid tracker for augmented reality visualization in laparoscopic liver surgery. In Dohi, T., Kikinis, R., eds.: Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI). Volume 2489 of Lecture Notes in Computer Science., Springer-Verlag (2002) 148–155
7. Hummel, J.B., Bax, M.R., Figl, M.L., Kang, Y., Maurer, Jr., C., Birkfellner, W.W., Bergmann, H., Shahidi, R.: Design and application of an assessment protocol for electromagnetic tracking systems. *Medical Physics* **32**(7) (2005) 2371–2379
8. Nafis, C., Jensen, V., Beauregard, L., Anderson, P.: Method for estimating dynamic em tracking accuracy of surgical navigation tools. In Cleary, K.R., Galloway, Jr., R.L., eds.: *Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display*. Volume 6141 of Proceedings of SPIE. (2006)
9. Kindratenko, V.V.: A survey of electromagnetic position tracker calibration techniques. *Virtual Reality: Research, Development, and Applications* **5**(3) (2000) 169–182
10. Birkfellner, W., Watzinger, F., Wanschitz, F., Enislidis, G., Truppe, M., Ewers, R., Bergmann, H.: Concepts and results in the development of a hybrid tracking system for cas. In Wells, I.W.M., Colchester, A.C.F., Delp, S.L., eds.: *Proceedings of the First International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Volume 1496 of Lecture Notes in Computer Science., Cambridge, MA, USA (1998) 343–351
11. Mucha, D., Kosmecki, B., Bier, J.: Plausibility check for error compensation in electromagnetic navigation in endoscopic sinus surgery. *International Journal of Computer Assisted Radiology and Surgery* **1**(Supplement 1) (2006) 316–318
12. Feuerstein, M.: *Augmented Reality in Laparoscopic Surgery - New Concepts for Intraoperative Multi-modal Imaging*. PhD thesis, Technische Universität München (2007)
13. Reichl, T.: *Online error correction for the tracking of laparoscopic ultrasound*. Master's thesis, Technische Universität München (2007)

14. Feuerstein, M., Reichl, T., Vogel, J., Schneider, A., Feussner, H., Navab, N.: Magneto-optic tracking of a flexible laparoscopic ultrasound transducer for laparoscope augmentation. In Ayache, N., Ourselin, S., Maeder, A., eds.: Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI). Volume 4791 of Lecture Notes in Computer Science., Brisbane, Australia, Springer-Verlag (2007) 458–466
15. Sielhorst, T., Feuerstein, M., Traub, J., Kutter, O., Navab, N.: Campar: A software framework guaranteeing quality for medical augmented reality. *International Journal of Computer Assisted Radiology and Surgery* **1**(Supplement 1) (2006) 29–30
16. Daniilidis, K.: Hand-eye calibration using dual quaternions. *International Journal of Robotics Research* **18** (1999) 286–298
17. Hartley, R., Zisserman, A.: *Multiple View Geometry in Computer Vision*. 2nd edn. Cambridge University Press (2003)
18. Yamaguchi, T., Nakamoto, M., Sato, Y., Konishi, K., Hashizume, M., Sugano, N., Yoshikawa, H., Tamura, S.: Development of a camera model and calibration procedure for oblique-viewing endoscopes. *Computer Aided Surgery* **9**(5) (2004) 203–214
19. Feuerstein, M., Mussack, T., Heining, S.M., Navab, N.: Intraoperative laparoscope augmentation for port placement and resection planning in minimally invasive liver resection. *IEEE Trans. Med. Imag.* **27**(3) (2008) 355–369
20. Treece, G.M., Gee, A.H., Prager, R.W., Cash, C.J.C., Berman, L.H.: High-definition freehand 3-d ultrasound. *Ultrasound in Medicine and Biology* **29**(4) (2003) 529–546
21. Climent, J., Marés, P.: Automatic instrument localization in laparoscopic surgery. *Electronic Letters on Computer Vision and Image Analysis* **4**(1) (2004) 21–31
22. Voros, S., Long, J.A., Cinquin, P.: Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder. In Larsen, R., Nielsen, M., Sporing, J., eds.: Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI). Volume 4190 of Lecture Notes in Computer Science., Springer (2006) 535–542
23. Hough, P.: Method and means for recognizing complex patterns. United States Patent 5063604 (1962)
24. Doignon, C., Nageotte, F., de Mathelin, M.: Segmentation and guidance of multiple rigid objects for intra-operative endoscopic vision. In: Workshop on Dynamic Vision, European Conference on Computer Vision. (2006)
25. Canny, J.: A computational approach to edge detection. *IEEE Trans. Pattern Anal. Machine Intell.* **8**(6) (1986) 679–698
26. Murray, R.M., Li, Z., Sastry, S.S.: *A Mathematical Introduction to Robotic Manipulation*. CRC Press (1994)
27. Ascension: microbird product brochure.  
<http://www.ascension-tech.com/docs/microBIRDbrochure021705.pdf> (2005)
28. Youden, W.J.: Index for rating diagnostic tests. *Cancer* **3**(1) (1950) 32–35
29. Langø, T.: *Ultrasound Guided Surgery: Image Processing and Navigation*. PhD thesis, Norwegian University of Science and Technology (2000)