Non-orthogonal tool/flange and robot/world calibration

Floris Ernst\(^1,2\)*
Lars Richter\(^1,3\)
Lars Matthäus\(^4\)
Volker Martens\(^1\)
Ralf Bruder\(^1\)
Alexander Schlaefer\(^3,5\)
Achim Schweikard\(^1\)

\(^1\)Institute for Robotics and Cognitive Systems, University of Lübeck, Ratzeburger Allee 160, 23538, Lübeck, Germany
\(^2\)HotSwap Lübeck GmbH, 23554, Lübeck, Germany
\(^3\)Graduate School for Computing in Medicine and Life Sciences, University of Lübeck, 23538, Lübeck, Germany
\(^4\)Eemagine Medical Imaging Solutions GmbH, 10243, Berlin, Germany
\(^5\)Medical Robotics Group, University of Lübeck, 23538, Lübeck, Germany

*Correspondence to: F. Ernst, Institute for Robotics and Cognitive Systems, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany. E-mail: ernst@rob.uni-luebeck.de

Abstract

**Background** For many robot-assisted medical applications, it is necessary to accurately compute the relation between the robot’s coordinate system and the coordinate system of a localisation or tracking device. Today, this is typically carried out using hand-eye calibration methods like those proposed by Tsai/Lenz or Daniilidis.

**Methods** We present a new method for simultaneous tool/flange and robot/world calibration by estimating a solution to the matrix equation \(AX = YB\). It is computed using a least-squares approach. Because real robots and localisation are all afflicted by errors, our approach allows for non-orthogonal matrices, partially compensating for imperfect calibration of the robot or localisation device. We also introduce a new method where full robot/world and partial tool/flange calibration is possible by using localisation devices providing less than six degrees of freedom (DOFs).

The methods are evaluated on simulation data and on real-world measurements from optical and magnetical tracking devices, volumetric ultrasound providing 3-DOF data, and a surface laser scanning device. We compare our methods with two classical approaches: the method by Tsai/Lenz and the method by Daniilidis.

**Results** In all experiments, the new algorithms outperform the classical methods in terms of translational accuracy by up to 80% and perform similarly in terms of rotational accuracy. Additionally, the methods are shown to be stable: the number of calibration stations used has far less influence on calibration quality than for the classical methods.

**Conclusion** Our work shows that the new method can be used for estimating the relationship between the robot’s and the localisation device’s coordinate systems. The new method can also be used for deficient systems providing only 3-DOF data, and it can be employed in real-time scenarios because of its speed. Copyright © 2012 John Wiley & Sons, Ltd.

**Keywords** hand-eye calibration; calibration and identification; medical robots and systems

Introduction

Tool/flange (or, equivalently, hand-eye or eye-in-hand calibration) and robot/world calibration is an important topic whenever it comes to a machine interacting with objects detected by some sort of camera, tracking system or other localisation device. In general, tool/flange and robot/world calibration boils down to solving a matrix equation of the type \(AX = YB\) (simultaneous tool/flange and robot/world calibration), where the matrices \(A\) and \(B\) are known and the matrices...
X and Y are unknown. Typically, A would be the pose matrix of a robot and B would be the position and orientation of a calibration object with respect to a tracking device. Consequently, the matrix X is the flange/tool matrix and Y is the robot/world matrix. This problem is usually tackled by taking measurements at multiple stations \((A, B)\), and eliminating one of the two unknown matrices to arrive at the simplified equation \(A^{-1}AX = XB^{-1}B\). The first works on solving this solution were devised independently by Shiu and Ahmad (27,28) and Tsai and Lenz (30,31). Both solutions make use of matrix algebra and the special properties of homogeneous matrices, separately determining the rotational and translational parts of the matrix X. An early review of these methods was given in (32), where the algorithm by Tsai and Lenz performed slightly better than the algorithm by Shiu and Ahmad. Other works, which also compute these components separately, are on the basis of quaternion algebra (7), screw motion analysis (5), the properties of the Euclidean group (21) and the solution of nonlinear equations (9,11,12). Li and Betis (17) compare their methods by using a geometric approach, a least-squares solution, and a nonlinear optimisation problem to the methods presented in (31) and (11). Again, the algorithm by Tsai and Lenz proved to be best and as good as their new nonlinear optimisation method. Interest was also sparked to simultaneously compute the rotational and translational parts of the matrix X. An early review of these methods was given in (32), where the algorithm by Tsai and Lenz performed slightly better than the algorithm by Shiu and Ahmad. Other works, which also compute these components separately, are on the basis of quaternion algebra (7), screw motion analysis (5), the properties of the Euclidean group (21) and the solution of nonlinear equations (9,11,12). Li and Betis (17) compare their methods by using a geometric approach, a least-squares solution, and a nonlinear optimisation problem to the methods presented in (31) and (11). Again, the algorithm by Tsai and Lenz proved to be best and as good as their new nonlinear optimisation method. Interest was also sparked to simultaneously compute the rotational and translational parts of the matrix X. This was carried out in (8) by a dual quaternion approach, in (36) and (35) by a one-stage iterative method minimising a non-linear cost function, in (22) by nonlinear minimisation of a sum of scalar products and in (29) by means of parameterisation of a stochastic model. In (38), the approach from (28) was extended to quaternion algebra (37) and used to simultaneously compute the parameters of the matrices X and Y. In (39), solving the hand-eye calibration problem is coupled with calibrating the robot’s kinematic parameters and the camera’s intrinsic parameters, resulting in a very large (up to 32 parameters) nonlinear optimisation problem. In (1), the authors refrain from using a calibration rig and attempt to solve the calibration problem by means of a structure-from-motion approach.

All those methods, however, have one thing in common: they expect that orthogonal homogeneous matrices \(^1\) \(X\) and \(Y\) can be found which optimally solve the relation \(AX = YB\). In reality, however, this is not necessarily the case: a typical robot will not be calibrated perfectly. Also, for new industrial robots deviations up to 2 – 3 mm can arise (2). Neither will an arbitrary tracking device deliver results that are exact. Optimal calibrated optical tracking systems will have a root mean square (RMS) error of 0.2 – 0.3 mm (34). For electromagnetic tracking systems, RMS errors of 1 – 1.5 mm can arise (10). The expected tracking accuracy for non-common tracking systems, used in medical applications, like 4D Ultrasound or 3D laser scanning systems is \(\approx 0.5\) and \(\approx 1 – 1.5\) mm, respectively.

In many medical applications, robots have become an important component and are required in many cases. For Transcranial Magnetic Stimulation (TMS), for instance, robotized systems are requested for precise and repeatable coil positioning on the head (16). A current development of a robot based system for TMS consists of an industrial robot for coil placement and a stereo-optical infrared tracking system for navigation (19). As for brain stimulation, only a small part of the robot workspace is used, local accuracy is more important. In particular, robot/world calibration algorithms should be adapted for robust real time calibration for robotized TMS (26).

Laser scanning systems are used for motion detection and patient registration in medical applications (14). In literature, it has been proposed to use a 3D laser scanning system to track human faces (15) and devices such as the GateRT system (VisionRT, London, UK) are used to detect a patient’s respiratory phase by tracking his/her chest. In general, laser scanning can also be used as a generic tracking device, especially suited for direct tracking, that is, when it is desirable that a target can be tracked without the need for attaching a marker. The laser scanning system directly scans the target and provides a 3D surface of the object. Even though accuracy of laser scanning systems is not as high as of stereo-optical infrared system, these systems can become a convenient alternative to standard tracking devices for medical robotic scenarios (25).

A relatively new tracking method, proposed in (3), makes use of a high-speed ultrasound station (GE Vivid7 Dimension, GE Healthcare, Little Chalfont, UK). It has been extended with tracking capabilities and can be used to locate a target volume by template matching or maximum intensity search up to 60 times per second. Combined with an industrial robot, the system is used for real time target tracking in radiotherapy (4). Furthermore, ultrasound is used for a laparoscopic liver surgery assistant system (18).

To take this into account, and to also allow for more freedom in the tracking modality used, we propose to extend tool/flange and robot/world calibration to allow non-orthogonal matrices, that is, to try to correct for system inaccuracies in the tool/flange and robot/world calibration matrices. It is also important that all calibration algorithms (except for (1), which requires a scene with more than one distinct feature) rely on matrices describing the full six degree of freedom (DOF) pose of a calibration rig in space. Consequently, tracking devices delivering information with less than six DOFs cannot be used.

In this work, we will present a new method with three different variations on the basis of a naïve least-squares solution of the equation system \(AX = YB\) (for simultaneous tool/flange and robot/world calibration). This approach combines the following important features:

- Because we deal with real-world localisation devices and imperfect robots, we allow the matrices \(X\) and \(Y\) to be non-orthogonal.
The methods compute simultaneously the rotational and translational parts of the matrices X and Y.

As calibration is not necessarily required for the full robot workspace for many medical applications, the algorithm aims for high local accuracy (i.e. required in (20)).

In the case of deficient tracking data, a partial solution is presented, that is, when the localisation device only provides translational data or does not provide full rotational data.

Methods

Tool/flange and robot/world calibration

First, the novel non-orthogonal calibration method is presented for tracking data providing full six DOF. Later, an adaptation of the method for tracking data with only partial tracking information having less than six DOF is described. The last variation deals with preconditioning of the equation system.

Non-orthogonal calibration

A naive approach to tool/flange and robot/world calibration is to look at the general relation

\[ R_{\mathbf{E}} \mathbf{E}_{\mathbf{M}} = R_{\mathbf{T}} \mathbf{T}_{\mathbf{M}}, \quad (1) \]

which is illustrated in Figure 1. Here, the matrices \( R_{\mathbf{E}} \) are the transforms from the robot’s end effector to the marker, and \( R_{\mathbf{T}} \), the transform from the robot’s base to the tracking system are unknown. To compute these matrices, \( n \) measurements for different robot poses are taken, resulting in \( n \) equations

\[ (R_{\mathbf{E}})_{i} \mathbf{E}_{\mathbf{M}} = R_{\mathbf{T}} (\mathbf{T}_{\mathbf{M}}), \quad i = 1, \ldots, n. \quad (2) \]

Typically, the robot’s end effector is moved to random points selected from inside a sphere with radius \( r \). Additionally, a random rotation of up to \( \pm d \) degrees in yaw, pitch and roll is added to the pose.

As a shorthand, let \( M_i = (R_{\mathbf{E}})_{i}, \quad X = \mathbf{E}_{\mathbf{M}}, \quad N_i = (\mathbf{T}_{\mathbf{M}}), \) and \( Y = \mathbf{R}_{\mathbf{T}} \). Consequently, Equation 2 can be written as

\[ M_i X - Y N_i = 0, \quad i = 1, \ldots, n. \quad (3) \]

By regarding the non-trivial elements of \( X \) and \( Y \) as components of a vector

\[ w = [x_{1,1}, x_{2,1}, \ldots, x_{3,4}, y_{1,1}, y_{2,1}, \ldots, y_{3,4}]^{T} \in \mathbb{R}^{24}, \]

Equation 3 can be combined into a system of linear equations

\[ A w = b, \quad (4) \]

where \( A \in \mathbb{R}^{12n \times 24} \) and \( b \in \mathbb{R}^{12n} \). More specifically,

\[ A = \begin{bmatrix} A_1 \vline A_2 \vline \ldots \vline A_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \vline b_2 \vline \ldots \vline b_n \end{bmatrix}, \quad (5) \]

where the matrices \( A_i \in \mathbb{R}^{12 \times 24} \) and the vectors \( b_i \in \mathbb{R}^{12} \) are determined as shown in Equation 6.

\[ A_i = \begin{bmatrix} \mathcal{R} (M_i)_{1,1} & \mathcal{R} (M_i)_{1,2} & \mathcal{R} (M_i)_{1,3} & Z_{3 \times 3}^{3} & \mathcal{R} (M_i)_{2,1} & \mathcal{R} (M_i)_{2,2} & \mathcal{R} (M_i)_{2,3} & Z_{3 \times 3}^{3} & E_{12} \mathcal{R} (M_i)_{3,1} & \mathcal{R} (M_i)_{3,2} & \mathcal{R} (M_i)_{3,3} & Z_{3 \times 3}^{3} & \mathcal{R} (M_i)_{4,1} & \mathcal{R} (M_i)_{4,2} & \mathcal{R} (M_i)_{4,3} & \mathcal{R} (M_i) \end{bmatrix}, \quad (6) \]

Here, \( \mathcal{R} (M_i) \in \mathbb{R}^{3 \times 3} \) is the rotational part of \( M_i \), \( T (M_i) \in \mathbb{R}^{3} \) is the translational part of \( M_i \), \( Z_{m \times n} \) is the \( m \times n \) zero matrix; and \( E_{k} \) is the \( k \times k \) identity matrix. This system can then be solved in a least squares sense by means of QR-factorisation. The resulting vector \( w \) now contains the non-trivial entries of the matrices \( X \) and \( Y \). Note that this method tries to find the optimal entries for \( X \) and \( Y \) in terms of the quadratic error, that is, it minimises

\[ \sum_{i=1}^{n} || M_i X - Y N_i ||_{F}^{2}, \quad (7) \]

whence the resulting matrices are not necessarily orthogonal. Here, \( || \cdot ||_{F} \) is the Frobenius norm. We call this method the \( QR24 \) calibration algorithm because 24

Figure 1. Principles of tool/flange and robot/world calibration: A marker \( M \) is attached to the robot’s end effector \( E \) and measured by the tracking system or an ultrasound station \( T \). Tool/flange and robot/world calibration is used to determine the unknown transforms \( R_{\mathbf{E}} \), the transform from the robot’s base to the tracking system, and \( \mathbf{E}_{\mathbf{M}} \), the transform from the end effector’s local system to the marker’s local system.
elements have to be estimated, and QR-factorisation is used for solving the equation system.

Note that, the QR24 algorithm solves for 24 parameters instead of 12 parameters used by standard calibration approaches. As we consider inaccuracies by tracking and robot in the calibration, shearing and scaling can occur in the matrices besides rotation and translation. Thus, the QR24 calibration problem is not over-parameterized when using non-orthogonality.

Calibration using only partial tracking information
In some situations, the tracking system might not deliver full 6-DOF poses. This will happen whenever the tracking system does not determine the pose of a rigid body but the location of one point in space, such as the position of one light-emitting diode (LED) or the position of maximum intensity in a US volume or the probe does not provide full rotational information. For instance, NDI’s 5-DOF magnetic sensors (Northern Digital Inc., Waterloo, Ontario, Canada) do not provide information about the roll angle. In such cases, the classical calibration algorithms such as the ones presented in (31) and (8) cannot be used. The QR24 calibration algorithm, however, can be adapted to deal with these partial measurements. A closer look at Equation 3 shows that it can be written as shown in Equation 8.

\[
\begin{bmatrix}
R[M_i] & T[M_i] \\
Z_{1\times 3} & 1
\end{bmatrix}
= \begin{bmatrix}
R[X] & T[X] \\
Z_{1\times 3} & 1
\end{bmatrix}
- \begin{bmatrix}
R[N_i] & T[N_i] \\
Z_{1\times 3} & 1
\end{bmatrix}
= Y_{1\times 4}.
\]

Using only the upper right part of this matrix equation,

\[
\]

shows that we can determine Y and T[X] by only using T[N] and M, that is, without using rotational information from the tracking system. Consequently, the matrices \( A_i \) and the vectors \( b_i \) from Equation 6 are changed as shown in Equation 10.

\[
A_i = \begin{bmatrix}
R[M_i](N_i)_{1,4} & R[M_i](N_i)_{2,4} & R[M_i](N_i)_{3,4} & R[M_i] - E_3
\end{bmatrix}
\]

\[
b_i = -T[M_i]
\]

This method is called the QR15 calibration algorithm because we can estimate only 15 elements.

Preconditioning the equation system
One drawback of the QR24 algorithm is its sensitivity with respect to the unit of the translational component of the pose matrices: all elements of X and Y are treated equally, that is, errors in the rotational part are treated with the same importance as errors in the translational part. Consequently, there will be a difference between the calibration results obtained using inches, metres or millimetres.

We have thus extended the QR24 algorithm to allow for an arbitrary scaling factor of the translation. We have performed all experiments with translational values given in mm. Additionally, we used a scaling factor of 0.001 in a second run of the QR24 algorithm to demonstrate the effect of preconditioning (these results are denoted with QR24m).

Orthonormalisation
Because—as described earlier—X and Y do not need to be orthogonal during computation, poses calibrated using these matrices must be orthonormalised for the final result. Otherwise, rigid body transformations by using the found calibration matrices would not be applicable.

To avoid the bias incurred by Gram-Schmidt orthonormalisation, we propose to use orthonormalisation by means of singular value decomposition (SVD). To this end, let us assume that N is a pose measured by the tracking system, and we want to know this pose in robot coordinates. We then compute YN, a non-orthogonal matrix, and subsequently determine the ‘closest’ orthogonal matrix, \((YN)^{-1}\). Let \(USV^T = YN\) be the SVD of YN. Then \((YN)^{-1}\) can be computed as \(UV^T\). \(^2\)

Computing calibration errors
Using the calibration matrices \( R_{\mathcal{T}} \) and \( E_{\mathcal{M}} \), computed by any calibration method, we can determine the calibration quality by looking at Equation 1. In the situation of perfect calibration, this equation can be transformed to

\[
(E_{\mathcal{M}})^{-1}(R_{\mathcal{T}})^{-1}R_{\mathcal{T}}^T E_{\mathcal{M}} = E_4.
\]

Note that in reality, equality in Equation 1 will not hold, and, consequently, deriving Equation 11 is not frame-invariant: there are eight possible ways of expressing this relation. Because all frames of reference are equally valid, it has been proposed in literature (29) to use the average of the errors from all possible frames of reference. But when the tracking system does not deliver full 6-DOF measurements, the matrices \( E_{\mathcal{M}} \) and \( R_{\mathcal{T}} \) cannot be

\(^2\)Note that we do not actually perform orthonormalisation of YN but of \(R[YN]\). See also footnote 1.
inverted. Consequently, we have to use Equation 11 for 6-DOF data and the following equation for 3-DOF data to ensure comparable results:

$$
(\mathbf{R}^E \mathbf{I}_6)^{-1} \mathbf{R}^T \mathbf{T} \mathbf{I}_3 = \mathbf{T} \mathbf{I}_3 = \mathbf{Z}_{3 \times 1}.
$$

(12)

To determine calibration quality, we introduce the following measures:

- Translational accuracy and
- Rotational accuracy.

Let $A = \left( (\mathbf{E} \mathbf{I}_3)^{-1} (\mathbf{R}^E \mathbf{I}_6)^{-1} \mathbf{R}^T \mathbf{T} \mathbf{I}_3 \right)$. Then its translational error is defined as

$$
e_{\text{trans}}[A] = \sqrt{a_1^2 + a_2^2 + a_3^2}.
$$

(13)

To determine the rotational accuracy, let $(k, \theta)$ be the axis-angle representation of $\mathbf{R}[A]$. Then the rotation error is defined as

$$
e_{\text{rot}}[A] = |\theta|.
$$

(14)

In this metric (as was carried out in (29), where more details can be found), we regard the axis of rotation as irrelevant.

Clearly, this approach is only possible if both calibration matrices $\mathbf{R}^T$ and $\mathbf{E} \mathbf{I}_3$ as well as 6-DOF tracking information is available. If the tracking system only delivers 3-DOF or 5-DOF information, we cannot compute the rotation errors, and the formula for computation of the translation error is changed to

$$
e_{\text{trans}}[a] = \sqrt{a_1^2 + a_2^2 + a_3^2},
$$

(15)

where

$$
a = \left( (\mathbf{R}^E \mathbf{I}_6)^{-1} \mathbf{R}^T \mathbf{T} \mathbf{I}_3 \right) - \mathbf{T} \mathbf{I}_3
$$

(16)

### Experimental validation—collecting data

To evaluate the accuracy and robustness of the proposed calibration methods, we have performed multiple calibration tests. In the first test, simulated data were used. In the second test, calibration was performed using four real world set-ups as shown in Figure 2 and using the algorithm of Tsai and Lenz, Daniilidis’ dual quaternion method as well as the new methods proposed.

The robot was controlled using our robot client/server framework, presented in (24).

**Computer-generated data**

To test the calibration algorithms, random orthogonal matrices $X$ and $Y$ were generated. To create realistic values, the elements of $\mathbf{T}[X]$ were selected randomly from $[-20, 20]$ mm, and the elements of $\mathbf{T}[Y]$ from $\pm [500, 2000]$ mm. Using these matrices, 500 completely random orthogonal robot poses $M_i$ were generated and, subsequently, the corresponding tracking matrices $\mathbf{N}_i$ were computed. To simulate the actual behaviour of an optical tracking system, a tool with four markers was defined as

$$
\text{tool} = \begin{bmatrix}
0 & 50 & 50 & 0 \\
0 & 0 & 50 & 50 \\
0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1
\end{bmatrix}.
$$

To determine the distorted tracking matrices $\tilde{\mathbf{N}}_i$, the tool was moved to the pose described by $\mathbf{N}_i$, and the matrix $\tilde{\mathbf{N}}_i$ was computed as the best fit computed by Horn’s algorithm without scaling (13), that is,

$$
\tilde{\mathbf{N}}_i = \text{horn} \left( \mathbf{N}_i \text{tool} + \Sigma \right).
$$

Here,

$$
\Sigma_i = \begin{bmatrix}
\hat{s}_{i1.1} & \hat{s}_{i1.2} & \hat{s}_{i1.3} & \hat{s}_{i1.4} \\
\hat{s}_{i2.1} & \hat{s}_{i2.2} & \hat{s}_{i2.3} & \hat{s}_{i2.4} \\
\hat{s}_{i3.1} & \hat{s}_{i3.2} & \hat{s}_{i3.3} & \hat{s}_{i3.4} \\
0 & 0 & 0 & 0
\end{bmatrix}
$$

and $\hat{s}_{i,k}$ were drawn from the Gaussian process with $\sigma = 0.05$ mm.

The robot matrices were distorted in a different way: their rotational part was disturbed by multiplication with a random rotation matrix around an arbitrary axis $a_i$. The rotation angles $\theta_i$ were selected from the Gaussian distribution with $\sigma = 0.05^\circ$, and their translational part was disturbed by additive Gaussian noise:

$$
\tilde{M}_i = \begin{bmatrix}
\mathbf{R}[M_i] \cdot \text{rot}_m (\theta_i) & \mathbf{T}[M_i] + \Sigma \\
0 & 1
\end{bmatrix},
$$

where

$$
\Sigma_i = [\hat{s}_{i1.1, \hat{s}_{i1.2, \hat{s}_{i1.3}}}]^T
$$

and $\hat{s}_{i,k}$ were drawn from the Gaussian process with $\sigma = 0.05$ mm.

**Data from tracking systems**

Real data were generated using two tracking systems: an optical device and a magnetic tracking system (NDI’s Polaris Spectra and Aurora systems, Northern Digital Inc., Waterloo, Ontario, Canada). Both systems are capable of delivering full 6-DOF tracking data.

**Optical tracking**

The spectra system was mounted to a KUKA KR 16 robot (KUKA Roboter GmbH, Augsburg, Germany), and an active 4-LED marker was attached to the end effector of an Adept Viper s850 robot (Adept Technology, Inc., Livermore, California) (Figure 2(a)). The marker and the camera were positioned such that the marker’s tracked position was in the centre of the camera’s working volume (in this case, at $x = 0$, $y = 0$ and $z = -1500$ mm). Subsequently, the marker was moved to 500 poses around the initial pose $P_0$. These poses were determined by
\[ P_i = P_0T(t)R_x(\theta_1)R_y(\theta_2)R_z(\theta_3), \quad i = 1, \ldots, 500, \]  

where \( R_a(\theta) \) is a rotation matrix around the a-axis by \( \theta \), and \( T \) is a translation matrix with translation vector \( t \). 

The values of the magnetic test were generated similarly: the field generator was placed on a plastic table, and a 6-DOF sensor coil was attached to an acrylic plate (1 m long) that was attached to the robot’s end effector (Figure 2(b)). Using this set-up, 500 pose pairs were collected with \( r \) set to 75 mm and \( \theta_{\text{max}} \) set to 5°. Again, 100 measurements were averaged at each calibration station.

**Other modalities**

To determine the applicability of the calibration algorithms to other, less common, imaging/tracking modalities, we have also collected data by using a 3D laser scanning system (LAP GALAXY) and a 4D ultrasound station (GE Vivid7 Dimension).

**Laser scanning**

In our case, we used the LAP GALAXY system (LAP Laser Applikationen GmbH, Lüneburg, Germany) and attached a special calibration tool to the robot. This tool is then scanned with the laser scanning system. To determine its pose, the calibration tool is measured relatively to a reference image. The pose matrices determined can be used with the proposed calibration methods (23).

Using the laser scanner is not as simple as using an off-the-shelf tracking system. It is time consuming to collect a large set of data points for calibration. This is because a special calibration tool has to be used as a marker, and the laser scanning system does not directly provide the pose matrix of the tool. To determine its pose, we require a reference image of the calibration tool. Then the pose matrix, relating the actual position and orientation of the scanned tool to the reference image, can be computed with, for example, an iterative closest point (ICP) algorithm (6). This indirect approach is necessary because the laser scanning system only provides a point cloud of the measured surface. Figure 3 shows the MATLAB graphical user interface (The Mathworks, Natick, MA, USA) used for landmark-based preregistration and ICP registration as well as a typical result.

Consequently, we have used a set of only \( n = 50 \) randomly distributed data points to test calibration of the laser scanning system. Figure 2(c) shows the set-up used to collect the data.

**Ultrasound**

We have used a simple lead ball phantom, moved by the robot, to calibrate the robot’s coordinate system to the coordinate system of the ultrasonic head. Note that this
approach will result in 3-DOF data. Using the same approach, it would also be possible to calibrate the probe with respect to an optical marker tracked by an optical tracking system. It must be noted that, because of varying speeds of sound in different tissues, calibration in a water tank must result in a non-unity scaling factor (this factor should be about 0.96, because the speed of sound in water is 1480 m/s and, on average, 1540 m/s in human tissue).

Tracking the lead ball was performed using a maximum intensity algorithm running directly on the ultrasound machine. The data (500 pose pairs) were acquired in the same way as for the optical and magnetic tracking systems. Because of the limited field of view of the ultrasound probe, however, $r$ was set to 15 mm, and $\theta_{\text{max}}$ was set to 3°. Figure 2(d) shows the set-up used.

To allow sub-voxel tracking, not the voxel with maximum intensity but the position in space corresponding to the centre of mass of the tracked object was used. This position was determined by fitting a quadratic polynomial to the intensities of the voxels around the brightest voxel and using this polynomial's extremal value as tracked position.

By using this method, an accuracy of better than 0.5 mm can be achieved throughout the volume.

Results

Evaluation speed and implementation

Evaluation of the algorithms was carried out on a standard business notebook (Intel Core i5-2540 M CPU with 8 GB of RAM, running Windows 7 × 64). The algorithms were all implemented in MATLAB R2011a. Sample evaluation times for the algorithms are less than 0.1 s for the QR24, QR15 and Tsai–Lenz methods and below 0.5 s for the dual quaternion method. The methods by Tsai–Lenz and the dual quaternion method were taken from (33).

Calibration errors

To compare the quality of the calibration algorithms, calibration was performed using the poses $P_1, \ldots, P_n$ for $n = 5, \ldots, 250$. For all such cases, the determined calibration matrices were evaluated using the remaining poses $P_{251}, \ldots, P_{500}$. The rotation and translation errors were computed for each testing pose, and the average of all 250 testing poses was determined. Figure 4 shows these...
averages for the optical set-up (top), the magnetic set-up (centre) and the simulated data (bottom).

We can clearly see that using more than about 50 poses hardly influences the translational error of the QR24 and QR15 algorithms in all three experiments. Even more, their errors are lowest by far (approximately 0.09 mm for the optical system, 0.27 mm for the magnetic system and 0.123 mm for the simulated data). On the synthetic data, we can see that the resulting error comes very close to the error obtained using the correct matrices. It remains within 1%. When looking at the rotational errors, however, the picture is somewhat different: the dual quaternion algorithm performs best on the real data but all algorithms perform similarly on the simulated data. The difference, however, is not as pronounced as for the translational errors. When using the scaled method (i.e. $QR_{24M}$), we can see that we gain rotational accuracy at the expense of higher translational errors, with results comparable with the dual quaternion algorithm. Very interesting, however, is that calibrating using more pose pairs does not necessarily result in improved results for all algorithms. This can be seen most clearly for the translational error of the Tsai–Lenz algorithm when using data from optical tracking.

Additionally, Table 1 shows the statistics of the errors for the optical and magnetic tests obtained using the optimal matrices, that is, those matrices resulting in minimal RMS error on the test poses. In Figure 5, the results of the optical tracking system are visualised. Again, the QR-based algorithms clearly outperform the Tsai–Lenz method.

Other modalities

For the other two modalities described before, laser scanning and volumetric ultrasound, the proposed calibration algorithms perform very satisfactorily.

Laser scanner: For the data collected with our laser scanning system, the results are given in Table 2(a) and Figures 6 and 7. It is clear that, in terms of translational accuracy, the proposed methods strongly outperform the algorithms by Tsai and Lenz and the Dual Quaternion method: the improvement in

---

Figure 4. Calibration errors for the QR, Tsai–Lenz, and dual quaternion algorithms. The algorithms used $n = 5, \ldots, 250$ poses to compute the calibration matrices that were then tested on 250 other poses. The top graphs show the mean errors by using optical tracking, the centre graphs show the mean errors by using magnetic tracking and the bottom graphs show the mean errors by using simulated data. In the bottom graphs, the turquoise lines show the calibration errors obtained using the correct matrices $X$ and $Y$. Note that, for the translation errors, the results for the QR24 and QR15 algorithms are nearly identical, whence the corresponding lines coincide.
The median error is around 80%. The rotation errors are similar for all systems, their median values ranging from 0.81 to 0.95°. One thing, however, is interesting: preconditioning the QR24 algorithm massively decreases translational accuracy while only slightly reducing rotational errors.

In general, given the laser scanner’s accuracy of approximately 0.5 mm and the average accuracy of the ICP matching of 1.1 mm, a median translational calibration error of 1.3–1.4 mm is convincing.

### Ultrasound

By using the pose pairs collected, it was possible to compute the matrices $Y$ and $T[X]$. Because the data recorded from the ultrasound station features only three DOF, the only calibration method we can use is the QR15 algorithm. Again, we computed the calibration matrices by using $n = 5, \ldots, 250$ pose pairs and computed the translational errors on the other 250 pose pairs. We can see from Figure 8 that the error’s start is below 1 mm even when using only five pose pairs and drops to below 0.6 mm when using more than 25 pose pairs. The error eventually stabilises around 0.52 mm. This result is well in line with the expected tracking accuracy of approximately 0.5 mm. Additionally, Figure 8 also shows the distribution of the calibration error amongst the 250 pose pairs used for testing when the median error is around 80%. The rotation errors are similar for all systems, their median values ranging from 0.81 to 0.95°. One thing, however, is interesting: preconditioning the QR24 algorithm massively decreases translational accuracy while only slightly reducing rotational errors.

In general, given the laser scanner’s accuracy of approximately 0.5 mm and the average accuracy of the ICP matching of 1.1 mm, a median translational calibration error of 1.3–1.4 mm is convincing.

### Ultrasound

By using the pose pairs collected, it was possible to compute the matrices $Y$ and $T[X]$. Because the data recorded from the ultrasound station features only three DOF, the only calibration method we can use is the QR15 algorithm. Again, we computed the calibration matrices by using $n = 5, \ldots, 250$ pose pairs and computed the translational errors on the other 250 pose pairs. We can see from Figure 8 that the error’s start is below 1 mm even when using only five pose pairs and drops to below 0.6 mm when using more than 25 pose pairs. The error eventually stabilises around 0.52 mm. This result is well in line with the expected tracking accuracy of approximately 0.5 mm. Additionally, Figure 8 also shows the distribution of the calibration error amongst the 250 pose pairs used for testing when the median error is around 80%. The rotation errors are similar for all systems, their median values ranging from 0.81 to 0.95°. One thing, however, is interesting: preconditioning the QR24 algorithm massively decreases translational accuracy while only slightly reducing rotational errors.

In general, given the laser scanner’s accuracy of approximately 0.5 mm and the average accuracy of the ICP matching of 1.1 mm, a median translational calibration error of 1.3–1.4 mm is convincing.

### Table 1. Error statistics of the calibration algorithms using optimal matrices on the test data

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Min</th>
<th>25th p</th>
<th>Median</th>
<th>75th p</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translation error (mm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tsai−Lenz</td>
<td>0.0350</td>
<td>0.2112</td>
<td>0.2950</td>
<td>0.4101</td>
<td>0.9289</td>
</tr>
<tr>
<td>Dual quaternion</td>
<td>0.0156</td>
<td>0.1610</td>
<td>0.2239</td>
<td>0.3625</td>
<td>0.8305</td>
</tr>
<tr>
<td>QR24</td>
<td>0.0388</td>
<td>0.1685</td>
<td>0.2255</td>
<td>0.3253</td>
<td>0.6833</td>
</tr>
<tr>
<td>QR24M</td>
<td>0.0223</td>
<td>0.0923</td>
<td>0.1317</td>
<td>0.1792</td>
<td>0.3095</td>
</tr>
<tr>
<td>QR15</td>
<td>0.0223</td>
<td>0.0920</td>
<td>0.1314</td>
<td>0.1790</td>
<td>0.3089</td>
</tr>
<tr>
<td>Rotation error (°)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tsai−Lenz</td>
<td>0.0042</td>
<td>0.0625</td>
<td>0.0979</td>
<td>0.1312</td>
<td>0.3758</td>
</tr>
<tr>
<td>Dual quaternion</td>
<td>0.0072</td>
<td>0.0586</td>
<td>0.0857</td>
<td>0.1187</td>
<td>0.3832</td>
</tr>
<tr>
<td>QR24</td>
<td>0.0117</td>
<td>0.0666</td>
<td>0.0961</td>
<td>0.1342</td>
<td>0.3860</td>
</tr>
<tr>
<td>QR24M</td>
<td>0.0130</td>
<td>0.0627</td>
<td>0.0940</td>
<td>0.1270</td>
<td>0.4069</td>
</tr>
<tr>
<td>QR15</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

The numbers shown are minimum, 25th percentile, median, 75th percentile and maximum. Minimal values for each column are marked in bold. (a) shows the data for the optical system, (b) for the magnetic system.

### Figure 5. These plots show the translation (left) and rotation (right) errors on the optical test data set by using optimal calibration matrices. The number of pose pairs used was $n = 5$ for the Tsai−Lenz algorithm, $n = 221$ for the dual quaternion algorithm, $n = 14$ for the QR24M algorithm and $n = 229$ for the QR24 and QR15 algorithms. The corresponding numbers are given in Table 1(a).
optimal calibration matrix (i.e. the one resulting in the lowest average error) is applied. We can see here that the 75th percentile lies at 0.70 mm, and an error of 1.2 mm is never exceeded.

As expected, a typical calibration matrix does exhibit non-unity scaling factors (for \( n = 174 \) pose pairs, the factors are 0.970, 0.969 and 0.978), and the matrix is not completely orthogonal (the angles between the axes...
of the coordinate system spanned by the matrix are 90.56°, 90.47° and 90.24°). This confirms that, in the case of a more complex tracking modality, a more robust calibration algorithm is required. For comparison, when looking at the optimal calibration results (QR24 algorithm) of the spectra system, the scaling factors are 1.0010, 1.0018 and 0.9967, and the matrix angles are 90.08°, 90.39° and 90.11°.

Convergence properties

To determine the convergence properties of the different algorithms, the synthetic data were evaluated again. It is clear from Figure 4, bottom row, that all algorithms seem to converge to the error obtained using the correct calibration matrices (the turquoise lines). Additionally, it was possible to determine if and how fast the algorithms converge to the correct matrices X and Y, which are known in this case. The results are shown in Figure 9. Again, we can see that the QR24 and QR15 algorithms converge quickly, stabilising at a translational error of about 0.02 mm on X and Y. The rotational errors show a similar picture: an accuracy of less than 0.01° is achieved on both matrices. The dual quaternion and Tsai–Lenz algorithms, however, do not show proper convergence. The errors may increase strongly when more poses are used in the calibration routine. The QR24M algorithm, however, results in higher errors on the matrix Y and similar errors on the matrix X when compared with the QR24 algorithm.

Increased workspace sizes

For many medical robotic applications—as already mentioned—the selected workspace sizes are in the desired range. However, to show the quality of our algorithms, we have performed additional calibration experiments with the Polaris Spectra system using
increased workspace sizes. (Note that the Aurora system’s workspace is only marginally larger than the workspace used in the first experiment.) We used the following five experiments with changed workspace sizes:

1. \(r = 100\) mm and \(\theta = 10^\circ\),
2. \(r = 200\) mm and \(\theta = 10^\circ\),
3. \(r = 300\) mm and \(\theta = 20^\circ\),
4. \(r = 400\) mm and \(\theta = 20^\circ\) and
5. \(r = 500\) mm and \(\theta = 30^\circ\),

with radius of the workspace \(r\) and a maximal rotation angle \(\theta\).

In all cases, the robot was ordered to acquire 1000 samples, which were successfully recorded in the first two experiments. Because of limitations of feasible joint angles, in the third experiment, 971 measurements, in the fourth experiment, 904 measurements and in the fifth experiment, 784 measurements were collected.

Figure 10 displays the found calibration errors for this experiment. Interestingly, the general calibration accuracy remained with increased workspace size. Although the errors slightly increased for all calibration methods, the presented calibration methods (QR24 and QR15) were still best.

Discussion

We have presented and evaluated a novel method for tool/flange and robot/world calibration. The method has been validated using simulated data, measured data from optical and magnetic tracking systems, and using data from surface laser scanning and volumetric ultrasound. The results were compared with calibration obtained with standard algorithms such as those by Tsai—Lenz and Daniilidis.

We have shown that, when using synthetic data, all methods perform very well with errors below 0.25 mm and 0.1°. This shows that all methods are suitable for typical calibration tasks. In our specific set-up, however, where off-the-shelf tracking systems and robots as well as somewhat uncommon devices for localisation were used, we found that the accuracy of our method was typically 50% better than the accuracy of the standard algorithms. It has to be noted, however, that the new method can only be used when

- it is acceptable or required to allow non-orthogonality in the calibration matrix/matrices, and
- the calibrated system is used in the same spatial area where calibration was performed.

Additionally, we have seen in all tests that calibration using our method comes very close to the accuracy we can expect from the systems: the accuracies of the systems are approximately 0.1 mm (optical tracking\(^3\)), 0.5 mm (magnetic tracking and volumetric ultrasound) and 1 – 2 mm (laser scanner data), and the median calibration accuracies are 0.13 mm, 0.26 mm, 0.48 mm and 1.35 mm, respectively.

We have shown that our method is well-suited to handle deficient tracking data, that is, to systems that do not deliver full 6-DOF pose data (such as volumetric ultrasound or 5-DOF magnetic sensors). The new calibration algorithm can be used to determine the robot/world calibration matrix even in this case, and it has been shown to be equivalent to the full method providing simultaneous tool/flange and robot/world calibration.

\(^3\)Note that we used a relatively small working volume and averaged multiple measurements to achieve this spatial accuracy. Typical accuracy of the device across its full measurement volume is approximately 0.3 mm RMS.
Concluding, we can say that for many medical applications, where calibration is not necessarily required for the full workspace of the robot and the need for high local accuracy is paramount, our new methods provide a viable alternative to the classical approaches.

References

5. Chen HH. A screw motion approach to uniqueness analysis of dimension.
