

# Hand-Eye Calibration without Hand Orientation Measurement Using Minimal Solution

Zuzana Kukelova<sup>1</sup> Jan Heller<sup>1</sup> Tomas Pajdla<sup>2</sup>

<sup>1</sup>Center for Machine Perception, Department of Cybernetics  
Faculty of Elec. Eng., Czech Technical University in Prague  
166 27 Prague 6, Technicka 2, Czech Republic

<sup>2</sup>Neovision, s.r.o., Barrandova 409  
143 00 Prague 4, Czech Republic

**Abstract.** In this paper we solve the problem of estimating the relative pose between a robot’s gripper and a camera mounted rigidly on the gripper in situations where the rotation of the gripper w.r.t. the robot global coordinate system is not known. It is a variation of the so called hand-eye calibration problem. We formulate it as a problem of seven equations in seven unknowns and solve it using the Gröbner basis method for solving systems of polynomial equations. This enables us to calibrate from the minimal number of two relative movements and to provide the first exact algebraic solution to the problem. Further, we describe a method for selecting the geometrically correct solution among the algebraically correct ones computed by the solver. In contrast to the previous iterative methods, our solution works without any initial estimate and has no problems with error accumulation. Finally, by evaluating our algorithm on both synthetic and real scene data we demonstrate that it is fast, noise resistant, and numerically stable.

## 1 Introduction

The problem of estimating the relative position and orientation of a robot gripper and a camera mounted rigidly on the gripper, known as *hand-eye calibration problem*, has been studied extensively in the past [23, 24, 17, 2, 25, 9]. This problem arises in wide range of applications not only in robotics but also in automotive or medical industries.

The standard formulation of this problem leads to solving a system of equations of the form

$$\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{B}, \quad (1)$$

where known  $\mathbf{A}$  and  $\mathbf{B}$  and unknown  $\mathbf{X}$  are homogeneous transformation matrices of the form

$$\begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix}, \quad (2)$$

with  $3 \times 3$  rotation matrix  $\mathbf{R} \in SO(3)$  and  $3 \times 1$  translation vector  $\mathbf{t} \in \mathbb{R}^3$ . It has been shown in [24] that at least two motions with non-parallel rotation axes are

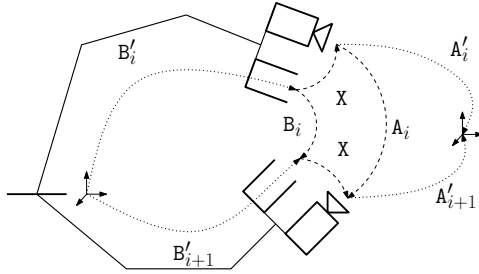


Fig. 1: The relative movement of the camera-gripper rig

required to solve hand-eye calibration problem. In practice, several motion are performed and the overconstrained system

$$A_i X = X B_i, \quad i = 1, \dots, n \quad (3)$$

is solved as a minimization problem, with every method trying to minimize a different error criteria. The existing methods can be divided into three groups.

The first group of methods [23, 24, 17, 2, 15] solves System 3 by decomposing it into matrix equations depending only on rotations and vector equations depending both on rotations and translations. In this way the methods decouple the rotation from the translation and solve for them separately, *i.e.*, first for the rotation and then for the translation. The drawback of such an approach is that the rotation estimation errors propagate to the translation errors. To address the problem of error propagation several methods for simultaneous estimation of rotation and translation appeared [25, 9, 5, 21]. These methods search for the the unknown transformation  $X$  by solving the overconstrained System 3 using different linear or non-linear minimization methods. The methods mostly differ in the error function which is minimized and in the used minimization method. Methods that use iterative optimization techniques suffer from the inherent problems of iterative algorithms, *i.e.*, problems with convergence and the necessity of a good initial estimate of  $X$ . In [19] authors proposed a method that uses tracked image points rather than matrices  $A_i$ . Recently, another group of methods appeared [8, 7, 16]. These methods use image correspondences instead of matrices  $A_i$  and employ global optimization to minimize different error functions in  $L_\infty$ -norm.

In this paper we are concerned with a variation of hand-eye calibration problem that has been scarcely addressed in the literature so far—hand-eye calibration with unknown hand rotation. This problem arises when the robot is not calibrated or the information from the robot is not available. In these situation one has to measure the robot's pose by an external measurement device. In many cases such a measurement device is not able to measure the whole pose, but only the translational part of it, since translation is much easier to measure than rotation. Without the hand rotation measurements none of the previously discussed methods can be used. A method presented in [25] addresses this prob-

lem by nonlinear optimization and estimates simultaneously both rotational and translational parts. However, it requires a good initial estimate of  $\mathbf{X}$ .

In case of two relative motions, we solve this problem by formulating it as a system of seven equations in seven unknowns and solving it using the Gröbner basis method for solving systems of polynomial equations. This provides an exact algebraic solution and has none of the problems of the former numerical minimization methods, *i.e.*, problems with convergence or the necessity of having a good initial estimate. In case of three or more motions, we use a residual function to select an initial solution among the candidates provided by the Gröbner basis method to initialize the method of [25]. By evaluating our solution on both synthetic and real scene data, we demonstrate that it is efficient, fast, and numerically stable. Further, we show that in case of more than two motions it provides a good estimate for nonlinear optimization.

## 2 Problem Formulation

First, let us consider the classical hand-eye calibration problem. The goal is to estimate the relative pose, *i.e.*, the rotation and the translation of the camera w.r.t. the gripper, see Figure 1. We will describe this transformation by the homogeneous transformation matrix

$$\mathbf{X} = \begin{bmatrix} \mathbf{R}_X & \mathbf{t}_X \\ \mathbf{0}^\top & 1 \end{bmatrix}, \quad (4)$$

where  $\mathbf{R}_X \in SO(3)$  is the unknown rotation from the camera to the gripper and  $\mathbf{t}_X \in \mathbb{R}^3$  the unknown translation.

Let us consider the  $i^{\text{th}}$  pose of the robot and denote the transformation matrix from the camera to the world coordinate system by  $\mathbf{A}'_i$  and the transformation matrix from another coordinate system in the world—usually placed in the robot’s base—to the robot’s gripper by  $\mathbf{B}'_i$ , see Figure 1. Camera’s transformations  $\mathbf{A}'_i$  can be obtained using the well known absolute pose solvers [6, 14] and transformations  $\mathbf{B}'_i$  from the robot’s positioning software.

Figure 1 shows that by knowing two poses of the robot we can get  $\mathbf{X}$  from the following equation

$$\mathbf{A}_i \mathbf{X} = \mathbf{X} \mathbf{B}_i, \quad (5)$$

where  $\mathbf{A}_i = \mathbf{A}'_i{}^{-1} \mathbf{A}'_{i+1}$  and  $\mathbf{B}_i = \mathbf{B}'_{i+1} \mathbf{B}'_i{}^{-1}$  are homogeneous transformation matrices representing the respective relative movements. Equation 5 can be decomposed into a matrix and a vector equation

$$\mathbf{R}_{\mathbf{A}_i} \mathbf{R}_X = \mathbf{R}_X \mathbf{R}_{\mathbf{B}_i}, \quad (6)$$

$$\mathbf{R}_{\mathbf{A}_i} \mathbf{t}_X + \mathbf{t}_A = \mathbf{R}_X \mathbf{t}_{\mathbf{B}_i} + \mathbf{t}_X. \quad (7)$$

At least two motions with non-parallel rotation axes are required to solve this system of equations. With two or more motions known, we obtain an over-constrained system of polynomial equations.

In situations where one does not have the information from the robot’s positioning software or the robot is not precisely calibrated transformations  $\mathbf{B}'_i$  are not readily known. To recover them, one has to use some external measurement equipment. In this paper we are interested in situations where such a measurement device does not allow to recover the whole pose of the robotic gripper, but only its translational part.

Typically, the external measurement devices are able to recover absolute gripper’s positions  $\mathbf{t}'_{\mathbf{B}}$  w.r.t. robot’s base. However, in Equation 7 relative translations  $\mathbf{t}_{\mathbf{B}}$  appear. In order to compute the relative translations  $\mathbf{t}_{\mathbf{B}}$  there has to be at least one position where the full pose of the robot can be recovered, *i.e.*, where the rotation  $\mathbf{R}'_{\mathbf{B}}$  is known as well. Even for an uncalibrated robot, the robot’s home position can be used as such *a priori* known pose. By constructing the relative movements in such a way as to always end in a position with a known rotation  $\mathbf{R}'_{\mathbf{B}}$ , relative translations  $\mathbf{t}_{\mathbf{B}}$  can be recovered. Since the positions with *a priori* known poses are usually hard to come by, it is advantageous for a method to be able to calibrate from a minimal number of movements possible.

### 3 Minimal Problem

First, let us suppose that we can measure two gripper’s relative translations  $\mathbf{t}_{\mathbf{B}_i}$  and  $\mathbf{t}_{\mathbf{B}_j}$  and two respective relative camera motions  $\mathbf{A}_i$  and  $\mathbf{A}_j$ . Now, let us note that the vector Equation 7 does not contain the unknown gripper’s rotations  $\mathbf{R}_{\mathbf{B}_i}$ . By parametrizing the rotation  $\mathbf{R}_{\mathbf{X}}$  by the unit quaternion  $q = a + bi + cj + dk$  as

$$\mathbf{R}_{\mathbf{X}} \equiv \mathbf{R}_{\mathbf{X}}^q = \begin{bmatrix} a^2 + b^2 - c^2 - d^2 & 2bc - 2ad & 2ac + 2bd \\ 2ad + 2bc & a^2 - b^2 + c^2 - d^2 & 2cd - 2ab \\ 2bd - 2ac & 2ab + 2cd & a^2 - b^2 - c^2 + d^2 \end{bmatrix} \quad (8)$$

and substituting it into the vector Equation 7 we get three polynomial equations in seven unknowns, *i.e.*, three translation parameters for  $\mathbf{t}_{\mathbf{X}}$  and four rotation parameters  $a, b, c,$  and  $d$ . Now we can apply this substitution to the two motions  $i$  and  $j$  and by adding the equation defining the unit quaternion  $q$  we obtain the following system of equations:

#### Problem 1 (Minimal Hand-Eye Calibration)

$$\begin{aligned} & \textit{Given} \quad \mathbf{R}_{\mathbf{A}_i}, \mathbf{R}_{\mathbf{A}_j}, \mathbf{t}_{\mathbf{A}_i}, \mathbf{t}_{\mathbf{A}_j}, \mathbf{t}_{\mathbf{B}_i}, \mathbf{t}_{\mathbf{B}_j} \\ & \textit{find} \quad \mathbf{R}_{\mathbf{X}} \in SO(3), \mathbf{t}_{\mathbf{X}} \in \mathbb{R}^3 \\ & \textit{subject to} \quad \mathbf{R}_{\mathbf{A}_i} \mathbf{t}_{\mathbf{X}} + \mathbf{t}_{\mathbf{A}_i} = \mathbf{R}_{\mathbf{X}}^q \mathbf{t}_{\mathbf{B}_i} + \mathbf{t}_{\mathbf{X}}, \\ & \quad \mathbf{R}_{\mathbf{A}_j} \mathbf{t}_{\mathbf{X}} + \mathbf{t}_{\mathbf{A}_j} = \mathbf{R}_{\mathbf{X}}^q \mathbf{t}_{\mathbf{B}_j} + \mathbf{t}_{\mathbf{X}}, \\ & \quad a^2 + b^2 + c^2 + d^2 = 1. \end{aligned}$$

Problem 1 is a well-constrained system of seven equations in seven unknowns. To solve it for the unknown hand-eye calibration  $\mathbf{X}$ , the Gröbner basis method can be readily used. This leads to a fast and non-iterative solution with no need for an initial solution estimate. Note that the minimal number of two relative movements without rotations  $\mathbf{R}_{B_i}$  and  $\mathbf{R}_{B_j}$  is needed to construct the system.

In case rotations  $\mathbf{R}_{B_i}$  and  $\mathbf{R}_{B_j}$  need to be recovered as well, by substituting the solutions for the rotation  $\mathbf{R}_X$  into the Equation 6 we get the rotations as

$$\mathbf{R}_{B_i} = \mathbf{R}_X^{-1} \mathbf{R}_{A_i} \mathbf{R}_X, \quad (9)$$

$$\mathbf{R}_{B_j} = \mathbf{R}_X^{-1} \mathbf{R}_{A_j} \mathbf{R}_X. \quad (10)$$

### 3.1 Gröbner Basis Method

The Gröbner basis method for solving systems of polynomial equations has recently become popular in computer vision and it has been used to create very fast, efficient and numerically stable solvers to many difficult problems. The method is based on polynomial ideal theory and is concerned with special bases of these ideals called Gröbner bases [3]. Gröbner bases have the same solutions as the initial system of polynomial equations defining the ideal but are often easier to solve. Gröbner bases are usually used to construct special multiplication (action) matrices [18], which can be viewed as a generalization of the companion matrix used in solving one polynomial equation in one unknown. The solutions to the system of polynomial equations is then obtained from the eigenvalues and eigenvectors of such action matrices. See [3, 4] for more on Gröbner basis methods and [20, 10, 1] for their applications in computer vision.

Since general algorithms [3] for computing Gröbner basis are not very efficient for solving problems which appear for example in computer vision, an automatic generator of specific polynomial equations solvers based on the Gröbner basis method has been proposed in [11]. These specific solvers often provide very efficient solutions to a class of systems of polynomial equations consisting of the same monomials and differing only in the coefficients.

Computer vision problems—like the hand-eye calibration problem presented in this paper—share the convenient property that the monomials appearing in the set of initial polynomials are always the same irrespective of the concrete coefficients arising from non-degenerate measurements. Therefore it is possible to use efficient specific solvers instead of less efficient general algorithms [3] for constructing the Gröbner bases.

The process of creating the specific solvers consists of two phases. In the first “offline” phase, the so-called “elimination templates” are found. These templates decide the elimination sequence in order to obtain all polynomials from the Gröbner basis or at least all polynomials necessary for the construction of the action matrix. This phase is performed only once for a given problem. In the second “online” phase, the elimination templates are used with coefficients arising from the specific measurements to construct the action matrix. Then, eigenvalues and eigenvectors of the action matrix provide solutions to the original polynomial equations. The automatic generator presented in [11] performs

the offline phase automatically and for an input system of polynomial equations outputs an efficient online solver.

### 3.2 Gröbner Basis Solver

To create an efficient solver for Problem 1 we used the automatic generator proposed in [11]. The Gröbner basis solver of the proposed hand-eye calibration problem starts with seven equations in seven unknowns, *i.e.*, three translation parameters for  $\mathbf{t}_X$  and four rotation parameters  $a, b, c,$  and  $d$ .

From the generator we obtained an elimination template which encodes how to multiply the seven input polynomials by the monomials and then how to eliminate the polynomials using the Gauss-Jordan (G-J) elimination process to obtain all polynomials necessary for the construction of the action matrix. In our case the automatic generator created the action matrix  $M_a$  for multiplication by  $a$ .

To get the elimination template the generator first generated all monomial multiples of the initial seven polynomial equations up to the total degree of four. This resulted in 252 polynomials in 330 monomials. Then the generator removed all unnecessary polynomials and monomials, *i.e.*, polynomials and monomials that do not influence the resulting action matrix. This resulted in matrix a  $182 \times 203$   $\mathbb{Q}$  representing the polynomials for the construction of the action matrix  $M_a$ , *i.e.*, the elimination template.

The online solver then only performs one G-J elimination of matrix  $\mathbb{Q}$  from the elimination template identified in the offline stage. This matrix contains coefficients which arise from specific measurements, *i.e.*, rotations  $R_{A_i}$  and  $R_{A_j}$  and translations  $\mathbf{t}_{A_i}, \mathbf{t}_{A_j}, \mathbf{t}_{B_i},$  and  $\mathbf{t}_{B_j}$ . After G-J elimination of matrix  $\mathbb{Q}$ , action matrix  $M_a$  can be created from its rows. The solutions to all seven unknowns can be found from the eigenvectors of the action matrix  $M_a$ . The online stage takes about 1 ms to finish in case of Problem 1.

This gives us a set  $\mathcal{X}_{ij}$  of up to 16 real solutions of  $\mathbf{X}$ . However each of these solutions appears twice, *i.e.*, there are double roots. Therefore we have only up to 8 different real solutions. Usually only one to four of them are *geometrically feasible*, *i.e.*, are real and of a reasonable length of the translation. The correct one can be chosen from the feasible solutions manually using some prior knowledge about the transformation  $\mathbf{X}$  or automatically using an additional set of solutions for different relative movements. The next section describes an automatic procedure for selecting the correct transformation.

## 4 Automatic Solution Selection

In order to automatically select the geometrically correct solution among the algebraically correct ones in  $\mathcal{X}_{ij}$ , at least one more set of solutions to Problem 1 for a different combination of relative movements is needed. Let  $\mathcal{X}_{k\ell}$  be such a set for two additional movements  $k$  and  $\ell$ . Supposing that the movements  $i, j$  and  $k, \ell$  form a geometrically non-degenerate configuration, we will find the

geometrically correct solution as  $\mathcal{X}_{ij} \cap \mathcal{X}_{kl}$ . In the presence of noise however, the intersection  $\mathcal{X}_{ij} \cap \mathcal{X}_{kl}$  will most likely be an empty set. In this case we have to select a solution from the union  $\mathcal{X}_{ij} \cup \mathcal{X}_{kl}$  that best fits the equations of Problem 1 for different motions. We will measure the fitness of a solution  $\mathbf{X}$  by the residual error of Equation 7

$$\mathbf{e}_i(\mathbf{X}) = \mathbf{R}_{A_i} \mathbf{t}_X + \mathbf{t}_{A_i} - \mathbf{R}_X \mathbf{t}_{B_i} - \mathbf{t}_X. \quad (11)$$

Now let us formalize the idea of selecting the best solution and to extend it to the case of more than two solution sets. Let  $n$  be the number of available relative movements and let  $I$  be a set of pairs of indexes of the relative movements

$$I \subset \{\{i, j\} : i, j \leq n\}, \quad |I| \geq 2. \quad (12)$$

Let  $\mathcal{X}$  be a set of solutions to Problem 1 for the pairs from the index set  $I$ ,

$$\mathcal{X} = \bigcup_{\{i, j\} \in I} \mathcal{X}_{ij}. \quad (13)$$

We select the geometrically correct solution among the solutions in  $\mathcal{X}$  by solving the following problem:

**Problem 2 (Minimal Hand-Eye Calibration for  $n$  Movements)**

$$\begin{aligned} & \text{Given } \mathbf{R}_{A_i}, \mathbf{t}_{A_i}, \mathbf{t}_{B_i}, I, i = 1, \dots, n \\ & \text{and a set of solutions } \mathcal{X} = \bigcup_{\{i, j\} \in I} \mathcal{X}_{ij} \\ & \text{find } \mathbf{X}^* = \arg \min_{\mathbf{X} \in \mathcal{X}} \sum_{i=1}^n \mathbf{e}_i(\mathbf{X})^\top \mathbf{e}_i(\mathbf{X}) \end{aligned}$$

As we can see from the above formulation, solving Problem 2 amounts to selecting a minimum from a set of  $|\mathcal{X}|$  real numbers.

In the presence of noise and in case  $n > 2$ , we can further refine the solution by an optimization method. For our experiments, we chose the method of Zhuang and Shiu [25] which requires a good initial estimate  $\mathbf{X}^0$ . By setting  $\mathbf{X}^0 \equiv \mathbf{X}^*$ , we can refine the solution by solving the following minimization problem:

**Problem 3 (Zhuang [25])**

$$\begin{aligned} & \text{Given } \mathbf{R}_{A_i}, \mathbf{t}_{A_i}, \mathbf{t}_{B_i}, i = 1, \dots, n \\ & \text{and an initial solution estimate } \mathbf{X}^0 \\ & \text{find } \mathbf{X}_{\text{opt}}^* = \arg \min \sum_{i=1}^n \mathbf{e}_i(\mathbf{X})^\top \mathbf{e}_i(\mathbf{X}) \\ & \text{subject to } \mathbf{R}_X \in SO(3) \end{aligned}$$

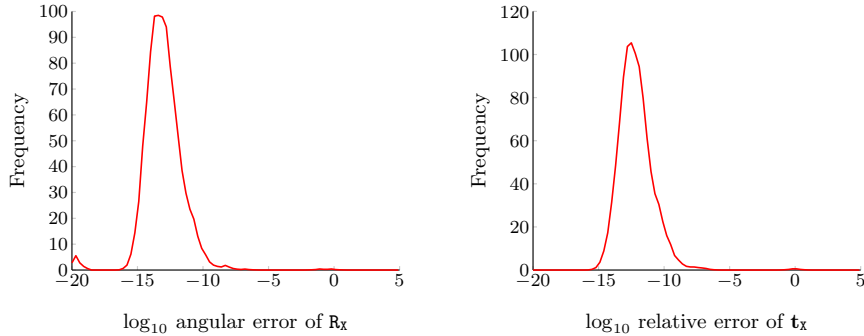


Fig. 2:  $\log_{10}$  angular error of the estimated rotation  $\mathbf{R}_x$  (Left) and  $\log_{10}$  translation error of  $\mathbf{t}_x$  (Right) for noise free data.

## 5 Experiments

To experimentally validate the proposed solutions, we use both synthetically generated and real world calibration scenarios. First, we use synthetically generated ground truth scenes to study the numerical stability of the proposed solution to Problem 1. Next, we study the behavior of the solutions to Problem 2 and Problem 3 on synthetic scenes consisting of 4 non-degenerate poses. Finally, we show the viability of the minimal solution in a real life experiment with a Mitsubishi MELFA-RV-6S serial manipulator with four draw-wire encoders attached to its end effector to recover the translations  $\mathbf{t}_{B_i}$ .

In all of the experiments we scaled the lengths of the input translation vectors  $\mathbf{t}_{B_i}$  and  $\mathbf{t}_{A_i}$  by the length of the largest one of them prior to running the Gröbner basis solver. We observe that this scaling improves the numerical stability of the solution.

The experiments were run on a 3GHz Intel Core i7 based desktop computer running 64-bit Linux. The Matlab implementation of the proposed method used in the experiments is available at <http://cmp.felk.cvut.cz/minimal/handeye.php>.

### 5.1 Experiments with Synthetic Data

**Numerical Stability Experiment.** First, we studied the behavior of the proposed Gröbner basis solver of Problem 1 to check its numerical stability. We generated 1000 random scenes with 100 points  $\mathbf{P}^k$ ,  $k = 1, \dots, 100$ , evenly distributed in the unit ball. Each scene consisted of 3 random absolute camera poses  $A'_i$ . The cameras were positioned to (i) be facing the center of the scene, (ii) see the scene points from the field of view (FOV) ranging from  $40^\circ$  to  $80^\circ$ . For every scene ground truth transformation  $\mathbf{X}_{\text{gt}}$  was generated so that the angle and the axis of  $\mathbf{R}_{\mathbf{X}_{\text{gt}}}$  were random and uniformly distributed and that  $\|\mathbf{t}_{\mathbf{X}_{\text{gt}}}\| \approx 0.1$ .



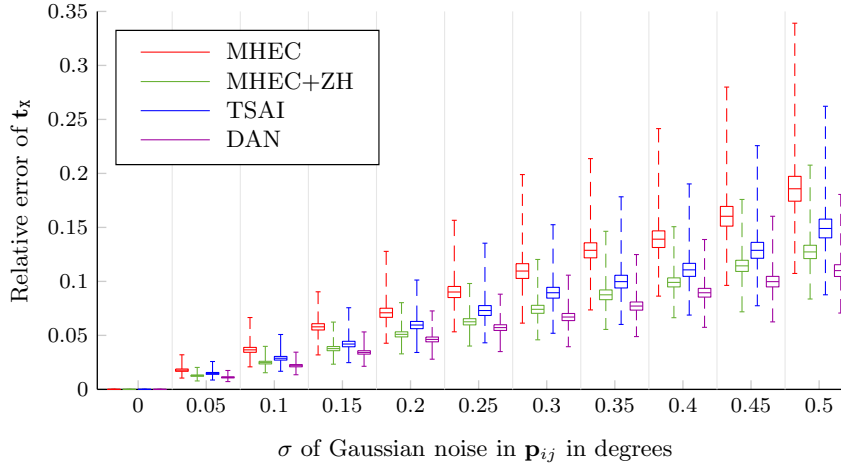


Fig. 3: Relative error of recovered translation  $\mathbf{t}_x$  for different levels of Gaussian noise.

Absolute robot poses  $\mathbf{B}'_i$  were determined by chaining  $\mathbf{x}_{\text{gt}}^{-1}$  and the generated absolute camera positions. For every combination of ground truth  $\mathbf{R}_{\mathbf{x}_{\text{gt}}}$ ,  $\mathbf{t}_{\mathbf{x}_{\text{gt}}}$  and the recovered  $\mathbf{R}_x$ ,  $\mathbf{t}_x$  we measured the error of the rotation as the angle  $\theta$  of the rotation  $\mathbf{R}_x^\top \mathbf{R}_{\mathbf{x}_{\text{gt}}}$ , such that  $0 \leq \theta \leq \pi$  and the error of translation as the relative error  $\|\mathbf{t}_x - \mathbf{t}_{\mathbf{x}_{\text{gt}}}\| / \|\mathbf{t}_{\mathbf{x}_{\text{gt}}}\|$ . Figure 2 shows the histograms of the respective errors, certifying the numerical stability of the solver.

**Calibration Experiment.** In this experiment we analyzed the performance with respect to image noise. We used the same scheme to generate random scenes as in Numerical Stability Experiment. This time, we generated four absolute robot poses in each scene and recovered the absolute camera positions by P3P algorithm [14].

We started by computing  $\mathbf{P}_i^k$ —the positions of the 100 random points  $\mathbf{P}^k$  with respect to the coordinate systems of the cameras  $\mathbf{A}'_i$ ,  $i = 1, \dots, 4$ . Further, we normalized  $\mathbf{P}_i^k$  to get only the directional vectors  $\mathbf{p}_i^k$  that were progressively corrupted with angular Gaussian noise. Finally, we used P3P in RANSAC loop to obtain noise corrupted absolute camera poses  $\mathbf{A}'_i$ ,  $i = 1, \dots, 4$ .

We experimented with 11 levels of angular Gaussian noise with the standard deviation  $\sigma$  ranging from 0 to 0.5 degrees, with the highest noise level translating to  $\sigma$  of ca. 20–40 pixels for a 8MP camera with 40°–80° field of view. We generated and recovered camera poses for 1000 random scenes for every noise level.

We recovered hand-eye calibrations  $\mathbf{X}$  by four different methods. The first method MHEC identifies the results obtained by the Gröbner basis solver with the solution selected according to Problem 2. The second method MHEC+ZH

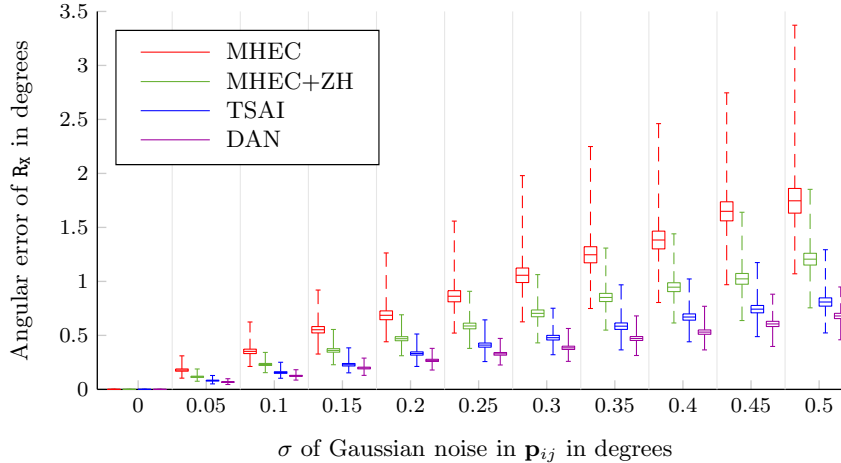


Fig. 4: Angular error of recovered rotation  $R_x$  for different levels of Gaussian noise.

stands for the results obtained by the method [25] (Problem 3) when initialized by the results of MHEC. For completeness sake, we include results obtained by the methods [24] labeled as TSAI and [5] labeled in the figures as DAN. These methods are not the direct competitors, since they require known robot rotations  $R_B$ . However, they can be used to gauge the accuracy of the results obtained by MHEC and MHEC+ZH.

Figures 3, 4, and 5 show the statistics of the obtained solutions using the Matlab `boxplot` function depicting values 25% to 75% quantile as a box with horizontal line at median. Figures 3 and 4 show the respective errors of  $\mathbf{t}_x$  and  $R_x$  using the same measures as described in Numerical Stability Experiment. Figure 5 shows the mean distance between the points  $\mathbf{P}_i^k$  transformed into the coordinate system of the gripper using the ground truth hand-eye transformation and the same points transformed into the coordinate system of the gripper using the estimated  $X$ . Note that the points were generated into the unit ball, *i.e.*, considering the diameter of this ball to be one meter means that the errors in Figure 5 are in meters.

## 5.2 Real Scene Data Experiment

In order to acquire a real scene calibration data, four draw-wire encoders were connected to the gripper of a Mitsubishi MELFA-RV-6S serial manipulator. A Canon 350D digital SLR camera with a Sigma 8 mm lens (cca.  $130^\circ$  field of view) was also attached to the gripper to form a hand-eye system.

The robot was instructed to move the gripper to (i) the home position with the known rotation w.r.t. the robot base, (ii) the four positions (backward, for-

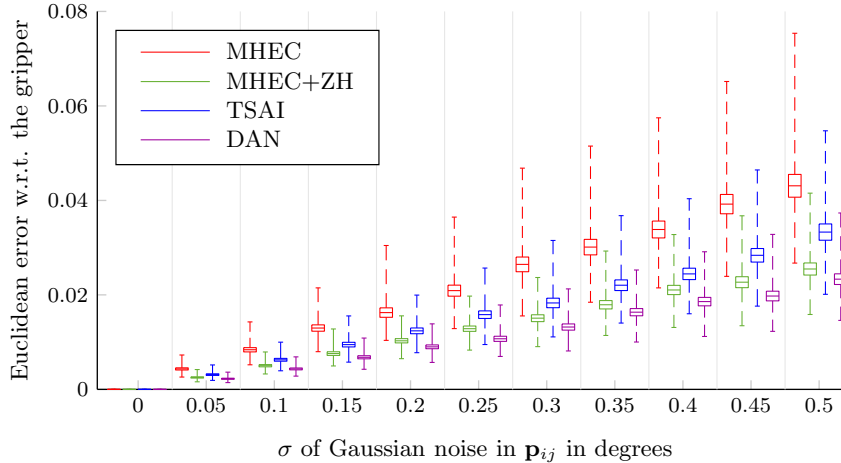


Fig. 5: Euclidean error of recovered calibration  $\mathbf{X}$  for different levels of Gaussian noise.

ward, left, right) distant approximately 400 mm at 10 degree pitch, (iii) the same four positions at 20 degree pitch, (iv) the position approximately 250 mm under the home position, and (v) the four positions at this height at 10 and 20 degree pitch again. While the robot was moving, the camera was remotely triggered to acquire  $2,592 \times 1,728$  pixels large images of a circular view field with 1,040 pixels radius.

The internal calibration of the camera in the form of a 2-parameter equi-angular model [12] was obtained using an image of a checkerboard with manually labeled corners. Then, a state-of-the-art sequential structure-from-motion pipeline [22] was used to automatically generate MSER, SIFT, and SURF feature points, perform approximate nearest neighbor matching in the descriptor space, verify the matches by pairwise epipolar geometries estimated by the 5-point algorithm [13] in a RANSAC loop, and create tracks and triangulated 3D points from verified matches spanning several images. The reconstructed 3D model was scaled to millimeter units by knowing the real dimensions of the checkerboard and measuring the distance of the corresponding 3D points in the model.

We used the system of four draw-wire encoders to determine the absolute positions of the gripper w.r.t. the robot base. For the experiment we chose 2 motions ending in the robots home position. Since the rotation of the robot in the home position is known, it is possible to transform the positions provided by the draw-wire encoders into the home position coordinate system and obtain translations  $\mathbf{t}_{B_1}$  and  $\mathbf{t}_{B_2}$ . We used  $\mathbf{t}_{B_1}$  and  $\mathbf{t}_{B_2}$  in combination with  $\mathbf{A}_1$ ,  $\mathbf{A}_2$  obtained from structure-from-motion to compute the hand-eye transformation  $\mathbf{X}$  and the relative gripper rotations  $\mathbf{R}_{B_1}$  and  $\mathbf{R}_{B_2}$ .

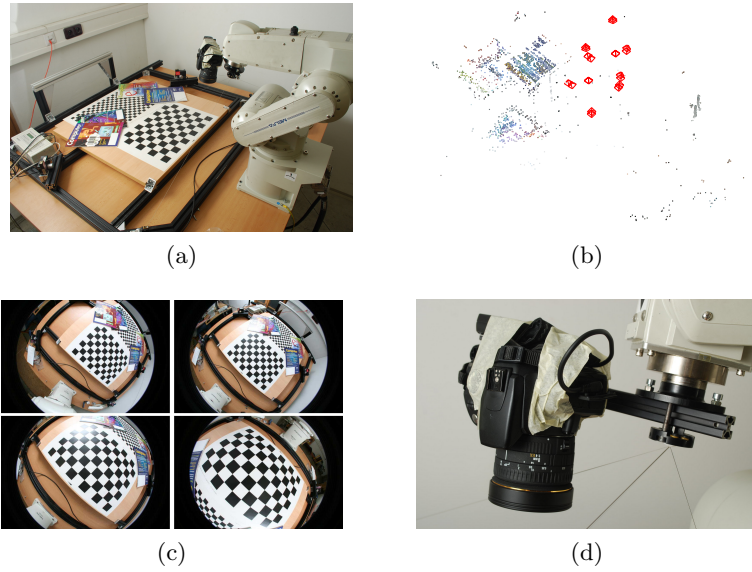


Fig. 6: Real data experiment. (a) A Mitsubishi MELFA-RV 6S serial manipulator used to acquire the data for the experiment. (b) The 3D model obtained from SfM. (c) Sample images of our scene taken by the camera mounted on the gripper of the robot. (d) Close up of the camera-gripper rig with draw-wire encoders.

For comparison, we also used  $\mathbf{t}_{B_{gt2}}$  and  $\mathbf{t}_{B_{gt1}}$  from robots positioning software with the same camera motions  $A_1$  and  $A_2$  to compute hand-eye transformation  $\bar{X}$ ,  $\bar{R}_{B_1}$ , and  $\bar{R}_{B_2}$ .

Since the robot was calibrated, we can also compare the computed gripper rotations  $R_{B_1}$ ,  $R_{B_2}$ ,  $\bar{R}_{B_1}$ , and  $\bar{R}_{B_2}$  with the rotations  $R_{B_{gt1}}$  and  $R_{B_{gt2}}$  from the robots positioning software, see Table 1.

Table 1: Angular rotation errors of estimated gripper rotations in degrees.

	$R_{B_1}$	$R_{B_2}$	$\bar{R}_{B_1}$	$\bar{R}_{B_2}$
$R_{B_{gt1}}$	0.84	—	0.89	—
$R_{B_{gt2}}$	—	0.61	—	1.09

Finally, let us express the obtained translations from the gripper to the camera center using the translation from the draw-wire encoders  $-\mathbf{R}_X^\top \mathbf{t}_X = (110.2, 26.2, 47.9)$ , and using the translation from the robot,  $-\mathbf{R}_X^\top \mathbf{t}_{\bar{X}} = (126.5, 28.7, 51.1)$ .

These results are consistent with each other as well as with the rough physical measurement of the mechanical reduction and show the validity of the obtained results.

## 6 Conclusion

We presented the first minimal problem of hand-eye calibration for the situations where the gripper's rotations are not known. We formulated the problem as a system of seven equations in seven unknowns and solved it using the Gröbner basis method for solving systems of polynomial equations providing the first exact algebraic solution to the problem. This solution uses the minimal number of two relative movements. Further, we showed how to select the geometrically correct solution using additional relative movements. Finally, our experiments showed that the proposed solver is numerically stable, fast and—since it can handle noisy inputs—that its results can be successfully used as initialization of subsequent minimization methods.

## Acknowledgment

The authors were supported by the EC under projects FP7-SME-2011-285839 De-Montes and FP7-288553 CloPeMa and by Grant Agency of the CTU Prague project SGS12/191/OHK3/3T/13. The authors would also like to thank Michal Havlena and Martin Meloun for their help with the real-data experiment.

## References

1. Martin Bujnak, Zuzana Kukelova, and Tomas Pajdla. A general solution to the P4P problem for camera with unknown focal length. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.
2. Jack C. K. Chou and M. Kamel. Finding the position and orientation of a sensor on a robot manipulator using quaternions. *International Journal of Robotics Research*, 10(3):240–254, 1991.
3. D.A. Cox, J.B. Little, and D. O'Shea. *Using Algebraic Geometry*. Graduate Texts in Mathematics. Springer, 2005.
4. D.A. Cox, J.B. Little, and D. O'Shea. *Ideals, Varieties, And Algorithms: An Introduction to Computational Algebraic Geometry And Commutative Algebra*. Number v. 10 in Undergraduate Texts in Mathematics. Springer, 2007.
5. Konstantinos Daniilidis. Hand-eye calibration using dual quaternions. *International Journal of Robotics Research*, 18:286–298, 1998.
6. M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
7. Jan Heller, Michal Havlena, and Tomas Pajdla. A branch-and-bound algorithm for globally optimal hand-eye calibration. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

8. Jan Heller, Michal Havlena, Akihiro Sugimoto, and Tomáš Pajdla. Structure-from-motion based hand-eye calibration using  $l_\infty$  minimization. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3497–3503, 2011.
9. Radu Horaud and Fadi Dornaika. Hand-eye calibration. *The International Journal of Robotics Research*, 14(3):195–210, 1995.
10. Zuzana Kukelova and Tomas Pajdla. A minimal solution to the autocalibration of radial distortion. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.
11. Zuzana Kukelova, Martin Bujnak, and Tomas Pajdla. Automatic generator of minimal problem solvers. In *Proceedings of European Conference on Computer Vision*, pages 302–315, Berlin, Heidelberg, 2008.
12. Branislav Micusik and Tomas Pajdla. Structure from motion with wide circular field of view cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006.
13. D. Nistér. An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):756–770, 2004.
14. David Nistér and Henrik Stewénius. A minimal solution to the generalised 3-point pose problem. *Journal of Mathematical Imaging and Vision*, 27(1):67–79, 2007.
15. F.C. Park and B.J. Martin. Robot sensor calibration: solving  $AX=XB$  on the euclidean group. *IEEE Transactions on Robotics and Automation*, 10(5):717–721, 1994.
16. Thomas Ruland, Tomas Pajdla, and Lars Kruger. Globally optimal hand-eye calibration. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2012.
17. Y.C. Shiu and S. Ahmad. Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form  $AX=XB$ . *IEEE Transactions on Robotics and Automation*, 5(1):16–29, 1989.
18. Hans J. Stetter. *Numerical polynomial algebra*. SIAM, 2004.
19. H. Stewénius, K. Aström. Hand-eye calibration using multilinear constraints. In *Proceedings of the Asian Conference on Computer Vision*, 2004.
20. H. Stewénius, F. Schaffalitzky, and D. Nister. How hard is 3-view triangulation really? In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 686–693, 2005.
21. K. H. Strobl and G. Hirzinger. Optimal Hand-Eye Calibration. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4647–4653, Beijing, China, 2006.
22. Akihiko Torii, Michal Havlena, and Tomáš Pajdla. Omnidirectional image stabilization for visual object recognition. *International Journal of Computer Vision*, 91(2):157–174, 2011.
23. R.Y. Tsai and R.K. Lenz. Real time versatile robotics hand/eye calibration using 3D machine vision. In *IEEE International Conference on Robotics and Automation*, pages 554–561 vol.1, 1988.
24. R.Y. Tsai and R.K. Lenz. A new technique for fully autonomous and efficient 3D robotics hand/eye calibration. *IEEE Transactions on Robotics and Automation*, 5(3):345–358, 1989.
25. Hanqi Zhuang and Yiu Cheung Shiu. A noise tolerant algorithm for wrist-mounted robotic sensor calibration with or without sensor orientation measurement. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 2, pages 1095–1100, 1992.