

Calibrating a Robot Camera

Dekun Yang and John Illingworth,
Department of Electronics and Electrical Engineering,
University of Surrey, Guildford. GU2 5XH

Abstract

This paper addresses the problem of calibrating a camera mounted on a robot arm. The objective is to estimate the camera's intrinsic and extrinsic parameters. These include the relative position and orientation of camera with respect to robot base as well as the relative position and orientation of the camera with respect to a pre-defined world frame. A calibration object with a known 3D shape is used together with two known movements of the robot. A method is presented to find calibration parameters within an optimisation framework. This method differs from existing methods in that 1) it fully exploits information from different displacements of the camera to produce an optimal calibration estimate, and 2) it uses an evolutionary algorithm to attain the optimal solution. Experimental results on both synthetic and real data are presented.

1 Introduction

Camera calibration involves establishing for a given camera configuration the relationship between 3D scene points and their corresponding 2D image coordinates. It is essential for performing 3D metric measurements from images. A camera calibration permits identification of the 3D ray of scene points which could have produced a given 2D image point. Camera calibration consists of determining both the intrinsic parameters that are the camera's optical characteristics and also the extrinsic parameters that describe the relative position and orientation of the camera with respect to a coordinate system of interest. For a static camera, the extrinsic parameters correspond to a single transformation: camera-to-world. However, if the camera is mounted on a movable robot arm then the extrinsic parameters are composed of a camera-to-robot and a robot-to-world transformation. The camera-to-robot transformation is a function of the motion of the robot arm. The extrinsic parameters are important for making use of controlled camera motion within an active vision system as robot motion is usually specified in a robot coordinate system while 3D metric measurements are derived from measurements made with respect to the camera coordinate system.

A large body of work exists on camera calibration in both the photogrammetry and computer vision communities. A survey of existing techniques was given by Tsai [10]. The intrinsic parameters and the camera-to-world transformation are usually estimated using a calibration object with known 3D reference points in the world. A method presented by Tsai [9] obtains the extrinsic parameters from a system of linear equations and then recovers the intrinsic parameters. Faugeras

and Toscani [4] suggested a different route whereby they first compute the perspective transformation matrix which relates 3D world coordinates to 2D image coordinates and then use this to determine the intrinsic and extrinsic parameters. Recently, self-calibration methods for determining the intrinsic parameters of a camera have been developed by several researchers. They usually exploit camera motion. Faugeras *et al.* [3] presented a method which is applicable to the general case of unknown camera motion. Dron [1] developed a method using translations of a camera while Du and Brady [2] developed a method using rotations. However, these self-calibration methods cannot estimate the extrinsic camera-to-world parameters because they do not use 3D reference points. The extrinsic camera-to-robot parameters are usually estimated using controlled robot motions. Two known robot motions are required to obtain a unique solution. Shiu and Ahmad [8] formulated this problem as a set of quadratically constrained linear regression equations and solved it by a least squares technique. Tsai and Lenz [11] proposed a simpler and more efficient closed-form method. Wang [12] developed a method which was similar to Tsai and Lenz.

A difficulty of estimating intrinsic parameters is that existing methods seek the solution through some intermediate unknowns that are composites of the intrinsic parameters. Although the least square technique for estimating the intermediate unknowns is linear, the procedure for recovering the intrinsic parameters from the intermediate unknowns is not. Small errors in the estimation of the intermediate unknowns induce large errors in the intrinsic parameter estimation. However, since the intrinsic parameters are invariant to camera location their estimation can be improved by using several displacements of the camera. Most of the existing methods of calibration have closed-form expressions for the solution. These methods are simple and efficient but can have numerical instabilities due to the inadequacy of low level vision for extracting image features. It is necessary to use optimization techniques to achieve more robust calibration results.

In this paper, we present a method for optimally estimating the intrinsic and extrinsic parameters of a single camera by fully exploiting information from three locations of the camera. The intrinsic parameter estimation can be improved by the information redundancy inherent in the use of three locations of the camera. We formulate the calibration problem as a non-linear optimisation problem. The optimal solution is the one that minimizes the difference between estimated projections of 3D reference points and the corresponding measured image coordinates under the constraint that the orthogonality of rotation matrixes is preserved. An evolutionary algorithm, a systematic multi-agent stochastic search technique, is proposed to find the optimal calibration parameters.

The paper is organised as follows: In section 2 a camera model is described and the calibration problem is stated. In section 3 the calibration problem is formulated as a nonlinear optimization problem. In section 4 the determination of the initial intrinsic and extrinsic parameters for the optimization task is considered. In section 5 an evolutionary algorithm is presented to find the optimal calibration parameters. Experimental results on synthetic and real data are provided in section 6 and a summary in section 7 concludes the paper.

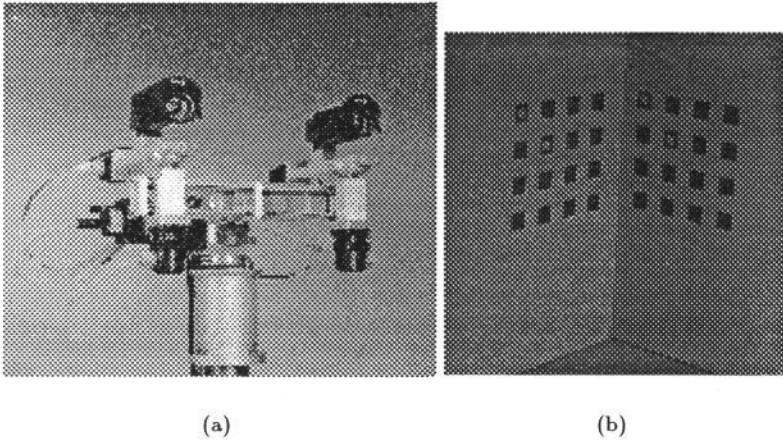


Figure 1: (a) GETAFIX: Surrey VSSP Group Robot Head (b) Calibration Chart

2 Camera Model and Problem Statement

The accuracy of scene measurement is a function of the complexity of the camera model adopted. The best result is obtained when a complete model including nonlinear distortions due to imperfections and misalignment of the optical system is used. In this paper a pinhole camera model without correction for lens distortion is used. Given a 3D point M , its image m is the intersection of the image plane and the line through M and the optical center C . The optical axis is the line which goes through optical center C and is perpendicular to the image plane. The camera coordinate system $X_c Y_c Z_c$ is defined such that its origin is the optical center, the Z axis is the optical axis and the X and Y axes are parallel to those of the image coordinate system. Consequently the coordinate of a 3D point (x_c, y_c, z_c) with respect to the camera coordinate system and its image coordinates (u, v) are related by

$$\begin{pmatrix} \omega u \\ \omega v \\ \omega \end{pmatrix} = \begin{bmatrix} f/s_u & 0 & u_0 \\ 0 & f/s_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} \quad (1)$$

where ω is an arbitrary scalar and the elements in the 3×3 matrix are composed from the camera's intrinsic parameters: f is the focal length of the camera, s_u and s_v are horizontal and vertical pixel size units and (u_0, v_0) is the the principal point of the camera, i.e., the intersection between the optical axis and the image plane. Since the focal length f cannot be separated from pixel size s_u or s_v the calibration deals with the ratios $k_u = f/s_u$ and $k_v = f/s_v$. Thus four intrinsic parameters have to be estimated.

In this paper the problem of calibrating a movable camera is considered. In our laboratory the camera is one of two which form the GETAFIX stereo robot head, see Figure 1 (a). For the purpose of the current paper the robot is only allowed to rotate about a vertical axis and a horizontal axis, pan and tilt. These movements

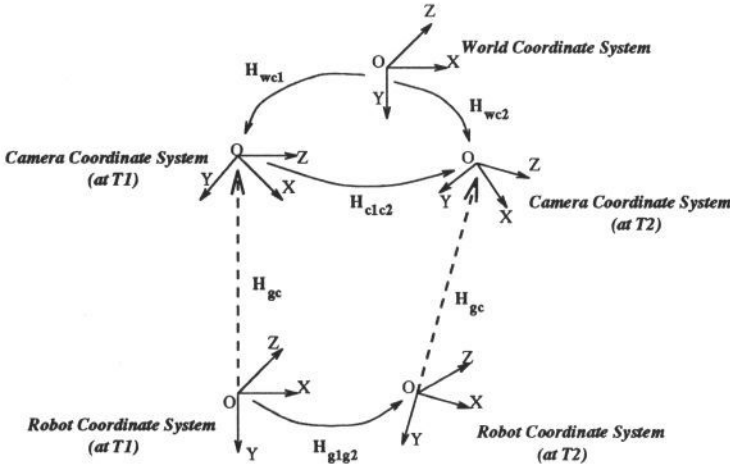


Figure 2: Relationship among coordinate systems for two robot locations.

are specified as pure rotations about the x and y axes of the robot coordinate system. The camera is rigidly attached to the robot, i.e., the relative position and orientation between the camera and robot remains unchanged as the robot moves to different locations. For calibration the chart shown in Figure 1 (b) is placed in view of the camera. The transformation between two coordinate systems, say W and C , is described by a homogeneous transform matrix, \mathbf{H}_{wc} i.e. coordinates of a 3D point in system C , (x_c, y_c, z_c) , are related to coordinates (x_w, y_w, z_w) in W by

$$(x_c, y_c, z_c, 1)^t = \mathbf{H}_{wc}(x_w, y_w, z_w, 1)^t \quad (2)$$

where superscript t denotes transpose and

$$\mathbf{H}_{wc} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

$R = \{r_{ij}\}$ and $T = (t_1, t_2, t_3)^t$ are the rotation matrix and the translation vector, respectively, from system W to coordinate system C .

In this paper calibration is achieved by making controlled motions of the robot and observing the changes in the image coordinates of the calibration chart. Figure 2 shows the relationships between various coordinate systems, world W , robot G and camera C , for two robot locations.

In order to achieve camera calibration, two known movements of the robot are made to obtain three different locations of the camera. The problem to be addressed can now be stated as follows:

given n ($n \geq 7$) reference points in the world coordinate system (x_i, y_i, z_i) , ($i = 1, \dots, n$) and their image coordinates for three different camera locations (u_{ij}, v_{ij}) , ($i = 1, 2, 3, j = 1, \dots, n$), and given two known motions of the robot \mathbf{H}_{g1g2} and \mathbf{H}_{g2g3} , estimate both the intrinsic f, k_u, k_v, u_0, v_0 , and extrinsic camera parameters $\mathbf{H}_{gc}, \mathbf{H}_{wc1}$.

3 Optimal Calibration Formulation

The standard methods for determining the intrinsic and extrinsic parameters of the camera-to-world transformation are in closed-form expressions given by Tsai [9] and Faugeras and Toscani [4]. The accuracy of the closed-form solution to the intrinsic and extrinsic parameters depends on several factors: (i) the uncertainty of image feature coordinates; and (ii) the non-linearity of analytical methods for computing the parameters. Since the parameters are recovered from intermediate unknowns non-linearly then small errors in the estimation of the intermediate unknowns may result in large errors in parameter estimates. Thus, it is desirable to improve the estimate using optimization techniques.

The intrinsic parameters and camera-to-robot transformation are invariant to different locations of the camera. Thus it is possible to exploit the redundancy of data obtained from three locations of the camera to improve the estimation of these parameters. This can be achieved by directly estimating them using optimization techniques with an initial estimate.

In order to reduce the number of parameters to be optimized and preserve the orthogonality of the rotation matrix, the rotation matrix can be represented by three independent parameters: the angles of rotation (α, β, γ) around the x, y and z axes respectively i.e.

$$R = \begin{bmatrix} \cos\beta\cos\gamma & \sin\alpha\sin\beta\cos\gamma + \cos\alpha\sin\gamma & -\cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma \\ -\cos\beta\sin\gamma & -\sin\alpha\sin\beta\sin\gamma & -\cos\alpha\sin\beta\sin\gamma + \sin\alpha\cos\gamma \\ \sin\beta & -\sin\alpha\cos\beta & \cos\alpha\cos\beta \end{bmatrix}$$

as where $\sin\beta = r_{31}$, $\tan(\alpha) = -r_{32}/r_{33}$ and $\tan(\gamma) = -r_{21}/r_{11}$. Thus the parameters to be estimated are

$$\mathbf{p} = \{k_u, k_v, u_0, v_0, R_{gc}, T_{gc}, R_{wc1}, T_{wc1}\} \quad (4)$$

This is a 16-dimensional parameter space.

The calibration problem is formulated as a non-linear minimization problem. The optimal solution is the one that minimizes the difference between projections of 3D reference points using calibrated parameters and the actual corresponding image coordinates. Consider the j -th 3D point in the world coordinate system (x_j, y_j, z_j). It can be seen from equation (3) that the corresponding image coordinates in the three camera coordinate systems are respectively:

$$(x_{1j}, y_{1j}, z_{1j}, 1)^t = \mathbf{H}_{wc1}(x_j, y_j, z_j, 1)^t \quad (5)$$

$$(x_{2j}, y_{2j}, z_{2j}, 1)^t = \mathbf{H}_{gc}\mathbf{H}_{g1g2}\mathbf{H}_{gc}^{-1}\mathbf{H}_{wc1}(x_j, y_j, z_j, 1)^t \quad (6)$$

$$(x_{3j}, y_{3j}, z_{3j}, 1)^t = \mathbf{H}_{gc}\mathbf{H}_{g2g3}\mathbf{H}_{g1g2}\mathbf{H}_{gc}^{-1}\mathbf{H}_{wc1}(x_j, y_j, z_j, 1)^t \quad (7)$$

In each camera coordinate system, $(x_{ij}, y_{ij}, z_{ij})^t, (i = 1, 2, 3)$ is related by its image coordinate $(u_{ij}, v_{ij})^t$ by

$$u_{ij} = k_u \frac{x_{ij}}{z_{ij}} - u_0 \quad \text{and} \quad v_{ij} = k_v \frac{y_{ij}}{z_{ij}} - v_0$$

Therefore the optimal calibration parameters $\hat{\mathbf{p}}$ can be found by performing the following non-linear minimization.

$$\hat{\mathbf{p}} = \text{Min}_{\mathbf{p}} f(\mathbf{p})$$

where

$$f(\mathbf{p}) = \sum_{i=1}^3 \sum_{j=1}^n [(u_{ij} - x_{ij}/z_{ij})^2 + (v_{ij} - y_{ij}/z_{ij})^2] \quad (8)$$

It is note that this criterion has a clear physical meaning as the minimization of a Euclidean distance in the image plane.

4 Setting Initial Calibration Parameters

Firstly we describe an alternative closed-form solution to the intrinsic parameters and the camera-to-world transformation. Given n 3D points in the world coordinate system $(x_i, y_i, z_i) (i = 1, \dots, n)$ and their corresponding image coordinates $(u_i, v_i) (i = 1, \dots, n)$, the following equations can be derived by combining equation (1), (2) and (3) and assuming that $t_3 \neq 0$

$$\begin{bmatrix} u_1 x_1 & u_1 y_1 & u_1 z_1 & -x_1 & -y_1 & -z_1 & -1 \\ u_2 x_2 & u_2 y_2 & u_2 z_2 & -x_2 & -y_2 & -z_2 & -1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -1 \\ u_n x_n & u_n y_n & u_n z_n & -x_n & -y_n & -z_n & -1 \end{bmatrix} \begin{pmatrix} r_{31}/t_3 \\ r_{32}/t_3 \\ r_{33}/t_3 \\ (k_u r_{11} + r_{31} u_0)/t_3 \\ (k_u r_{12} + r_{32} u_0)/t_3 \\ (k_u r_{13} + r_{33} u_0)/t_3 \\ (k_u t_1 + u_0 t_3)/t_3 \end{pmatrix} = \begin{pmatrix} -u_1 \\ -u_2 \\ \cdot \\ \cdot \\ \cdot \\ -u_n \end{pmatrix} \quad (9)$$

where $R = \{r_{ij}\}$ and $T = (t_1, t_2, t_3)^t$ are the camera-to-world rotation and translation. For $n \geq 7$, this yields an overdetermined system of linear equations. The seven unknowns $a_1 = r_{31}/t_3$, $a_2 = r_{32}/t_3$, $a_3 = r_{33}/t_3$, $a_4 = (k_u r_{11} + r_{31} u_0)/t_3$, $a_5 = (k_u r_{12} + r_{32} u_0)/t_3$, $a_6 = (k_u r_{13} + r_{33} u_0)/t_3$, $a_7 = (k_u t_1 + u_0 t_3)/t_3$ can be obtained using a standard linear least squares technique. Since the rotation matrix \mathbf{R} is orthogonal, i.e. $\mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I}$ in which there are six independent relations, the degree of freedom of \mathbf{R} is three. Using the orthogonality property, the calibration parameters can be obtained from the eleven intermediary variables $a_i, i = 1, \dots, 7$ as follows.

$$|t_3| = (a_1^2 + a_2^2 + a_3^2)^{-\frac{1}{2}}$$

where the sign of t_3 can be determined by the actual setup of the calibration chart.

$$\begin{aligned} r_{3i} &= a_i/t_3, \quad (i = 1, 2, 3) \\ u_0 &= t_3(a_4 r_{31} + a_5 r_{32} + a_6 r_{33}) \\ k_u &= \sqrt{(t_3 a_4 - r_{31} u_0)^2 + (t_3 a_5 - r_{32} u_0)^2 + (t_3 a_6 - r_{33} u_0)^2} \\ r_{1i} &= (t_3 a_{i+3} - r_{3i} u_0)/k_u, \quad (i = 1, 2, 3) \\ t_1 &= t_3(a_7 - u_0)/k_u \end{aligned}$$

Other calibration parameters v_0 , k_v , t_2 and r_{2i} , $i = 1, 2, 3$ can be obtained in a similar way.

Secondly we can obtain a closed-form solution to the camera-to-robot transformation once we have the camera-to-world transformations for three different locations of the camera. For each camera location the camera-to-world transformations is estimated H_{wc_1} , H_{wc_2} and H_{wc_3} . This permits calculation of the transformations between different camera locations, i.e., $H_{c_1c_2}$ and $H_{c_2c_3}$. The transformation algebra yields the following relations among transformations related by the robot motions:

$$R_{c_1c_2}R_{gc} = R_{gc}R_{g_1g_2} \quad (10)$$

$$(R_{c_1c_2} - I)T_{gc} = R_{gc}T_{g_1g_2} - T_{c_1c_2} \quad (11)$$

where $(R_{g_1g_2}, T_{g_1g_2})$ is the transformation between two locations of the robot. It has been shown that [8, 10, 12] the (R_{gc}, T_{gc}) can be uniquely determined when two displacements $H_{c_1c_2}$ and $H_{c_2c_3}$ are available. Let the rotation R be specified by an axis k and an angle θ . The rotation axis and angle for R_{gc} can be obtained from equation (10) [10, 12] as follows:

$$\kappa_{gc} = \frac{(\kappa_{c_1c_2} - \kappa_{g_1g_2}) \times (\kappa_{c_2c_3} - \kappa_{g_2g_3})}{\|(\kappa_{c_1c_2} - \kappa_{g_1g_2}) \times (\kappa_{c_2c_3} - \kappa_{g_2g_3})\|}$$

$$\begin{aligned} \theta_{gc} = & \operatorname{sgn}((\kappa_{c_1c_2} - \kappa_{g_1g_2})^t (\kappa_{gc} \times (\kappa_{c_1c_2} + \kappa_{g_1g_2}))) \tan^{-1} \left(\frac{\|\kappa_{c_1c_2} - \kappa_{g_1g_2}\|}{\|(\kappa_{gc} \times (\kappa_{c_1c_2} + \kappa_{g_1g_2}))\|} \right) \\ & + \operatorname{sgn}((\kappa_{c_2c_3} - \kappa_{g_2g_3})^t (\kappa_{gc} \times (\kappa_{c_2c_3} + \kappa_{g_2g_3}))) \tan^{-1} \left(\frac{\|\kappa_{c_2c_3} - \kappa_{g_2g_3}\|}{\|(\kappa_{gc} \times (\kappa_{c_2c_3} + \kappa_{g_2g_3}))\|} \right) \end{aligned}$$

where $\operatorname{sgn}(\bullet)$ is the sign function, i.e., $\operatorname{sgn}(x) = 1$ if $x \geq 0$ and $\operatorname{sgn}(x) = -1$ if $x < 0$. Once R_{gc} is estimated T_{gc} can be obtained from equation (11) using a standard least squares technique.

5 Minimization by an Evolutionary Algorithm

The success of the calibration method based on optimization framework depend on the existence of algorithms for minimizing the objective function in equation (8). The minimization of the nonlinear function is a difficult task because the function is not convex and contains complex local minima. Gradient-descent methods are not suitable for this task because: (i) they are susceptible to becoming trapped at local minima; and (ii) the gradient information of the objective function is not available. Simulated annealing is an optimization method to escape the local optima via a stochastic search mechanism, but it is not efficient as it performs an essentially serial and random search in the solution space.

Recently, evolutionary algorithms [5] have been studied for solving complex optimization tasks. Evolutionary algorithms are population-based search methods: successive populations of candidate solutions are generated in a stochastic manner and optimization is performed by multi-agent stochastic search. Although evolutionary algorithms have been shown to be promising in many applications (see [6]

and references therein), it is still new to researchers in computer vision community. Our aim here is to propose an evolutionary algorithm to find the optimal solution to the calibration problem.

Given an initial solution of 16-dimension $\mathbf{p} = (p(1), \dots, p(16))$ obtained from an analytical method described in previous section, we first form an initial population $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N\}$, where N is the size of population, by perturbing the initial solution along 16 directions respectively. We propose an evolutionary algorithm which has the following two operators:

1. Mutation: For each candidate solution $\mathbf{p}_i \in \mathbf{P}, i = 1, \dots, N$, we generate a new candidate solution \mathbf{p}'_i by perturbing it in terms of the following function

$$p'_i(j) = p_i(j) + N(0, \lambda f(\mathbf{p}_i)), \quad j = 1, \dots, 16 \quad (12)$$

where λ is the scale factor controlling the perturbation size and $N(\mu, \sigma^2)$ is a Gaussian variable with mean μ and variance σ^2 .

2. Selection: For each pair of candidate solutions \mathbf{p}_i and \mathbf{p}'_i , we select one of them which has lower cost f as a member of next generation of the population.

The evolutionary algorithm performs optimization by repeating the above operations until a stopping criteria are satisfied. The stopping criterion is that the lowest cost f_{min} remains the same for a sufficiently long time. The mutation operator leads to an increase in the diversity of the population induced through a stochastic operation. Thus, the mutation provides a capacity to escape local minima during search. In contrast, the selection operator retains the best candidate solutions in the population, which guarantees asymptotic convergence. The combination of the above mutation and selection operators leads to the global minimum of the objective function in equation (8). More detailed theoretical analysis of general evolutionary algorithms can be found in [7].

6 Experimental Results and Discussion

Firstly we need set up the parameters in the evolutionary algorithm. In the experiments described below, we set the population size $N = 5$ and the perturbation size $\lambda = 0.6$. Under these parameters the evolutionary algorithm converges within 20 generations.

Synthetic data was used to test the validity and evaluate the performance of the proposed method. 3D reference points were chosen and three synthetic images were simulated corresponding to the two movements of the robot. Uniform random noise from the interval $[0, k]$ is added independently to both the u and v coordinates of the projections. Accordingly, the maximum perturbation of point location is $k\sqrt{2}$ pixels. The proposed method was tested under several noise conditions $k = 0.5, 1.0, 1.5, 2.0$. The accuracy of the method was taken as the average distance between the projection of reference points and the actual image data. From the results obtained, conclusions about accuracy as a function of level of noise and the size of robot movements can be made. When noise level is low, i.e., $k \leq 1$ and their is a relatively large robot rotation angle ($\geq 5^\circ$) then the proposed method provides good estimation, i.e. the average error is within 1 pixel. When robot rotation

Intrinsic Parameters					World/Camera Transformation					
	k_u	k_v	u_0	v_0	α	β	γ	t_x	t_y	t_z
N	Initial Estimate				Initial Estimate					
1	13829	13850	25.0	21.0	-0.0472	-0.7553	-0.2787	-32	-118	968
2	13686	13765	-5.4	11.3	-0.1557	-0.7278	-0.3745	-25	-189	938
3	13950	13865	18.8	-3.7	-0.1166	-0.6468	-0.3495	-92	-189	915
	Final Estimate				Final Estimate					
1	13834	13835	8.8	5.5	-0.0473	-0.7552	-0.2787	-26	-151	956

Robot/Camera Transformation						
	α	β	γ	t_x	t_y	t_z
Initial Estimate	0.00056	0.26274	-0.00031	203.4	198.7	164.6
Final Estimate	-0.00014	0.26173	-0.00041	200.5	201.3	161.4

Table 1: Result of the intrinsic and the extrinsic parameters, N is the index of camera locations

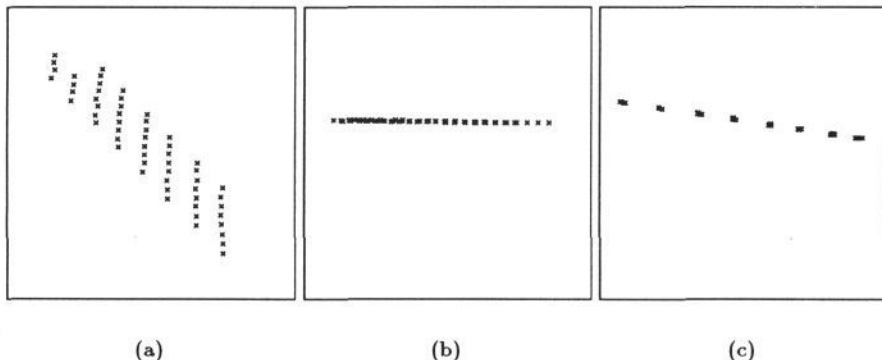


Figure 3: Recovered co-planar points at different viewpoints

angles are small the noise has a strong effect on the estimation of the camera-to-robot transformation. This is reasonable because small rotations mean that the coefficient matrixes of equations (10) and (11) are close to linearly dependent.

Experiments on real data were performed with the Surrey active stereo head. Calibration for one camera is considered. The robot can perform pure rotations with respect to the robot coordinate system around its x -axis and the y -axis independently. The motion of the robot is driven by a stepper motor with accuracy of 140 steps per degree i.e. $0.0072^\circ/\text{step}$. Two robot motions were considered: a 5° rotation around the x -axis followed by a 5° rotation around the y -axis. The camera viewed a calibration chart consisting of a regular array of black squares. The corners of the squares are chosen as the reference points and their corresponding image points are detected using standard edge detection and corner finding method. There is an uncertainty of 1 – 2 pixels for the locations of the corners.

Table 1 shows the result of the intrinsic and extrinsic parameters for a set of real data using the closed-form method. It can be seen that the estimated intrinsic parameters differ for different locations of the camera. The average value of the intrinsic parameters was used as the initial estimate to the proposed evolutionary algorithm. The estimated camera-to-world and robot-to-camera transformations

were used as the initial extrinsic parameter estimates. Table 1 shows the calibration result using the evolutionary algorithm. The root mean square distances using these final estimates were 0.7, 1.0 and 0.7 pixels, respectively. Thus the evolutionary algorithm has considerably improved the estimation of the intrinsic and extrinsic parameters.

In order to show the accuracy of calibration parameters in inferring 3D information, both cameras of the stereo head were calibrated using the proposed method. The transformation between cameras was then calculated by combining the extrinsic parameters of each camera. Figure 3 shows the projections of the recovered co-planar reference points from different viewpoints. It can be seen that the recovered reference points are located on the same plane, which implies that the estimation of intrinsic and extrinsic parameters is reasonably good.

7 Conclusions

It has been shown how to estimate the intrinsic and extrinsic parameters of a camera mounted on a robot arm within an optimization framework. The method fully exploits information from two accurate movements of the robot arm. An evolutionary algorithm was proposed to find an optimal calibration estimate. Experimental results obtained on both synthetic and real data demonstrate its usefulness.

Acknowledgement: This work was carried out under the ESPRIT project 7108 "Vision as Process II".

References

- [1] L. Dron. Dynamic camera self-calibration from controlled motion sequences. In CVPR1993, pages 501–506.
- [2] F. Du and M. Brady. Self-calibration of the intrinsic parameters of cameras for active vision systems. In CVPR1993, pages 477–482.
- [3] O.D. Faugeras, Q.T. Luong, and S.J. Maybank. Camera self-calibration: Theory and experiments. In ECCV199, pages 321–334.
- [4] O.D. Faugeras and G. Toscani. The calibration problem for stereo. In CVPR1986, pages 15–20. IEEE, 1986.
- [5] D. B. Fogel. *System Identification through Simulated Evolution: A Machine Learning Approach to Modeling*. Ginn Press, Needham Heights, MA, 1991.
- [6] D. B. Fogel. An introduction to simulated evolutionary optimization. *IEEE Transactions on Neural Network*, 5(1):3–14, 1994.
- [7] X. Qi and F. Palmieri. Theoretical analysis of evolutionary algorithms with an infinite population size in continuous space part I: Basic properties of selection and mutation. *IEEE Transactions on Neural Network*, 5(1):102–119, 1994.
- [8] Y.C. Shiu and S. Ahmad. Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form $AX = XB$. *IEEE Trans. Robotics and Automation*, 5:16–29, 1989.
- [9] R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 3:323–344, 1987.
- [10] R. Y. Tsai. Synopsis of recent progress on camera calibration for 3D machine vision. In O. Khatib et al., editor, *The Robotics Review*. MIT Press, 1989.
- [11] R.Y. Tsai and R.K. Lenz. A new technique for fully autonomous and efficient 3D robotics hand/eye calibration. *IEEE Trans. Robotics and Automation*, 5:345–358, 1989.
- [12] C.C. Wang. Extrinsic calibration of a vision sensor mounted on a robot. *IEEE Trans. Robotics and Automation*, 8:161–175, 1992.