

An Approach to Improve Online Hand-Eye Calibration

Fanhuai Shi, Jianhua Wang, and Yuncai Liu

Inst. Image Processing & Pattern Recognition, Shanghai Jiao Tong University
Shanghai 200030, P.R. China
{fhshi, jian-hua.wang, whomliu}@sjtu.edu.cn

Abstract. Online implementation of robotic hand-eye calibration consists in determining the relative pose between the robot gripper/end-effector and the sensors mounted on it, as the robot makes unplanned movement. With noisy measurements, inevitable in real applications, the calibration is sensitive to small rotations. Moreover, degenerate cases such as pure translations are of no effect in hand-eye calibration. This paper proposes an algorithm of motion selection for hand-eye calibration. Using this method, not only can we avoid the degenerate cases, but also the small rotations to decrease the calibration error. Thus, the procedure lends itself to an online implementation of hand-eye calibration, where degenerate cases and small rotations frequently occur in the sampled motions. Simulation and real experiments validate our method.

1 Introduction

The calibration of robotic hand-eye relationship is a classical problem in robotics, which concerns the relative position and orientation between the robot gripper/end-effector and sensors, such as a camera mounted rigidly on the gripper. Hand-eye calibration is an important task for robot applications involving 3-D vision measurement, visual servoing, and tactile sensing.

On this problem, much work has been done by solving the homogeneous transformation equation $\mathbf{AX}=\mathbf{XB}$ [1]-[7], which states that when the robot gripper undergoes a rigid motion \mathbf{A} and the corresponding camera motion is \mathbf{B} , the two motions are conjugated by the hand-eye transformation \mathbf{X} . Malm and Heyden [8] perform hand-eye calibration using normal derivatives of the image flow field instead of traditional point correspondences. And some methods simultaneously calibrate the camera and hand-eye relationship [9]-[11].

All the mentioned work requires an iterative approach, leading to offline least-square solutions. Angeles et al. [12] and Andreff et al. [13][14] first proposed the technique of online implementation of hand-eye calibration, allowing reducing human supervision required in classical calibration methods. Based on the linear invariants of rotation matrices, the former proposed a solution by recursive linear least squares. The latter derived a new linear formulation of the hand-eye problem, inspired by Sylvester equation: $\mathbf{UV}+\mathbf{VW}=\mathbf{T}$. Moreover, this method is extended to work with pose estimation by structure from motion. Therefore, it allows to get rid of the target objects required by standard approaches and use unknown scenes instead.

Note that whichever method is used, the hand-eye calibration problem intrinsically requires at least two motions with non-parallel rotation axes. This has been shown

algebraically [2] and geometrically [3]. So, hand-eye transformation from two independent motions sometimes can not be obtained when there exists a degenerate case such as pure translation or pure rotation etc., detailed algebraic analysis of the results for two independent motions refer to [14]. In addition, Tsai and Lenz made detailed analysis on critical factors affecting the accuracy of hand-eye calibration and got five observations, see [2].

During the online implementation of hand-eye calibration, hand-eye transformation is computed from continuously sampled motions, which are unplanned. So the problem is that, the sampled motions may be pure translation, pure rotation or their combinations, from which we only get partial calibration [14]. Furthermore, sampled motions may include a rotation with a tiny rotation angle, or the angle between two rotation axes is very small, both of which will bring on a large error in noisy measurements according to *observation 1* and *observation 2* in [2]. On the other side, according to *observation 4* in [2], small translation in gripper motion will be much useful in calibration.

In this paper we propose an algorithm of motion selection for online hand-eye calibration, which can not only avoid the degenerate cases in calibration, but also decrease the calibration error by selecting appropriate motion pairs. The remainder of this paper decomposes as follows. Section 2 describes the objective problem. Then, given golden rules for motion selection, the detailed algorithm of motion selection for online hand-eye calibration is presented in Section 3. Section 4 conducts some simulated and real experiments to validate the proposed algorithm.

2 Problem Formulation

We use upper-case boldface letters for matrices, e.g. \mathbf{X} , and lower-case boldface letters for 3-D vectors, e.g. \mathbf{x} . The angle between two vectors is denoted by $\angle(\mathbf{x}, \mathbf{y})$. $\|\cdot\|$ means the Frobenius norm of a vector or a matrix. Rigid transformation is represented with a 4×4 homogeneous matrix \mathbf{X} , which is often referred to as the couple (\mathbf{R}, \mathbf{t}) . At the i -th measurement, the camera pose with respect to reference object is denoted by 4×4 homogeneous matrix \mathbf{P}_i , and the recorded gripper pose relative to robot base is homogeneous matrix \mathbf{Q}_i .

The usual way to describe the hand-eye calibration is by means of homogeneous transformation matrices. We denote the transformation from gripper to camera by $\mathbf{X}=(\mathbf{R}_x, \mathbf{t}_x)$, the i -th motion matrix of the gripper by $\mathbf{A}_i=(\mathbf{R}_{a,i}, \mathbf{t}_{a,i})$, and the i -th motion matrix of the camera by $\mathbf{B}_i=(\mathbf{R}_{b,i}, \mathbf{t}_{b,i})$. The motion of the gripper is computed directly from the joint-angle readings by simple composition:

$$\mathbf{A}_i = \mathbf{Q}_i^{-1} \mathbf{Q}_{i+1} \quad (1)$$

With the known intrinsic camera parameters, the camera poses \mathbf{P}_i and \mathbf{P}_{i+1} relative to reference object are estimated, then the motion of the camera can also be determined by

$$\mathbf{B}_i = \mathbf{P}_i^{-1} \mathbf{P}_{i+1} \quad (2)$$

When dealing with an unknown scene (such as the scene without special calibration object), we can use a structure from motion algorithm [13][14] to estimate the camera

motion directly. Thus, the well-known hand-eye equation of $A_i X = X B_i$ can be established [1][2], which yields one matrix and one vector equation:

$$R_{a,i} R_x = R_x R_{b,i} \quad (3)$$

$$(R_{a,i} - I) t_x = R_x t_{b,i} - t_{a,i} \quad (4)$$

As we know that two motions with non-parallel rotation axes are necessary to determine the hand-eye transformation, so another group of motion equations should be obtained,

$$R_{a,i+1} R_x = R_x R_{b,i+1} \quad (5)$$

$$(R_{a,i+1} - I) t_x = R_x t_{b,i+1} - t_{a,i+1} \quad (6)$$

Eqs. (3)-(6) can be combined into the following linear form [13][14]:

$$\begin{pmatrix} I_9 - R_{a,i} \otimes R_{b,i} & 0_{9 \times 3} \\ I_3 \otimes (t_{b,i})^T & I_3 - R_{a,i} \\ I_9 - R_{a,i+1} \otimes R_{b,i+1} & 0_{9 \times 3} \\ I_3 \otimes (t_{b,i+1})^T & I_3 - R_{a,i+1} \end{pmatrix} \begin{pmatrix} \text{vec}(R_x) \\ t_x \end{pmatrix} = \begin{pmatrix} 0_{9 \times 1} \\ t_{a,i} \\ 0_{9 \times 1} \\ t_{a,i+1} \end{pmatrix} \quad (7)$$

where the \otimes product is the Kronecker product and operator vec reorders (one line after the other) the elements of a $m \times n$ matrix into a mn vector. Thus, given a pair of motions, hand-eye transformation can be linearly computed.

In practice, however, the movements of robot gripper vary with applications, not for hand-eye calibration. So, translations and small rotations usually occur in the sampled data, from which we can only get partial calibration or calibrate with big error. In order to make the online calibration practicable, we propose an approach of motion selection.

3 Motion Selection for Online Hand-Eye Calibration

We firstly give the golden rules for motion selection, and then describe two kinds of motion selection methods for online hand-eye calibration according to configurations.

3.1 Golden Rules

As a rotation matrix R can be expressed as a rotation around a rotation axis k by an angle θ , the relations between θ , k and R are given by Rodrigues theorem. Moreover, R_a and R_b have the same angle of rotation [1]. We can rewrite R_a and R_b as $\text{Rot}(k_a, \theta)$ and $\text{Rot}(k_b, \theta)$ respectively. Our aim in this paper is to sequentially find the pairs of consecutive motions (A_i, B_i) and (A_{i+1}, B_{i+1}) for hand-eye computation by motion selection from sampled motion series. In this procedure, we should obey the following golden rules inferred from Tsai and Lenz's observations [2], that is

Rule 1: Try to make $\angle(k_{a,i}, k_{a,i+1})$ (which is equal to $\angle(k_{b,i}, k_{b,i+1})$ [2]) large, the minimal threshold is set to be α .

Rule 2: Try to make θ_i large, the minimal threshold is β .

Rule 3: Try to make $\|t_{a,i}\|$ small, the maximal threshold is d .

(As *observation* 3 and 5 in [2] are relative to the system configuration, we don't consider them in this paper). Thus, we can avoid degenerate motions and small rotations in solving hand-eye relationship to obtain higher calibration accuracy.

3.2 Motion Selection Algorithms

We denote the i -th sampled hand-eye pose and motion by (P_i, Q_i) and (A_i, B_i) respectively in this section. (α, β, d) are threshold factors determined by experience and the detailed definitions are given in Section 3.1. (A', B') and (A'', B'') are selected motion pairs for calibration (see Fig. 1). For A' and A'' , the rotation axis, rotation angle and translation are denoted by (k'_a, θ', t'_a) , (k''_a, θ'', t''_a) respectively.

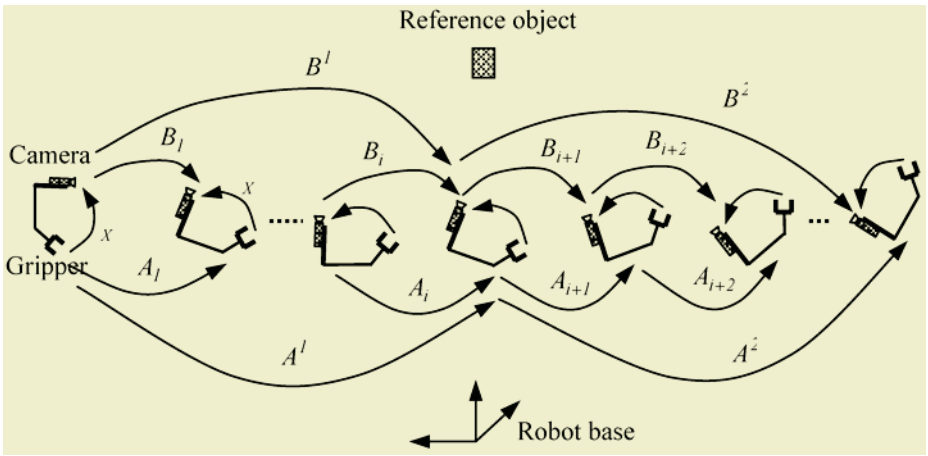


Fig. 1. Algorithm of motion selection for online hand-eye calibration.

Firstly, we consider the case when the camera pose P_i in each time instant can be estimated during the procedure.

At the beginning of the calibration process, we need to estimate (A', B') . The (A', B') is recovered from (P_1, Q_1) and (P_2, Q_2) according to Eqs. (1)-(2). If $\theta' \geq \beta$ and $\|t'_a\| \leq d$, we claim that the (A', B') has been found. Or else, we continue to compute (A', B') from (P_1, Q_1) and (P_3, Q_3) and judge the value θ' and $\|t'_a\|$ in the same way as before. Repeat this procedure until θ' and $\|t'_a\|$ fulfill the given conditions. Here, we assume that the first (A', B') is estimated from (P_1, Q_1) and (P_i, Q_i) . After (A', B') has been found, another motion pair (A'', B'') can be sought starting from (P_i, Q_i) and (P_{i+1}, Q_{i+1}) in the similar way as that of (A', B') , but the constrained conditions are changed to be $\theta'' \geq \beta$, $\|t''_a\| \leq d$ and $\angle(k'_a, k''_a) \geq \alpha$. When both motion pairs are found, we can make one calibration by solving Eq. (7).

In the next calibration, we take the last motion pair (A'' , B'') as the new motion pair (A' , B'), and then continue to seek for new (A'' , B'') from the successive sampled series and make a new hand-eye calibration in the same way as before.

We have thus derived an online hand-eye calibration algorithm based on iterative motion selection:

Algorithm I

1. $i \leftarrow 2$;
2. $A' = Q_1^{-1} Q_i$, $B' = P_1^{-1} P_i$;
3. Compute θ' and t'_a from A' ;
4. If $\theta' \geq \beta$ and $\|t'_a\| \leq d$, then go to 6;
5. $i \leftarrow i + 1$, go to 2; (Sample one more motion)
6. $j \leftarrow i + 1$; (Begin to search for A'')
7. $A'' = Q_i^{-1} Q_j$, $B'' = P_i^{-1} P_j$;
8. Compute $\angle(k'_a, k''_a)$, θ'' and t''_a from A' and A'' ;
9. If $\angle(k'_a, k''_a) \geq \alpha$ and $\theta'' \geq \beta$ and $\|t''_a\| \leq d$, then go to 11;
10. $j \leftarrow j + 1$, go to 7; (Sample one more motion)
11. Make one hand-eye calibration by solving Eq.(7);
12. $A' \leftarrow A''$, $B' \leftarrow B''$;
13. $i \leftarrow j$, $j \leftarrow j + 1$, go to 7 for next calibration.

When dealing with an unknown scene, pose estimation has to be replaced by an algorithm of structure from motion. Therefore, instead of the camera motion B_i from pose estimation, a scaled motion $\tilde{B}_i = (R_{b,i}, t_{b,i} / \|t_{b,i}\|)$ can be estimated, details refer to [13][14]. In this case, B' and B'' can be computed by motion synthesis described in the following.

Considering two consecutive camera motions B_1 and B_2 , we can get the equivalent motion B as follows:

$$B = \begin{bmatrix} R_b & t_b \\ \theta^T & 1 \end{bmatrix} = \begin{bmatrix} R_{b,1} & t_{b,1} \\ \theta^T & 1 \end{bmatrix} \begin{bmatrix} R_{b,2} & t_{b,2} \\ \theta^T & 1 \end{bmatrix} \quad R_b = R_{b,1} R_{b,2}, t_b = R_{b,1} t_{b,2} + t_{b,1} \quad (9)$$

For convenience, we use symbol “ \oplus ” in the following to denote the motion synthesis operation, thus Eq. (9) can be briefly rewritten as:

$$B = B_1 \oplus B_2 \quad (10)$$

The corresponding algorithm for online hand-eye calibration from unknown scenes is as follows:

Algorithm II

1. $i \leftarrow 1$;
2. $A' = Q_1^{-1} Q_2$, $B' \leftarrow B_1$;

3. Compute θ' and t'_a from A' ;
4. If $\theta' \geq \beta$ and $\|t'_a\| \leq d$, then go to 7;
else, $i \leftarrow i + 1$; (Sample one more motion)
5. $A' = Q_1^{-1} Q_{i+1}$, $B' \leftarrow B' \oplus B_i$, go to 3;
6. $j \leftarrow i + 1$, $i \leftarrow i + 1$; (Begin to search for A'')
7. $A'' = Q_i^{-1} Q_{j+1}$, $B'' \leftarrow B_j$;
8. Compute $\angle(k'_a, k''_a)$, θ'' and t''_a from A' and A'' ;
9. If $\angle(k'_a, k''_a) \geq \alpha$ and $\theta'' \geq \beta$ and $\|t''_a\| \leq d$, then go to 11;
10. $j \leftarrow j + 1$, go to 7; (Sample one more motion)
11. Make one hand-eye calibration by solving Eq.(7);
12. $A' \leftarrow A''$, $B' \leftarrow B''$;
13. $i \leftarrow j + 1$, $j \leftarrow j + 1$, go to 7 for next calibration.

In *Algorithm I* and *Algorithm II*, motion pairs (A', B') and (A'', B'') are sought by an iterative procedure. To prevent the iteration from performing too many times, a threshold could be set to control the number of iterations.

4 Experiments

In this section, experiments on synthetic data and real scenes are carried out to validate our algorithm, where we adopt *Algorithm I*. To compare its performance, we make an additional experiment by directly solving Eq. (7) without any motion selection. In the following graphs, we denote the proposed method by “new method” and the direct approach by “traditional method”.

The motivation of the synthetic experiments is to test the performance of the new method by varying three threshold factors. The simulation is conducted as follows: we establish a consecutive motion series with 1000 hand stations Q_i . We add uniformly distributed random noise with relative amplitude of 0.1% on the rotation matrix and of 1% on the translation vector. We assume a hand-eye setup and compute the camera pose P_i , to which we also add uniformly distributed random noise as before.

For each factor, we calibrate the hand-eye relationship with different threshold while fixing the other two factors unchanged. In this way, we get the estimated rotation matrix \hat{R} and translation vector \hat{t} . To qualify the results, we take RMS of the errors in the rotation matrix and the RMS of the relative errors $\|t - \hat{t}\| / \|t\|$ in the translation, which are customary error metrics in the literature [2][6][7]. Fig. 2 shows the simulation results. On one hand, the motion selection approach exhibits better behavior than the method without motion selection when there exist noisy measurements. The characteristics of error varying with different threshold validate the rules in Section 3.1.

From the results of the experiment, we can find that the impact of variation of d on the error is unnoticeable. So, in the second experiment, we test the behavior of the new algorithm with fixed d and simultaneous variation in (α, β) . Fig. 3 shows the result of

the second test, from which we can see that the calibration error is hard to be decreased when factors (α , β) are great than 30° . Note that in practice, when the two threshold increase, the average number of motions in each selection will increase, which will decrease the performance of system in real-time applications.

We also demonstrate the foregoing algorithm on a real setup composed of an infrared marker and a pair of CCD cameras which are attached to the end-effector of a 6-DOF robot (MOTOMAN CYR-UPJ3-B00), see Fig. 4(a). After the stereo rig is precisely calibrated, we mount an infrared filter on each camera. Thus, we get an infrared navigation system with stereoscopic vision. Without loss of generality, we compute the hand-eye transformation between the left camera and the gripper.

In the test, the robot is fixed on a workbench and the moving cameras observe the static infrared mark. We randomly move the gripper to 24 locations with different relative rotation or/and translation controlled by program; repeat the test for 10 times. For each time instant, gripper pose Q_i can be read from robot controller and pose of reference object P_i relative to the camera can be solved by binocular vision. We perform the hand-eye calibration using the same methods as in the synthetic experiments. In the test with motion selection, the values of three factors need to be set by experience at first.

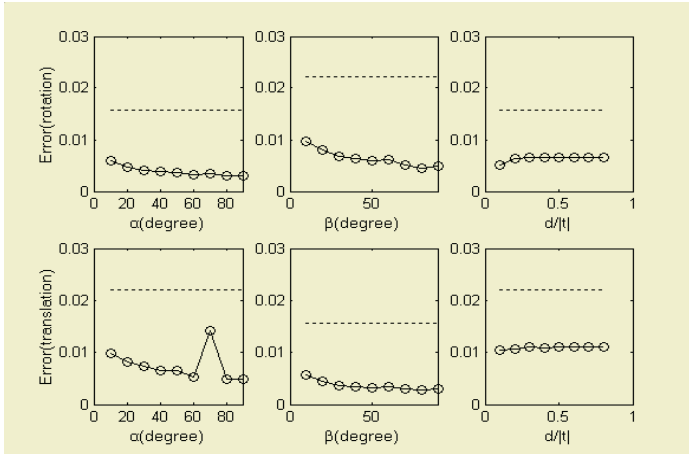


Fig. 2. Performances of the new algorithm with variation in α , β or d , which are compared to the traditional method without motion selection. The RMS rotation error is shown on the top and the RMS relative translation error is shown on the bottom, where the solid with label “o” denote new method and the dotted denote traditional method.

As no ground-truth value is available for comparison in such experiments, we compared $A_i X$ and XB_i for each motion i , and then gathered all these errors into RMS errors. This kind of measurement was also adopted by Andreff et al.[13]. The results of the experiments are shown in Fig 4(b), where the calibration times of traditional method does not include degenerate case. From Fig 4(b) we can see that, the average RMS error of proposed method is much lower than that of the traditional method.

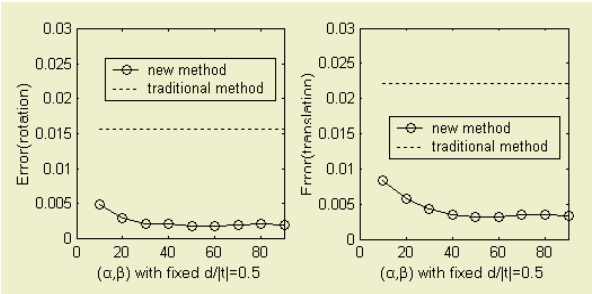


Fig. 3. Performance of the new algorithm with fixed d and simultaneous variation in (α, β) , which is compared to the traditional method without motion selection. The RMS rotation error is shown on the left and the RMS relative translation error is on the right.

5 Conclusion and Open Issue

In this paper, we propose an algorithm of motion selection for online hand-eye calibration, which can not only avoid the degenerate cases in hand-eye calibration, but also try to decrease the calibration error by selecting appropriate motion pairs. Experimental results from simulated data and real setup show that the method can greatly decrease the error of online hand-eye calibration.

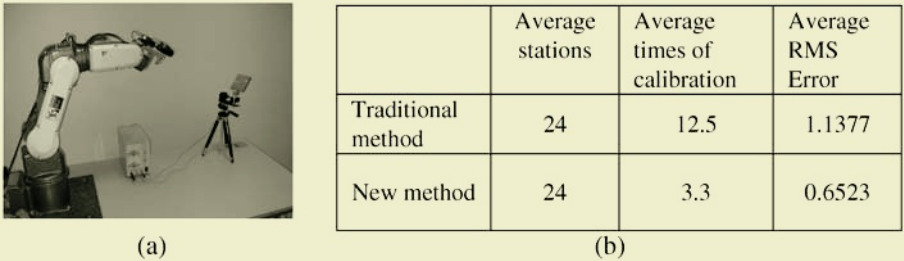


Fig. 4. Real experiment for online hand-eye calibration. (a). Experimental system. (b) Results of the real experiment.

However, the characteristics of gripper motion in different applications are diverse, so the threshold of the three factors should be chosen by experience according to the real setup. Open issue includes the following problem: “What is the optimal motion selection for online hand-eye calibration?”

References

- 1. Y. C. Shiu and S. Ahmad, "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form $AX = XB$," *IEEE Trans. Robot. Automat.*, vol. 5, pp. 16-29, Feb. 1989.
- 2. R. Y. Tsai and R. K. Lenz, "A new technique for fully autonomous and efficient 3d robotics hand/eye calibration", *IEEE Trans. Robot. Automat.*, vol. 5, pp. 345-358, 1989.

3. H. Chen. "A screw motion approach to uniqueness analysis of head-eye geometry". in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Maui, Hawaii, USA, pp. 145-151, June 1991.
4. C. Wang. "Extrinsic calibration of a robot sensor mounted on a robot". *IEEE Trans. Robot. Automat.*, 8(2):161-175, Apr. 1992.
5. Hanqi Zhuang and Yui Cheung Shiu, "A Noise-Tolerant Algorithm for Robotic Hand-Eye Calibration With or Without Sensor Orientation Measurement". *IEEE Trans. on System, Man and Cybernetics*, 23(4):1168-1175, 1993.
6. R. Horaud and F. Dornaika, "Hand-eye calibration," *Int. J. Robot. Res.*, 14(3):195-210, 1995.
7. K. Daniilidis. "Hand-eye calibration using dual quaternions". *Int. J. Robot. Res.*, 18(3):286-298, 1999.
8. H. Malm, A. Heyden,. "A new approach to hand-eye calibration", in *Proc. 15th Int. Conf. Pattern Recognition*, Vol. 1 , pp. 525-529, Sept. 2000.
9. S. Ma, "A self-calibration technique for active vision systems", *IEEE Trans. Robot. Automat.*, 12(1):114-120, Feb. 1996.
10. G. Wei, K. Arbter, and G. Hirzinger. "Active self-calibration of robotic eyes and hand-eye relationships with model identification". *IEEE Trans. Robot. Automat.*, 14(1):158-166, 1998.
11. H. Malm, and A. Heyden, "Simplified intrinsic camera calibration and hand-eye calibration for robot vision", in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, Vol.1, pp. 1037-1043, Oct.2003.
12. J. Angeles, G. Soucy and F. P. Ferrie, "The online solution of the hand-eye problem", *IEEE Trans. Robot. Automat.*, vol. 16, pp. 720-731, Dec. 2000.
13. N. Andreff, R. Horaud and B. Espiau, "On-line hand-eye calibration", in *Proc. Int. Conf. on 3-D Digital Imaging and Modeling*, pp. 430 - 436, Oct. 1999.
14. N. Andreff, R. Horaud, and B. Espiau, "Robot hand-eye calibration using structure-from-motion", *Int. J. Robot. Res.*, 20(3):228-248, 2001.