Proceedings of the 2010 IEEE
International Conference on Robotics and Biomimetics
December 14-18, 2010, Tianjin, China

# Robot Head-Eye Calibration Using the Minimum Variance Method

Sin-Jung Kim, Mun-Ho Jeong, Joong-Jae Lee, Ji-Yong Lee, Kang-Geon Kim, Bum-Jae You and Sang-Rok Oh

*Abstract*— **This paper presents a calibration method for a robot Head-Eye system which can be used in a humanoid robot vision system. This method estimates the transformation between the robot coordinate and the camera coordinate. The calibration procedure uses visual measurements and kinematic information as the inputs of nonlinear optimization. For this calibration method, an objective function for the optimization, that uses the *Minimum Variance* method is defined. The procedure of this method is very simple and intuitive. Besides, the result of this method is considerably precise and robust against the noisy environment. The performance is compared with earlier approaches and the results of simulation and actual experiment are followed.**

## I. Introduction

The calibration procedure for the transformation between a robot's end-effector and a camera mounted on the robot is very important. The relationship obtained from the calibration provides more diverse possibilities for the robot. A robot can receive information pertaining to objects in the robot's environment from its camera. Using this information, the robot can perform more precise and various tasks. This paper presents a robot Head-Eye calibration method to assign these benefits to a robot. A nonlinear optimization method is used for this calibration and the objective function is based on the *Minimum Variance* method which is quite intuitive and accurate.

This calibration method focus on the relation between robot and camera, and the camera is supposed well calibrated. The 3D positions of target are used for input of this method and these can be calculated easily by diverse ways, i.e. triangulation using calibrated stereo camera, 2D-to-2D matching using single camera([10]), IR sensor, laser sensor, etc. In this paper, the humanoid robot Mahru-Z of KIST is used in an experimental routine. This robot has a head-mounted stereo camera - see Fig. 1.

This work was supported in part by Ministry Knowledge Economy(MKE) & Institute for Information Technology Advancement(IITA) through IT Leading R&D support Project.

S. J. Kim is with the Korea University of Science and Technology(UST) and the Cognitive Robotics Center, Korea Institute of Science and Technology(KIST), Korea. (e-mail: caesarkim@kist.re.kr)

M. H. Jeong is with the Lab. of Robot Vision & Intelligence School of Robot Kwangwoon Univ., Korea. (e-mail: mhjeong@kw.ac.kr)

J. J. Lee is with the Machine Intelligence Group, Information & Technology Lab, LGE Advanced Research Institute, Korea. (e-mail: joongjae.lee@lge.com)

J. Y. Lee, K. G. Kim, B. J. You and S. R. Oh are with the Cognitive Robotics Center, Korea Institute of Science and Technology(KIST), Korea. (e-mail: wisrage@kist.re.kr danny@kist.re.kr ybj@kist.re.kr sroh@kist.re.kr)

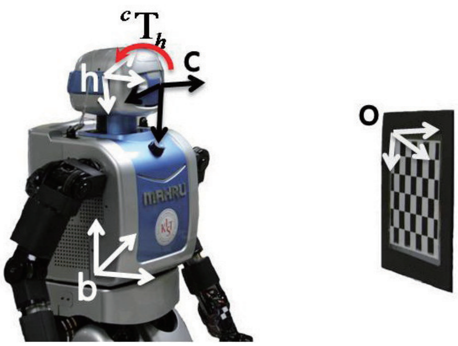


Fig. 1. The humanoid robot named Mahru-Z in Korea Institute of Science and Technology(KIST).

### A. Related works

In the past, there were many approaches to solve the calibration problem. These earlier solutions can be divided into two categories.

*1) The formulation $AX = XB$ :* The first class uses the formulation $AX = XB$ (henceforth "$XX$" for this). In this formula, $A$ is the transformation between the camera frames during the robot motions and $B$ is the one between the robot's end-effector frames, which are identical to the camera's frames. $X$ is the desired solution; the transformation between the camera frame and the robot's end-effector frame. The most popular and classical studies in this area were done by Tsai and Lenz [1] and Shiu and Ahmad [2]. Tsai and Lenz solved the problem of rotation first, and then translation using a least square fitting approach. Shiu and Ahmad provided a solution similar to [1] by using angle-axis representation based on eigenvalues and eigenvectors. Park and Martin [5] used the Euclidean group and nonlinear optimization method. Horaud and Dornaika [4] proposed a nonlinear optimization system for both rotation (using a quaternion formulation) and translation one-to-one. Fassi and Legnani [6] provided a solution based on a geometrical approach. They solved the rotation part first and, then the translation part using a least square method. And singular cases were also discussed.

*2) The formulation $AX = ZB$ :* The other category uses the formulation $AX = ZB$ (henceforth "$XZ$" for this). In this case, $A$ is the transformation between the target object frame and the camera frame and $B$ is that between the robot base frame and end-effector frame. $X$ is equivalent to its value in the first category. The last variable $Z$ is also an unknown, which denotes the transformation between the target object frame and the robot base frame. Wang [8] presented several methods and gave results comparable to [1] and [2]. Dornaika

978-1-4244-9318-0/10/$26.00 © 2010 IEEE

and Horaud [7] presented a nonlinear optimization method based on a quaternion. They used nonlinear optimization for both rotation and translation parts through the one-to-one minimization of Frobenius norms. Remy, Dhome, Lavest and Daucher [9] suggested the nonlinear optimization of a sum of scalar products which estimated $X$ and $Z$ simultaneously.

Strobl and Hirzinger [3] presented an efficient calibration method which can be used either in the positioning of a robot manipulator $XZ$ or in its motion $XX$.

### B. The proposed method : Minimum Variance

The proposed method neither conform to the previous formulation $XX$ nor $XZ$. Instead of both formulation, the method proposed here uses more simple and intuitive approach to solve the problem.

During the robot motion, stereo camera can obtain the 3D position of points on the target object, and if the transformation obtained from the Head-Eye calibration is definitely accurate, the 3D position of the points for every robot motion can be projected onto the robot base frame identically and accurately. However, it is not always possible to derive the values of the transform matrix elements accurately, due to the presence of noise and uncertainty. Therefore, this paper proposes an approach of robot Head-Eye calibration using the *Minimum Variance* method by virtue of which the sum of variances of 3D positions for every point is minimized and the resultant matrix is close to the accurate value. A simulation and an actual experiment results prove that this method is simple and highly precise in comparison with previous methods in the noisy environment.

## II. Description of The Problem

Fig. 1 shows the coordinate systems of the calibration problem. There are a number of notations to be predefined that will be used in the formula of calibration process. The coordinate systems $c$, $h$, $b$ and $o$ represent the system of camera, robot head, robot base and target object, respectively. Let ${}^{c}\mathbf{T}_{h}$ be the transformation from the camera frame to the robot's head frame and it is deemed to be $X$ in the both $XX$ and $XZ$ formulation. Let ${}^{b}\mathbf{T}_{h}$ be the transformation from the robot's base frame to the head frame. This is identical to the matrix of $B$ in the $XZ$ formulation. Let ${}^{o}\mathbf{T}_{c}$ be the transformation from the target object frame to the camera frame. This is identical to $A$ in the $XZ$ formulation. There are many methods used to estimate this value. The classical approach was described in the study [10], and another study [11] oulines one of the most popular methods of using nonlinear optimization. Let ${}^{o}\mathbf{T}_{b}$ be the transformation from the target object frame to the robot base frame. It is deemed to be $Z$ in the $XZ$ formulation. Let ${}^{c_i}\mathbf{T}_{c_j}$ be the transformation between the camera frames during the robot motions; it is equal to $A$ in the $XX$ formulation. Let ${}^{h_i}\mathbf{T}_{h_j}$ be the transformation between the robot's end-effector frames, corresponding also to the camera's frames; it is equal to $B$ in the $XX$ formulation.

### A. $AX = XB$ approach

This approach follows the commutative equation as shown below.

$$\begin{tikzcd} {}^{c}\mathbf{p}_j \arrow[r, "{}^{c}\mathbf{T}_h"] & {}^{h}\mathbf{p}_j \\ {}^{c}\mathbf{p}_i \arrow[u, "{}^{c_i}\mathbf{T}_{c_j}"] \arrow[r, "{}^{c}\mathbf{T}_h"] & {}^{h}\mathbf{p}_i \arrow[u, "{}^{h_i}\mathbf{T}_{h_j}"'] \end{tikzcd} \qquad {}^{c_i}\mathbf{T}_{c_j}\ {}^{c}\mathbf{T}_{h} = {}^{c}\mathbf{T}_{h}\ {}^{h_i}\mathbf{T}_{h_j} \tag{1}$$

The component ${}^{x}\mathbf{p}_{n}$ of (1) denotes the 3D point vector of $n$-th motion with respect to (henceforth "w.r.t." for this) the $x$ frame. The transform matrix ${}^{c_i}\mathbf{T}_{c_j}$ $({}^{h_i}\mathbf{T}_{h_j})$ can be calculated after estimation (computation) of ${}^{o}\mathbf{T}_{c}({}^{b}\mathbf{T}_{h})$ for some different location of the camera (head) as the robot moves. Usually, ${}^{o}\mathbf{T}_{c}$ can be estimated by least square fitting or nonlinear optimization. For ${}^{b}\mathbf{T}_{h}$, it can be computed using the forward kinematics. There are many approaches that conform to this equation (e.g., [1]-[6]).

### B. $AX = ZB$ approach

As in the equation of $XZ$, it is possible to express this equation as shown below.

$$\begin{tikzcd} {}^{c}\mathbf{p} \arrow[r, "{}^{c}\mathbf{T}_h"] & {}^{h}\mathbf{p} \\ {}^{o}\mathbf{p} \arrow[u, "{}^{o}\mathbf{T}_{c}"] \arrow[r, "{}^{o}\mathbf{T}_b"] & {}^{b}\mathbf{p} \arrow[u, "{}^{b}\mathbf{T}_{h}"'] \end{tikzcd} \qquad {}^{o}\mathbf{T}_{c}\ {}^{c}\mathbf{T}_{h} = {}^{o}\mathbf{T}_{b}\ {}^{b}\mathbf{T}_{h} \tag{2}$$

An estimation or a calculation process of the components of the equation above is similar to that used in the $XX$ approach. There are also many approaches that conform to this equation (e.g., [2],[7]-[9]). (As an exception [2], there is a nonlinear optimization method that can be used for both $XX$ and $XZ$).

## III. Solution of The Problem

### A. Classical method : Tsai and Lenz

The most popular and classical method referred to by many was proposed by Tsai and Lenz [1]. This method conforms to the formulation $XX$. The main concept of this paper is its use of closed form least square fitting. They calculated the rotation part first, and then the translation part using the rotation result. The principal equations are shown below.

$$\mathbf{Skew}({}^{g_i}\mathbf{e}_{g_j} + {}^{c_i}\mathbf{e}_{c_j})\,{}^{c}\mathbf{e}_{g}{}' = {}^{c_i}\mathbf{e}_{c_j} - {}^{g_i}\mathbf{e}_{g_j} \tag{3}$$

$$({}^{g_i}\mathbf{R}_{g_j} - I)\,{}^{c}\mathbf{t}_{g} = {}^{c}\mathbf{R}_{g}\ {}^{c_i}\mathbf{t}_{c_j} - {}^{g_i}\mathbf{t}_{g_j} \tag{4}$$

In these equations, the component $\mathbf{e}$ is the eigenvector (or principal vector) of the rotation matrix; it is related to the eigenvalue of 1. The components ${}^{g_i}\mathbf{R}_{g_j}$, ${}^{g_i}\mathbf{e}_{g_j}$ and ${}^{g_i}\mathbf{t}_{g_j}$ denote the rotation matrix, the eigenvector of the rotation matrix and the translation vector from the $i$-th to $j$-th frame of the gripper (end-effector), respectively. ${}^{c_i}\mathbf{R}_{c_j}$, ${}^{c_i}\mathbf{e}_{c_j}$ and ${}^{c_i}\mathbf{t}_{c_j}$ are the ones from the $i$-th to $j$-th frame of the camera, respectively, similar to the previous ones. The components ${}^{c}\mathbf{R}_{g}$, ${}^{c}\mathbf{e}_{g}$ and ${}^{c}\mathbf{t}_{g}$ are the ones from the camera to the gripper frame. Tsai and Lenz computed the transformation matrix ${}^{o}\mathbf{T}_{c}$ by the method proposed by Tsai [10]. (For more details, see [1])

### B. Recent method : Strobl and Hirzinger

Strobl and Hirzinger [3] introduced a more efficient method that involved a stochastic model to estimate the transformation for both $XX$ and $XZ$ formulation. They defined a new objective function for the nonlinear optimization that optimally reduces actual system error and allows for a natural weighting of the rotational and translational components. The objective function for the nonlinear optimization is shown below.

$$\begin{aligned}\{{}^tT_c,\ {}^hT_o\}^\star &= \arg\min_{{}^tT_c,\ {}^hT_o}\left(\sum_{i=1}^{n}\frac{(O_i^{rot})^2}{{}^\star\sigma_{rot}^2}+\frac{(O_i^{tra})^2}{{}^\star\sigma_{tra}^2}\right) \\ &= \arg\min_{{}^tT_c,\ {}^hT_o}\left(\sum_{i=1}^{n}(O_i^{rot})^2+\frac{(O_i^{tra})^2}{({}^\star\sigma_{tra}/{}^\star\sigma_{rot})^2}\right)\end{aligned} \quad (5)$$

In this equation, the metric $O_i^{rot}$ and $O_i^{tra}$ represent the rotation and translation error of the $i$-th iteration. The items ${}^\star\sigma_{rot}^2$ and ${}^\star\sigma_{tra}^2$ are the $2^{nd}$ moments of the independent Gaussian probability density functions for rotation and translation error. ${}^\star\sigma_{tra}/{}^\star\sigma_{rot}$ is the position/orientation precision ratio.

For an estimation of the transformation between the target object frame and the hand mounted camera frame, the well-known method proposed by Zhang [11] was used. (For more details, see [3])

### C. The Proposed Method : Minimum Variance

Almost all of the methods that follow the formulation $XX$ or $XZ$ require the transform matrix ${}^o\mathbf{T}_c$ and ${}^c\mathbf{T}_h$ for their calibration process. The matrix ${}^o\mathbf{T}_c$ can be calculated using the 3D positions of target object w.r.t. camera. Once the 3D positions of target object is obtained, there are numerous ways to obtain the relationship([10], [11]). However, there are some rounding error can be exist in the process of calculation for the matrix. The proposed method does not use this matrix but uses the information of 3D position data of target object, directly. Hence, the possibility of rounding error which can be occured during the calculation of the matrix is eliminated. Furthermore, proposed method has minimized the number of variables to be estimated or calculated. This can reduce the effects of rounding error.

The procedure of the proposed method has five simple steps. In Fig. 2, a flowchart of the method is described. Some definition of notations for the proposed method are given below.

- ${}^b\mathbf{p}_{fp}$ : 3D position of the $p$-th point at the $f$-th robot motion w.r.t. the robot base frame.
- ${}^c\mathbf{p}_{fp}$ : 3D position of the $p$-th point at the $f$-th robot motion w.r.t. the camera frame.
- ${}^b\mu_p$ : Average 3D position of the $p$-th points projected to the robot base frame which were measured at every robot motion.
- ${}^b\mathbf{T}_h$ : Transform matrix from the robot base frame to the head frame.
- ${}^c\mathbf{T}_h$ : The transform matrix from the camera frame to the robot head frame which will be estimated by this calibration method.
- ${}^c\hat{\mathbf{T}}_h$ : An initial guess of the transform matrix ${}^c\mathbf{T}_h$

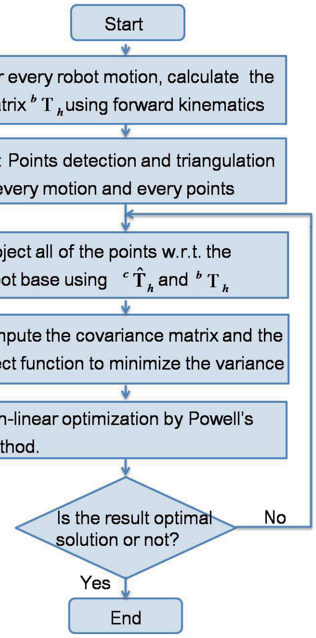


Fig. 2. A flowchart of the proposed method.

- $\mathbf{v}$ : A vector for the nonlinear optimization process.
- $\mathbf{v}^\star$ : The solution vector of the nonlinear optimization

*1) Transform matrix from Head to Base:* In this approach, $m$ motions of the robot and $n$ points in the target object are used. For every motion, the angles of robot neck joints can be measured using joint encoders. Subsequently, the forward kinematics formula, used in conjunction with these measurements, provides the transform matrix from the robot head frame to the base frame.

*2) Detection and triangulation:* At this step, it is necessary to compute the $m$ number of 3D positions for every $n$ point w.r.t. the camera frame for every robot motion. At first, a stereo camera captures a pair of images. A marker detection process of both images then follows. Once the image data consisting of the 2D positions of points in pixel units of each stereo image is detected, the next process is to project them onto the $m$ number of 3D positions for every $n$ point w.r.t. the camera frame by triangulation.

*3) Project onto the base:* For this process, two measured data and one presumed data are used. The transform matrix from the robot base frame to the head frame ${}^b\mathbf{T}_h$ and the $m$ number of 3D positions of $n$ point w.r.t. the camera frame for every robot motion ${}^c\mathbf{p}_{fp}$ denote the measured data, and the presumed data is the transformation between the camera frame and the robot head frame ${}^c\hat{\mathbf{T}}_h$. For the first iteration of the optimization process, the initial guess of the transform matrix can be used. After the first iteration, optimization process will update the transform matrix. In step 3, the 3D positions of all the points for every robot motion w.r.t. the camera frame, are projected to the ones w.r.t. the robot base frame using the two matrices described above.

$${}^b\mathbf{p}_{fp} = {}^b\mathbf{T}_{h_f}\ {}^c\hat{\mathbf{T}}_h^{-1}\ {}^c\mathbf{p}_{fp} \quad (\text{for } f=1\cdots m,\ p=1\cdots n) \quad (6)$$

After the projection process, there are $m$ number of 3D positions for every $n$ point, we have. Henceforth, these $m$ number of 3D positions of each point are referred to as "group".

*4) Variance of points:* This is the core step for the proposed method. In this step, the definition of an objective function, which will be used for the nonlinear optimization process of the Head-Eye calibration, is described. Through the projection process in step 3, there are $n$ number of point groups w.r.t. the robot base frame. However, the matrix used for projection is not the optimal solution but the initial guess. Moreover, feature detection process generates some noise. Hence, all $m$ number of position data for each group cannot be identical. Accordingly, an objective function is composed to optimize the matrix in spite of the noise. At this point, the calculation of the covariance matrix of 3D positions for every $n$ point group occurs.

$$\begin{gathered} C_{p_{xx}} = E[({}^b\mathbf{p}_{fp_x} - {}^b\mu_{p_x})({}^b\mathbf{p}_{fp_x} - {}^b\mu_{p_x})] \\ C_{p_{xy}} = E[({}^b\mathbf{p}_{fp_x} - {}^b\mu_{p_x})({}^b\mathbf{p}_{fp_y} - {}^b\mu_{p_y})] \\ \vdots \\ C_{p_{zz}} = E[({}^b\mathbf{p}_{fp_z} - {}^b\mu_{p_z})({}^b\mathbf{p}_{fp_z} - {}^b\mu_{p_z})] \\ \Sigma_p = \begin{bmatrix} C_{p_{xx}} & C_{p_{xy}} & C_{p_{xz}} \\ C_{p_{yx}} & C_{p_{yy}} & C_{p_{yz}} \\ C_{p_{zx}} & C_{p_{zy}} & C_{p_{zz}} \end{bmatrix} \quad (\text{for } p = 1 \cdots n) \end{gathered} \tag{7}$$

The covariance matrix $\Sigma_p$ has 3 different variance value of x, y and z in its diagonal, but the covariance matrix also has some other covariance value in its non-diagonal parts. Therefore, to minimize the variance, it is demanded to diagonalize the matrix. However, sum of the eigenvalues of covariance matrix is equal to the trace sum of diagonalized matrix. Thus, the objective function will be performed with the vector $\mathbf{v}$ to minimize the eigenvalue sum of covariance matrix of 3D positions for every $n$ group which has $m$ number of position data. Here vector $\mathbf{v}$ is composed by 6 components that represent the matrix ${}^c\mathbf{T}_h$, 3 for the of rotation part of the transform matrix ${}^c\mathbf{T}_h$ (e.g. yaw, pitch and roll), and another 3 for the translation vector of the same matrix (e.g. x, y and z). For every iteration of nonlinear optimization, the components of vector $\mathbf{v}$ is changed to minimize the variance, and finally, the optimized solution for matrix ${}^c\mathbf{T}_h$ is acquired. The objective function is defined as shown below.

$$\mathbf{v}^\star = \arg\min_{\mathbf{v}} \left( \sum_{p=1}^{n} \sum_{i=1}^{N} \lambda_{p_i} \right) \tag{8}$$

In this equation, $\lambda$ is the eigenvalue of $\Sigma_p$ and $N$ is the number of eigenvalues.

*5) Nonlinear optimization:* There are numerous optimization methods and each of them can be classified into a specific group depending on their characteristics.

The most universal one, originally formulated by Levenberg, Marquardt and Newton, can be categorized as a "Gradient search" method. This category of methods uses the differentiation of objective function to determine the direction of the optimization process. However, due to the necessities of the differentiation process, if there are many variables to be estimated, it may fail.

Powell and Brent formulated one of the "Direct search" methods. These methods do not require differentiation of objective function. Instead, they simply use the iteration memories of the optimization process. This greatly reduces the complexity of calculating the differentiation.

In this paper, due to the advantage of these "Direct search" methods, Powell's method is adopted for nonlinear optimization. This nonlinear optimization can provide the transform matrix ${}^c\mathbf{T}_h$ that minimize the variance of difference of points, which is the goal of this calibration.

## IV. Experiment Results

### A. Implementation Issue

In actual experiment environment, there are several noisy factors that causes error during the calibration process. First, the joint angles feedback from the encoder might have some error. This can make the forward kinematics formula incorrectly. Second, the possibility of difference between initial designed DH parameters and the ones of actual manufactured robot cannot be excluded (e.g. DH parameter $d$, $a$, $\alpha$ and $\theta$). Third, to calculate the 3D distance from the center of camera frame, feature detection algorithm (e.g. Harris Corner detection) is needed. Also, it may have unreliability base on the pixel detection error due to the blurred image, low resolution and so on. However, in case of the first, the error comes from robot joint encoder cannot influence the calibration. The resolution of encoder and gear ratio of drive in actual robot environment is so high, so the error by the encoder feedback can be ignored. The second factor cannot effect strongly to the calibration process, either. It was easily recognizable that both first and second case cannot have an effect on the calibration particulary, through the simulation that introduce the noise to both factors in the bounds of common sense. On the other hand, the factor of third case is obviously influence the calibration badly. Therefore, for the simulation experiment, noisy values of feature detection are adopted to create a condition similar to an actual experiment.

### B. Simulation

The information of the robot kinematics (i.e., the DH parameters and rotation angles for every motion of the robot), the intrinsic parameters of stereo camera and the target object (i.e., the relationship with the robot and the number of points in the target object and 3D position of each point w.r.t. the robot base) are predefined. There are 25 number of points in the target object and 25 number of robot motions are used for simulation. To make the noise environment close to an actual experiment, Gaussian noise is applied to the image data of the points. This can be considered as the measurement error that can be arises when detecting the points. Using these noise-introducing data points, a simulation of the proposed method was run, as described in Section III. For a comparison of performance levels, the method of Tsai and Lenz and the one of Strobl and Hirzinger are also simulated.

The performance of each method was compared in 2 different ways. For the first, Gaussian noise which has co-variance from $10^{-4}$ to $10^{-3}$ pixel is used for detailed contrast of matrices obtained from each method. Each method of calibration is performed 10 times and the standard deviations of obtained matrix are used for the comparison of performance levels. In the simulation, the correct answer of transform matrix between camera and robot head is known and it can be used for the mean value to calculate the standard deviation. Second, Gaussian noise which has covariance from $10^{-5}$ to $10^{-1}$ pixel is used for overall performance against the noise environment. In this task, point projection error is the object of comparison. After the calibration process, some points that were not used to obtain the calibration result, are employed for the measurements during some robot motion. The answer matrix of calibration can project these points into the robot base frame. The correct answer of projected points are already known in this simulation. Thus, the sum of distance error between projected points and correct values was compared for the performance.

Table I and Fig. 3 show the results of the first simulation and Table II and Fig. 4 shows the results of the second. In the table I and Fig. 3, $\sigma_{rot}$ of Strobl and Hirzinger is slightly smaller than *Minimum Variance*'s one. but both result are almost close to zero. Whereas in the case of $\sigma_{tra}$, the performance of *Minimum Variance* is much better than Strobl and Hirzinger's one. The values of Table II and Fig. 4 are the average value of the Euclidean distance error. The method formulated by Tsai and Lenz uses a closed form least square solution, which makes it more unstable than the nonlinear optimization method in a noisy environment, as used in this simulation. The method of Strobl and Hirzinger performs more robust; However, as mentioned in the previous section, both methods must estimate the transformation between the target object frame and the camera frame using the measurements of the 3D positions for every point. This estimation is repeated for every motion of the robot. In contrast, the proposed method does not require transform estimation, instead simply requiring the measurements of the 3D position of points w.r.t. the camera frame. Thus, the rounding error of this component is reduced. The result shows that the most stable result is that of the *Minimum Variance* method.

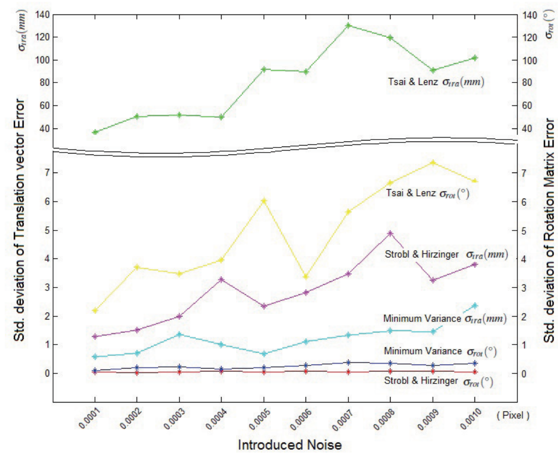


Fig. 3. Simulation result 1.

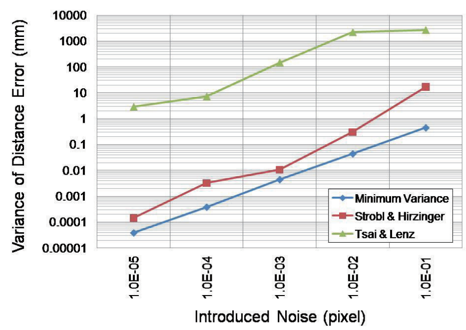


Fig. 4. Simulation result 2 in a Log scale graph.

## C. Robot Experiment

This experiment involved the humanoid robot Mahru-Z of KIST (Fig. 1). This robot has 2- DOF in the neck and uses a stereo camera (Bumble Bee II which provide precise intrinsic parameters.) on the head (the end-effector of the neck joint). In the robot experiment, the procedure of calibration was progressed following the proposed method same as the simulation. The method of Tsai and Lenz, and the one of Strobl and Hirzinger are also progressed for the comparison of performance. The contrast criteria of performance is identical to those ones of the second task of simulation. Once the estimation process of the transform matrix is completed, all of the points w.r.t. the camera frame (these are calculated by triangulation using 2D points of the stereo image obtained from the point detection) are projected onto the robot base frame using the obtained result. Unlike the simulation, the accurate values of the 3D positions of every point w.r.t. the robot base frame cannot be known. Therefore, evaluating the obtained transform matrix as to whether it is precise or not is challenging. However, similar to the simulation, if the obtained matrix is optimized solution, then the projected 3D position data of each point w.r.t. the robot base frame will have nearly identical values. Therefore, the variance of each point group will have the almost zero value. Hence, the sum of variances of 3D positions for every $n$ group which has $m$ number of position data, is used for the comparison of the result of methods in actual experiment. For the verification of the method, new feature points that were not used in the calibration process, are used.

Table III and Fig.5 show the results of the experiment. The values of Table III and Fig.5 are the sum of variances of 3D positions. Identical to the simulation result, concerning the nonlinear optimization, the proposed method and that of Strobl and Hirzinger are more precise and robust than the closed form least square solution. Moreover, the proposed

TABLE I
SIMULATION RESULT 1

| Noise | | 0.0001 | 0.0002 | 0.0003 | 0.0004 | 0.0005 | 0.0006 | 0.0007 | 0.0008 | 0.0009 | 0.0010 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Minimum Variance | $\sigma_{rot}$ (°) | 0.1026 | 0.2043 | 0.2262 | 0.1563 | 0.2048 | 0.2822 | 0.3688 | 0.3467 | 0.2762 | 0.3423 |
| | $\sigma_{tra}$ (mm) | 0.5826 | 0.6983 | 1.3525 | 0.9940 | 0.6766 | 1.1056 | 1.3334 | 1.4791 | 1.4530 | 2.3759 |
| Strobl and Hirzinger | $\sigma_{rot}$ (°) | 0.0340 | 0.0331 | 0.0458 | 0.0597 | 0.0555 | 0.0781 | 0.0492 | 0.0711 | 0.0817 | 0.0519 |
| | $\sigma_{tra}$ (mm) | 1.2894 | 1.5115 | 1.9870 | 3.2696 | 2.3506 | 2.8302 | 3.4750 | 4.8871 | 3.2515 | 3.8078 |
| Tsai and Lenz | $\sigma_{rot}$ (°) | 2.1793 | 3.7004 | 3.4820 | 3.9597 | 6.0271 | 3.3713 | 5.6417 | 6.6586 | 7.3648 | 6.6908 |
| | $\sigma_{tra}$ (mm) | 36.3923 | 50.2901 | 51.7221 | 50.1190 | 91.4947 | 90.0212 | 129.9815 | 119.3332 | 90.9091 | 101.7289 |

TABLE II
SIMULATION RESULT 2

| Noise | 0.00001 | 0.0001 | 0.001 | 0.01 | 0.1 |
|---|---|---|---|---|---|
| Minimum Variance | $3.82 \times 10^{-5}$ | $3.79 \times 10^{-4}$ | $4.40 \times 10^{-3}$ | $4.40 \times 10^{-2}$ | $4.48 \times 10^{-1}$ |
| Strobl and Hirzinger | $1.45 \times 10^{-4}$ | $3.24 \times 10^{-3}$ | $1.07 \times 10^{-2}$ | $3.03 \times 10^{-2}$ | 17.0538 |
| Tsai and Lenz | $2.90 \times 10^{-1}$ | $7.25 \times 10^{-1}$ | $1.49 \times 10^{2}$ | $2.20 \times 10^{3}$ | $2.68 \times 10^{3}$ |

[mm]

TABLE III
ROBOT EXPERIMENTAL RESULT

| Method | Minimum Variance | Strobl and Hirzinger | Tsai and Lenz |
|---|---|---|---|
| variance | $2.92 \times 10^{-3}$ | $1.61 \times 10^{-2}$ | 2.5633 |

[mm]

method outperforms the method of Strobl and Hirzinger. Since, the proposed method can reduce the calculation process as much as possible; the effect of noise and rounding error is minimized, and also, a precise result is obtained as well.

## V. CONCLUSION

This paper presents a method for robot Head-Eye calibration using the *Minimum Variance* method. This method provides a simpler process and more precise result in noisy environment as compared to the earlier methods. It neither follows the formulation of XX nor XZ, simply using the 3D information of points instead. After project the 3D position of points w.r.t. the camera to the robot frame, an objective function serves to minimize the sum of variances of 3D position data for every motion and every point in a nonlinear optimization process. With this method, unlikely others, it is not necessary to estimate the transform matrix between the target object frame and the camera frame. Rounding error exists in nearly all of the earlier approaches due to the use of estimation of that matrix. Additionally, with the many motions made by the robot for the calibration, the noises are increased and accumulated. The proposed method reduces the rounding error and is very robust in a noisy environment. And the entire process is very simple and intuitive. The results of a simulation and an actual experiment verified the robustness and accuracy of this method.

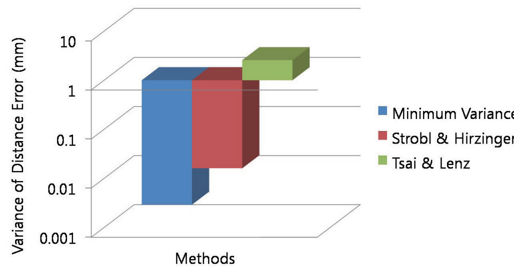


Fig. 5. Robot experimental results in a log scale graph.

## REFERENCES


[1] R. Y .Tsai and R. K. Lenz, "A New Technique for Fully Autonomous and Efficient 3D Robotic Hand/Eye Calibration", *IEEE Trans. On Robs. And Aut.*, vol. 5, no. 3, pp. 345-358, June 1989

[2] Y. C. Shiu and S. Ahmad, "Calibration of Wrist-Mounted Robotic Sensors by Solving Homogeneous Transform Equations of the Form AX=XB", *IEEE Trans. On Robs. And Aut.*, vol. 5, no. 1, pp. 16-29, February 1989

[3] K. H. Strobl and G. Hirzinger, "Optimal Hand-Eye Calibration", in *Proc. Of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems IROS*, Beijing, China, October 2006

[4] R. Horaud and F. Dornaika, "Hand-Eye Calibration", *Int. Journal of Robs. Research*, vol. 14, no. 3, pp. 195-210, June 1995

[5] F. Park and B. Martin, "Robot Sensor Calibration: Solving AX=XB on the Euclidean Group", *IEEE Trans. On Robs. And Aut.*, vol. 10, no. 5, pp. 717-721, October 1994

[6] I. Fassi and G. Legnani, "Hand to Sensor Calibration : A Geometrical Interpretation of the Matrix Equation AX=XB", *Journal on Robotics Systems*, vol. 22, no. 9, pp. 497-506, 2005

[7] F. Dornaika and R. Horaud, "Simultaneous Robot-World and Hand-Eye Calibration", *IEEE Trans. On Robs. And Aut.*, vol. 14, no. 4, pp. 617-622, August 1998

[8] C. C. Wang, "Extrinsic Calibration of a Vision Sensor Mounted on a Robot", *IEEE Trans. On Robs. And Aut.*, vol. 8, no. 2, pp 161-175, 1992

[9] S. Remy, M. Dhome, J. M. Lavest and N. Daucher, "Hand-Eye Calibration", in *Proc. Of the IEEE/RSJ Int.Conf. on Intelligent Robots and Systems IROS*, Grenoble, France, September 1997, pp.1057-1065

[10] R. Y. Tsai, "A Versatile Camera Calibration Technique for High-accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses", *IEEE J. Robotics Automat.*, vol. RA-3, no. 4, August 1987

[11] Z. Zhang, "A Flexible new Technique for Camera Calibration", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330-1334, November 2000