

# Camera Calibration and 3-D Measurement with an Active Stereo Vision System for Handling Moving Objects

Atushi Yamashita\*, Toru Kaneko\*, Shinya Matsushita\*\*, Kenjiro T. Miura\*  
and Suekichi Isogai\*\*\*

\* Department of Mechanical Engineering, Shizuoka University  
3-5-1 Hamamatsu-shi, Shizuoka 432-8561, Japan

E-mail address: yamashita@ieee.org

\*\* Nihon System Software

\*\*\* Denksha

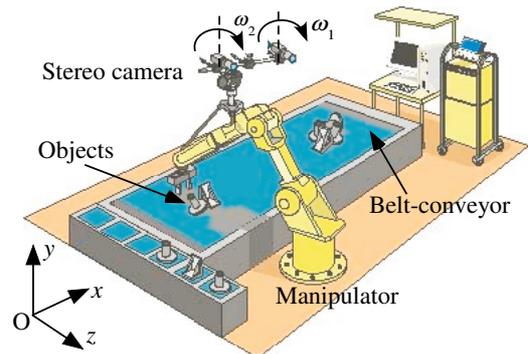
In this paper, we propose a fast and easy camera calibration and 3-D measurement method with an active stereo vision system for handling moving objects whose geometric models are known. We adopt the stereo vision system that can change its direction to follow the moving objects. To gain the extrinsic camera parameters in real time, a baseline stereo camera (parallel stereo camera) model and a projective transformation of stereo images are utilized by considering the epipolar constraints. To make use of 3-D measurement results of the moving object, the manipulator hand approaches the object. When the manipulator hand and the object are near enough for them to be in a single image, a very accurate camera calibration can be executed to calculate the manipulator size in the image. Our method does not need complicated image processing and can measure 3-D position and orientation of the object fast.

**key words:** stereo vision, camera calibration, 3-D measurement, epipolar constraints, moving object

## 1 Introduction

Robot vision is very important to handle moving objects with a robot manipulator. In general object-handling tasks, at first the system recognizes the type of an object, calculates its 3-D position and orientation, decides the grasping points on its surface, and then the robot hand approaches it while measuring 3-D position and orientation. There are a lot of researches about the object-handling task with a camera and manipulator system. As to the determination of grasping points on the object coordinate, an image based technique with images of a stereo vision system is proposed [1]. And many object-tracking and 3-D measurement methods with a stereo vision system are also proposed. For example, in [2], 3-camera stereo vision system that can track moving objects using geometric models was proposed. These studies assumed that the camera calibration was exactly done and if there existed some error about the calibration, the handling tasks could not be executed because the system recognized incorrect 3-D positions and orientations of the objects.

In this paper, we propose a 3-D measurement method with an active stereo vision system for handling moving objects (**Figure 1**). Each camera can change its direction about the Y-axis to follow the moving objects because the field of view of the camera is not wide enough to see the movable range of the objects. Our real-time camera calibration method is developed for the active camera system.



**Figure 1:** The manipulator handles the moving objects on the belt-conveyor by using the result of 3-D measurement with the active stereo vision system.

## 2 Active Stereo Vision System

### 2.1 Previous Works

To measure the exact position and orientation of the objects with a stereo vision system, the camera calibration is very important.

The calculation technique of a projective camera matrix mentioned in [3] is the most basic way of calibration. In [4], a fast calibration technique for a hand-eye system which can compute the relative position and orientation of a camera to a manipulator hand was proposed. However this technique cannot be applied to the moving object-tracking task. A self-calibration method for a moving camera based on a fundamental matrix [5, 6] and a method based on homography matrix [7] have been proposed. However these meth-

ods are generally weak for the reading errors of corresponding points and other image noise. Therefore a calibration method based on statistical model of image noise that decide the fundamental matrix from two uncalibrated perspective views was proposed [8]. This method can calculate the object’s position and orientation precisely, but it is not suitable for real-time object handling tasks because its computational cost is large. A fast algorithm aimed at the self-calibration of the stereo vision system based on the localization of specific markers placed on the object was proposed [9]. However this method cannot be used when there are no markers on the object. In [10, 11], a transformation that generates new pairs of images in which the epipolar lines are horizontal was utilized for a weakly-calibrated stereo vision system. This is a fast self-calibration method for rover navigation, but the accuracy is not good enough to grasp the moving objects.

On the other hands, as a visual servoing technique[12] is suitable for following a moving object or measuring by a moving camera, an object tracking method using data from a stereo vision system for a visual servoing task was proposed [13]. But main purpose of visual servoing manipulation is the dynamic control of a camera or a manipulator rather than the object recognition. When there are a lot of kinds of objects, the system must measure their 3-D positions and recognize their 3-D structures.

## 2.2 Challenging Points and Our Approach

In our study, all camera parameters must be calibrated in real time, because stereo camera’s directions always changes to track moving objects. In brief, the calibration methods that calculate the projective matrix, the fundamental matrix, or the homography matrix need a lot of computational costs to optimize these matrix. Hence these methods are not suitable for the real-time tasks. We can also calibrate all camera parameters and the position and the orientation of cameras’ axes of rotation in advance, and utilize the angle values from the motors’ encoders in real time. However, the parameters about each camera axis are six, and the extrinsic and the intrinsic camera parameters are eleven in this case. Each component of these matrices includes the extrinsic camera parameters and the image noise affects several parameters in a complicated way. Even if an optimal matrix is obtained, a good result of 3-D measurement can not necessarily always come out against the influence of image noise.

In this paper, we propose a fast and easy calibration method that utilizes the projective transformation and the epipolar constraints of the baseline stereo (parallel stereo) camera model. The epipolar constraints about the transformed pairs of images are considered. Moreover, the extremely exact 3-D measurement is not always required in object grasping tasks. When the cameras track the moving object, each camera never loses sight of the object by changing its direction to see the

object on the center of its image. The object tracking is robust against the calibration error, because it is based on the image-based visual servoing approach and uses no calibration results. When the manipulator hand grasps the object, high accuracy is needed. In that case, additional information can be obtained when the hand and the object is in the same image. Therefore, we propose a calibration technique that efficiently utilizes the manipulator’s size in the images and relative relationship between the manipulator hand and the object.

## 2.3 Overview of Handling Task

The outline of handling task is as follows. Each camera changes its direction about the  $Y$ -axis (see **Figure 1**; the  $Y$ -axis is perpendicular to the ground) to capture the image in which the moving object always locates on the center of each image plane. The characteristic points of the object’s silhouette on two images are extracted and then corresponding points are searched. After that, the projective transformation that generates new pairs of baseline stereo camera images is calculated by using the evaluation functions. From these new pairs of images, the 3-D position and orientation of the object is measured and the manipulator hand moves forward to the object. When the manipulator and the object are near enough to shot them in the same images, the more accurate calibration is performed to use additional information of the manipulator’s size.

## 2.4 Overview of Calibration

Each camera of our system can rotate about the  $Y$ -axis. The axis of rotation must be set perpendicular to the  $X$ - $Z$  plane (ground). But even when the great efforts are done about the camera system setting toward the world coordinate, the position and the direction of the rotation shafts necessarily deviates a little. Therefore, a deviation grows large according to the directional change of the active camera to follow the moving object. Unknown parameters of the active cameras increase because we must calculate the direction of each rotation shaft in addition to their initial relative position and initial direction between two cameras. This implies that an accurate calibration of each camera parameters becomes difficult and complicated.

In this paper, a two-step calibration technique without calculating all camera parameters precisely is proposed. In the first step, the rotation angles of the projective transformation about the  $X$ -,  $Y$ -, and  $Z$ -axes to make baseline stereo camera images are determined. The angles about the  $X$ - and  $Z$ -axes, and the change of baseline length are expressed as approximate functions of the change of the rotation angle about the  $Y$ -axis. This calibration is carried out before the grasping task. In the second step, the condition when the  $Y$  coordinate value of the corresponding points after transformation becomes equal is searched in real-time

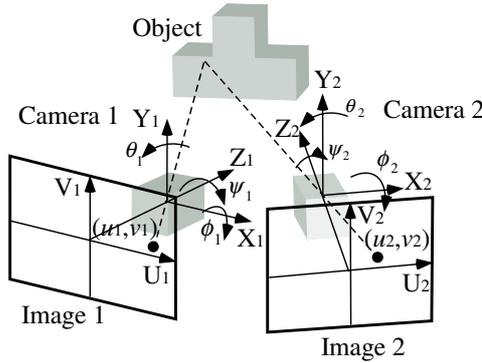
by using the evaluation functions. After that, the 3-D measurement is done. As for your being careful here, the real motor angle and the angle for the projective transformation are necessarily not the same. The reason why the original images are changed into those of the baseline stereo vision by using the projective transformation is to calibrate the camera parameters and to measure 3-D position faster compared with the method using the projective or the fundamental matrices. When the manipulator and the object are in the same image, more precise calibration is performed by using the manipulator's size additional to the epipolar constraints.

### 3 Calibration and 3-D Measurement

The angles for the projective transformation for the baseline stereo vision images are searched with the epipolar constraints that the epipolar lines are horizontal, and the 3-D measurement is carried out from the information of corresponding points.

#### 3.1 Principle of 3-D Measurement

The pairs of baseline stereo images are generated from the ordinary images with the projective transformation about the  $X$ -,  $Y$ -, and  $Z$ -axes (**Figure 2**).



**Figure 2:** Image planes of stereo vision system and coordinate systems of each camera.

The first projective transformation is about the  $Z$ -axis, secondly about the  $X$ -axis, and then about the  $Y$ -axis. The point  $(u_i, v_i)$  on each image ( $i = 1$ : left camera image,  $i = 2$ : right camera image) is projected to  $(x_i, y_i)$  with the projective transformation about the  $Z_i$ -axis (**Equations (1),(2)**).

$$x_i = u_i \cos \psi_i - v_i \sin \psi_i \quad (1)$$

$$y_i = u_i \sin \psi_i + v_i \cos \psi_i \quad (2)$$

where  $\psi_i$  is the angle of rotation about the  $Z_i$ -axis.

Then, the point  $(x_i, y_i)$  is projected to  $(x'_i, y'_i)$  with the projective transformation about the  $X_i$ -axis (**Equations (3),(4)**).

$$x'_i = \frac{\sqrt{1 + \tan^2 \phi_i}}{f - y_i \tan \phi_i} x_i \quad (3)$$

$$y'_i = \frac{f(\tan \phi_i + y_i)}{f - y_i \tan \phi_i} \quad (4)$$

where  $\phi_i$  is the angle of rotation about the  $X_i$ -axis, and  $f$  is the image distance.

Finally, the point  $(x'_i, y'_i)$  is projected to  $(x''_i, y''_i)$  with the projective transformation about the  $Y_i$ -axis (**Equations (5),(6)**).

$$x''_i = \frac{f(\tan \theta_i + x'_i)}{f - x'_i \tan \theta_i} \quad (5)$$

$$y''_i = \frac{\sqrt{1 + \tan^2 \theta_i}}{f - x'_i \tan \theta_i} y'_i \quad (6)$$

where  $\theta_i$  is the angle of rotation about the  $Y_i$ -axis.

If  $\psi_i$ ,  $\phi_i$  and  $\theta_i$  are known, the 3-D position of the corresponding point  $(u_i, v_i)$  on each image plane is calculated as follows with the baseline stereo camera model (**Equations (7) – (9)**).

$$X = \frac{b(x''_1 + x''_2)}{2d} \quad (7)$$

$$Y = \frac{b(y''_1 + y''_2)}{2d} \quad (8)$$

$$Z = \frac{bf}{d} \quad (9)$$

where  $b$  is the length of baseline, and  $d = x''_1 - x''_2$  is the disparity of the corresponding points.

Ideally, each motor angle  $\omega_i$  is equal to the angle  $\theta_i$  for the projective transformation, and  $\psi_i$  and  $\phi_i$  are always equal to 0. However there is a setting error of each motor's shaft and hence  $\omega_i$  is not equal to  $\theta_i$ . Hence, we use  $\psi_i$ ,  $\phi_i$  and  $\theta_i$  for the projective transformation.

#### 3.2 Evaluation of Projective Transformation

The epipolar lines of the baseline stereo camera images are horizontal. Therefore, two images are close to the baseline stereo vision images when the evaluation function  $E_{epi}$  has a small value (**Equation (10)**).

$$E_{epi} = \frac{1}{N} \sum_{i=1}^N (y''_{1,i} - y''_{2,i})^2 \quad (10)$$

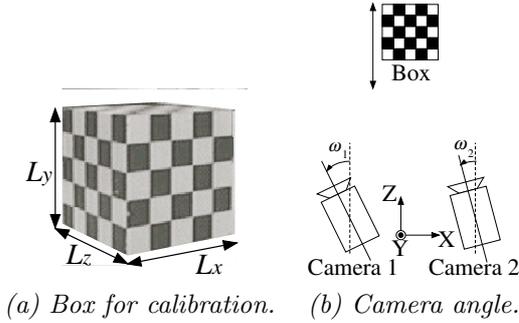
where  $y''_{1,i}$  and  $y''_{2,i}$  are the  $Y$  coordinate value of the corresponding point calculated with **Equation (6)**, and  $N$  is a number of the corresponding points.

Ideally, a good 3-D measurement can be done to use the angles  $\psi_i$ ,  $\phi_i$ ,  $\theta_i$  that minimize  $E_{epi}$ . However, a

good calibration never be done with the influence of image noise. Therefore, we propose a two-step calibration technique that considers not only the epipolar constraints but also the length information.

### 3.3 Calibration in Advance

In the first calibration step, each camera's feature is measured by using a box whose surface is painted with the planner pattern (**Figure 3**).



**Figure 3:** Calibration in advance.

The calibration is executed by changing each camera angle  $\omega_i$  and the position of the box along  $Z$ -axis. The surface of the box is set perpendicular to the optical axes of two cameras in the initial angle ( $\omega_i = 0$ ). At first, the image distance  $f$  and the baseline length  $b_0$  in the initial angle is calculated by changing the distance between the box and the camera system. After that, the left camera angle  $\omega_1$  is changed discretely and the pairs of images are gained while  $\omega_1$  is always set 0. The discrete value of the angle change is constant. The right camera angle  $\omega_2$  is changed in the same way while  $\omega_1 = 0$ . The same procedure is repeated by changing the distance between the box and the camera system. To minimize the disparity of length error between 3 axes,  $E_{dis}$  is introduced (**Equation (11)**).

$$E_{dis} = \frac{1}{3} \{ (l_x - l_y)^2 + (l_y - l_z)^2 + (l_z - l_x)^2 \} \quad (11)$$

where  $l_x$ ,  $l_y$ , and  $l_z$  are the length of the box's edge along  $X$ ,  $Y$ ,  $Z$ -axis from the result of 3-D reconstruction of the box shape.

An accurate calibration can be done to use the angles  $\psi_i$ ,  $\phi_i$ ,  $\theta_i$  that minimize  $E_1$  (**Equation (12)**).

$$E_1 = \alpha_1 E_{epi} + \beta_1 E_{dis} \quad (12)$$

where  $\alpha_1$  and  $\beta_1$  are weight coefficients.

When  $\omega_2$  is settled 0 and  $\omega_1$  is changed, the optimal  $\theta_1$ ,  $\psi_1$ , and  $\phi_1$  are calculated discretely according to the  $\omega_1$  value to minimize  $E_1$ . The change of baseline length  $\Delta b_1$  is calculated in **Equation (13)**. When  $\omega_1$  is settled 0, the optimal  $\theta_2$ ,  $\psi_2$ ,  $\phi_2$ , and  $\Delta b_2$  is calculated

discretely, too. The baseline length is calculated in **Equation (14)**.

$$\Delta b_i = \left( \frac{L_x + L_y + L_z}{l_x + l_y + l_z} - 1 \right) b_0 \quad (13)$$

$$b = b_0 + \Delta b_1 + \Delta b_2 \quad (14)$$

where  $L_x$ ,  $L_y$ ,  $L_z$  are the real length of the box, and  $l_x$ ,  $l_y$ ,  $l_z$  are measured length.

The discrete values of  $\theta_i$ ,  $\psi_i$ ,  $\phi_i$ , and  $\Delta b_i$  are fitted on approximate functions. To find optimal angle for the projective transformation fast in the second-step real-time calibration,  $\psi_i$ ,  $\phi_i$ , and  $\Delta b_i$  are expressed as the functions of  $\theta_i$  (**Equation (15)–(17)**).

$$\psi_i = f_i(\theta_i) \quad (15)$$

$$\phi_i = g_i(\theta_i) \quad (16)$$

$$\Delta b_i = h_i(\theta_i) \quad (17)$$

### 3.4 Calibration in Real Time

After calculating the approximate function, the real-time calibration for the handling task is done. To do a more accurate calibration, the length information is utilized. The marks are attached to the manipulator hand and the distance between two marks are known (**Figure 5** (a),(b)). The evaluation function of the difference between the real length and the measured length  $E_{len}$  is set as **Equation (18)**. To search  $\theta_i$  that minimizes  $E_2$ , the real-time 3-D measurement is executed. The computation costs of the optimal conditions become small because  $\psi_i$ ,  $\phi_i$ , and  $\Delta b_i$  are expressed as function of  $\theta_i$ , and all that the system must do is to decide  $\theta_i$  that minimizes  $E_2$  (**Equation (19)**).

$$E_{len} = \frac{1}{M} \sum_{i=1}^M (L_i - l_i)^2 \quad (18)$$

$$E_2 = \alpha_2 E_{epi} + \beta_2 E_{len} \quad (19)$$

where  $L_i$  is the real distance between two marks,  $l_i$  is the calculated distance by image processing,  $M$  is a number of mark distances, and  $\alpha_2$  and  $\beta_2$  are weight coefficients.

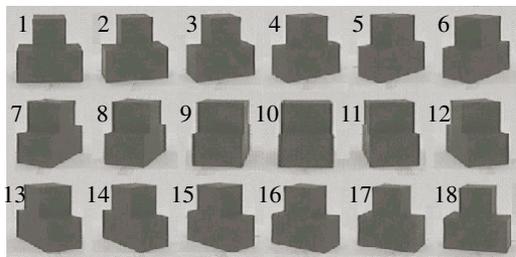
In the early time when each camera tracks the moving object, the manipulator is not in the images. In that case, the information about the mark distance cannot be used and  $\beta_2$  is set 0. The manipulator is guided close to the moving object by the result of 3-D measurement only considering  $E_{epi}$ . When the manipulator and the object are in the same image,  $\beta_2$  becomes a positive number. The accuracy of 3-D measurement becomes good and the relative distance between the manipulator and the object in the camera coordinate (not in the world coordinate) is utilized to control the position and the orientation of manipulator hand at the same time.

### 3.5 Object Recognition

We use a template matching technique for detecting the types of the objects and the corresponding points between two images. The sequential similarity detection algorithm (SSDA) is adopted as the template matching because it can work fast. There are two processes in our template matching using the object template and the point template respectively.

At first, the type of the object and the rough orientation recognition is done by using the object templates shown in **Figure 4**. In each object template, the type and the orientation of the object change discretely. The object models can be obtained either from the geometric models and from the real images. The most fitting template is selected with SSDA by searching the center position of each camera image. For example, object template 17 is selected in the left image (**Figure 5** (a)) and object template 15 is selected in the right image (**Figure 5** (b)).

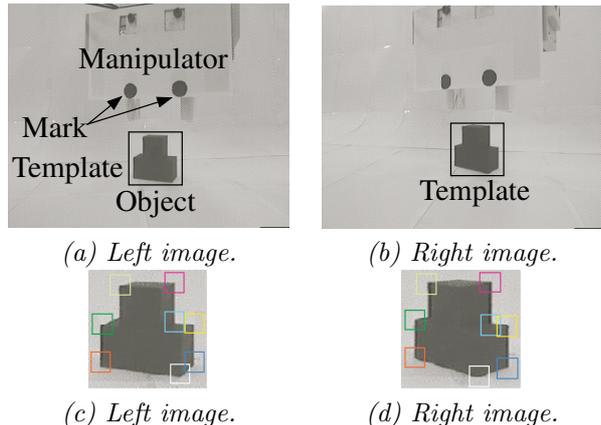
Secondly, exact corresponding points are searched with the point templates of each object template. Each point template has small template ranges in which characteristic area of object's silhouette exists (**Figure 5** (c),(d)). The point template can rotate around  $Z$ -axis, and a fine adjustment along the  $X$ - and  $Y$ -axes is done with SSDA to search corresponding points. After detecting corresponding points on the object and the manipulator, the angles for the projective transformation is calculated by **Equation** (19) and 3-D positions and orientations are measured.



**Figure 4:** Example of the object template.

## 4 Experiments

To verify the validity of our proposed method, experiments with a stereo camera system, a manipulator and a belt-conveyer were done. Each camera's image distance is fixed equally. The manipulator has six rotational joints and a parallel gripper hand with touch sensors. The camera system does not know the flow speed of the belt-conveyer. To increase the contrast between the moving objects and the other materials, the belt-conveyer and the manipulator are painted in white. The object templates were generated every 10deg on one object and each object template had eight



**Figure 5:** Template matching. (a),(b): Object kind and rough orientation recognition with the object template. (c),(d): Exact recognition of corresponding points with the point template.

position templates in maximum. One computer (CPU: Dual Pentium II 450MHz, Memory: 512MB) controls the manipulator and the camera system, and executes image processing.

### 4.1 Camera Calibration in Advance

The first-step camera calibration was done before the handling task. At first, the image distance  $f$  and the initial length of baseline  $b_0$  was measured by using the box shown in **Figure 3**. As the result of calibration,  $f = 4845.4$  pixel,  $b_0 = 152.6$  mm. The relationship between  $\theta_i$  and the other parameters were examined under the condition when  $\alpha_1 = 1$  and  $\beta_1 = 1$ . We changed the camera angle  $\omega_i$  by discrete value 5deg from  $-30\text{deg}$  to  $30\text{deg}$ . The results of camera calibration could be expressed as the polynomial functions whose degree are three in our stereo camera system (**Equations** (20) – (25)).

$$\begin{aligned} \psi_1 = & -3.0 \times 10^{-5}\theta_1^3 + 9.8 \times 10^{-4}\theta_1^2 \\ & + 5.7 \times 10^{-2}\theta_1 - 3.0 \times 10^{-1} \end{aligned} \quad (20)$$

$$\begin{aligned} \phi_1 = & 3.8 \times 10^{-7}\theta_1^3 + 3.9 \times 10^{-4}\theta_1^2 \\ & + 2.0 \times 10^{-3}\theta_1 - 8.1 \times 10^{-1} \end{aligned} \quad (21)$$

$$\begin{aligned} \Delta b_1 = & -1.8 \times 10^{-5}\theta_1^3 + 2.7 \times 10^{-4}\theta_1^2 \\ & - 4.2 \times 10^{-2}\theta_1 - 1.2 \times 10^{-2} \end{aligned} \quad (22)$$

$$\begin{aligned} \psi_2 = & -5.7 \times 10^{-6}\theta_2^3 - 3.6 \times 10^{-4}\theta_2^2 \\ & + 5.6 \times 10^{-2}\theta_2 + 2.6 \times 10^{-1} \end{aligned} \quad (23)$$

$$\begin{aligned} \phi_2 = & -9.8 \times 10^{-6}\theta_2^3 + 1.3 \times 10^{-4}\theta_2^2 \\ & - 4.2 \times 10^{-2}\theta_2 + 8.0 \times 10^{-1} \end{aligned} \quad (24)$$

$$\begin{aligned} \Delta b_2 = & -1.7 \times 10^{-5}\theta_2^3 - 4.3 \times 10^{-4}\theta_2^2 \\ & + 4.9 \times 10^{-2}\theta_2 - 4.9 \times 10^{-2} \end{aligned} \quad (25)$$

## 4.2 3-D Shape and Length Reconstruction and Object Handling

The results of 3-D reconstruction are shown in **Table 1**. Eight points on the box for calibration was measured by our active stereo camera system while the camera angles were changed. The results are the average values of the experiments of twelve times. In this table,  $X_e$ ,  $Y_e$ ,  $Z_e$  indicate the average length error along each axis and  $\eta_e$  indicates the average angle error between two axes when  $\omega_i$  is from  $-30\text{deg}$  to  $30\text{deg}$ . The real length is 60mm and the real angle is  $90\text{deg}$ . The  $Z$ -axis direction is the same as the optic axis direction in the initial motor setting ( $\omega_i = 0$ ).

In Strategy 1, the 3-D measurement is done by only using motor angle  $\omega_i$  for the projective transformation. In this case,  $\omega_i = \theta_i$ , and  $\psi_i$ ,  $\phi_i$ , and  $\Delta b_i$  are always equal to 0. This means that the calibration is not done in this strategy. In Strategy 2, the 3-D measurement is done by using **Equation (19)** under condition when  $\alpha_2 = 1$  and  $\beta_2 = 0$ . It is the case where the manipulator and the object are not in the same image. In Strategy 3,  $\alpha_2 = 1$ ,  $\beta_2 = 1$ . It is the case where they are in the same image and the length information is utilized for the camera calibration.

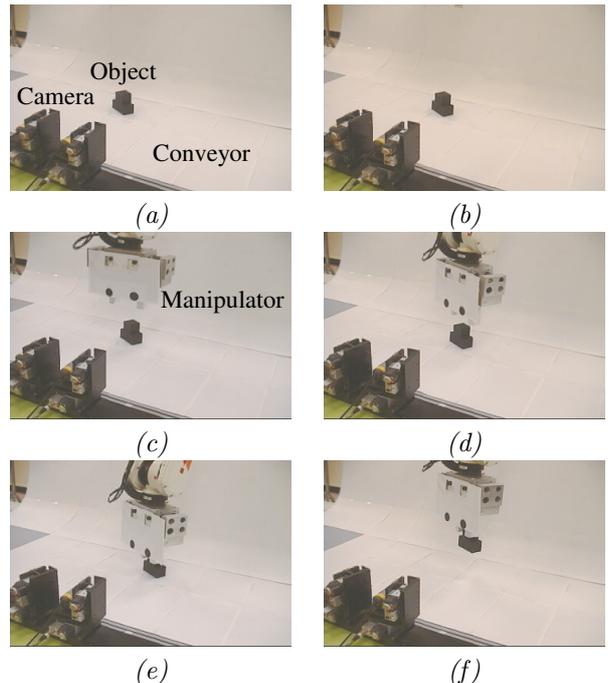
In our method, the correction of image distortions was not necessarily needed because cameras changed their directions to see the object on the center of the images, and the distortions of the image centers are vanishingly small. Even when the camera does not see the object on the center of the images, there is no problem about the non-linear image distortion to eliminate the influence of the distortion (for example, [14]) at first.

**Table 1:** The results of 3-D reconstruction.

Strategy	$X_e$	$Y_e$	$Z_e$	$\eta_e$
1	8.9mm (14.8%)	6.4mm (10.7%)	9.7mm (16.1%)	2.9deg (3.2%)
2	3.8mm (6.3%)	1.4mm (2.3%)	3.4mm (5.7%)	2.4deg (2.7%)
3	0.1mm (0.2%)	0.8mm (1.3%)	1.8mm (3.1%)	0.7deg (0.1%)

The accuracy of the 3-D reconstruction in Strategy 3 was the best and the maximum error was 1.8mm along the  $Z$ -axis (depth direction) when the distance between the stereo camera system and the object was within 500mm. The angle error was within 1.0deg. The system often failed the handling task when Strategy 1 was adopted. The main reason why the manipulator failed to grasp the object was that the accuracy of 3-D measurement was not high and the manipulator could not approach the object appropriately. To adopt Strategy 2 when the manipulator and the object are not in the same image and adopt Strategy 3 when they are in the same image, the manipulator could smoothly approach the object and grasp it.

The example of handling task is shown in **Figure 6**. From (a) to (b), the object and the manipulator were not in the same image, and the manipulator approached the object by the result of real-time calibration without the length information and 3-D measurement. In (c), they were in the same image, and the real-time calibration was done under the condition when  $\alpha_2 = 1$  and  $\beta_2 = 1$ . In (d), the precise positioning was done. In (e), the manipulator grasped the object by using the information of touch sensor, and in (f), it carried the object successfully.

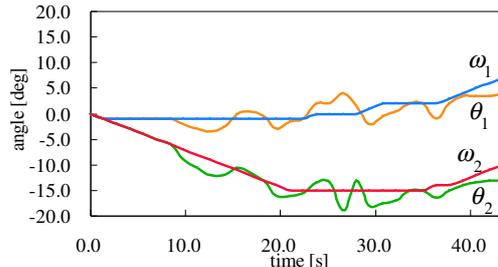


**Figure 6:** Experimental results of the object handling task.

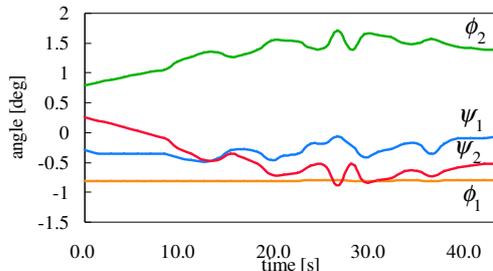
The changes of the motor angle, the angles for the projective transformation, and the baseline length are shown in **Figure 7**. From these results, it is seen that  $\omega_i$  was not equal to  $\theta_i$  and accurate 3-D measurement was executed by considering the epipolar constraints and the length information. The system sometimes mistook recognizing the corresponding points in template matching at a certain moment, but it could recover the high accuracy at the next moment. From these experimental results, our proposed method is for practical use when the manipulator handles the moving objects in real time.

## 5 Conclusions

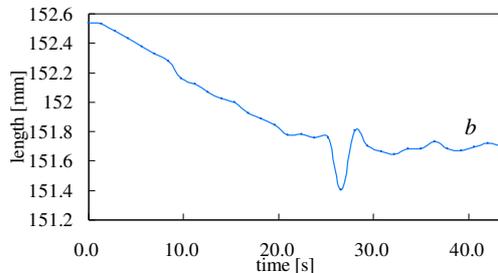
In this paper, we have proposed a new calibration and 3-D measurement method for handling moving objects. By using the projective transformation and expressing the relationship between angles for the transformation as the approximate functions, the fast calibration can be done. Experimental results show that



(a) Angle  $\omega_i, \theta_i$ .



(b) Angle  $\psi_i, \phi_i$ .



(c) Baseline length  $b$ .

**Figure 7:** The changes of the angles for the projective transformation and the baseline length.

3-D measurement has high accuracy.

As the future works, the template matching must become faster and more precisely by using the hardware that is good at the correlation calculation (for example, normalized cross-correlation) to improve the robustness against the changes of the lighting conditions.

## References

- [1] Alexa Hauck, Johanna Ruttinger, Michael Sorg and Georg Farber: “Visual Determination of 3D Grasping Points on Unknown Objects with a Binocular Camera System”, *Proceedings of the 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.272–278, 1999.
- [2] T. Matsushita, Y. Sumi, Y. Ishiyama and F. Tomita: “A Tracking Based Manipulation System Built on Stereo Vision”, *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.185–190, 1998.
- [3] Roger 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*, Vol.RA-3, No.4, pp.323–344, 1987.
- [4] Roger Y. Tsai and Reimark K. Lenz: “A New Technique for Fully Autonomous and Efficient 3D Robotics Hand/Eye Calibration”, *IEEE Transactions on Robotics and Automation*, Vol.5, No.3, pp.345–358, 1989.
- [5] Stephen J. Maybank and Olivier D. Faugeras: “A Theory of Self-Calibration of a Moving Camera”, *International Journal of Computer Vision*, Vol.8, No.2, pp.123–151, 1992.
- [6] Q.-T. Luong and O. D. Faugeras: “Self-Calibration of a Moving Camera from Point Correspondences and Fundamental Matrices”, *International Journal of Computer Vision*, Vol.22, No.3, pp.261–289, 1997.
- [7] Ricard Hartley, Rajiv Gupta and Tom Chang: “Stereo from Uncalibrated Cameras”, *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp.761–764, 1992.
- [8] Kenichi Kanatani: “Gauge-Based Reliability Analysis of 3-D Reconstruction from Two Uncalibrated Perspective Views”, *Proceedings of 15th International Conference on Pattern Recognition*, pp.76–79, 2000.
- [9] Alberto Broggi, Massimo Bertozzi and Alessandra Fascioli: “Self-Calibration of a Stereo Vision System for Automotive Applications”, *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, pp.3698–3703, 2001.
- [10] Luc Robert, Michel Buffa and Martial Hebert: “Weakly-Calibrated Stereo Perception for Rover Navigation”, *Proceedings of Image Understanding Workshop*, pp.1317–1324, 1994.
- [11] Eric Krotkov and Martial Hebert: “Mapping and Positioning for a Prototype Lunar Rover”, *Proceedings of the 1995 IEEE International Conference on Robotics and Automation*, pp.2913–2919, 1995.
- [12] Seth Hutchinson, Gregory D. Hanger and Peter I. Corke: “A Tutorial on Visual Servo Control”, *IEEE Transactions on Robotics and Automation*, Vol.12, No.5, pp.651–670, 1996.
- [13] Romuald Ginhoux and Jens-Steffen Gutmann: “Model-Based Object Tracking Using Stereo Vision”, *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, pp.1226–1232, 2001.
- [14] Juyang Weng, Paul Cohen and Marc Herniou: “Camera Calibration with Distortion Models and Accuracy Evaluation”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.14, No.10, pp.965–980, 1992.