We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Development of Pilot Assistance System with Stereo Vision for Robot Manipulation

Takeshi Nishida[1], Shuichi Kurogi[1], Koichi Yamanaka[1],
Wataru Kogushi[1] and Yuichi Arimura[2]
[1]Kyushu Institute of Technology
[2]tmsuk co., LTD
Japan

1. Introduction

Recently, the rescue robots called "T-52 Enryu" and "T-53 Enryu"(Fig. 1) that aimed at the rescue operation in danger zone such as disaster district etc. (Fig. 2) have been developed. The robots have two airms with eight DOFs for execution of various tasks, and it can be operated remotely in order to achieve the tasks such as removal, mining, and so on without sending the pilot to a dangerous place (Yokokohji, 2006). However, these operations were difficult at the beginning of development of the robot because there was not any pilot assistance system, and time for a great deal of training was necessary (Nishida et al. 2006). Namely, the pilot had to understand a situation of surrounding of the robot from a few monitors that display the images of the cameras installed in the robot. Thus, because an assistance system to help the understanding of the positions of target objects had been necessary to improve the efficiency of the remote operation of the robot, we have developed a vision system that assists the pilot by tracking multiple targets and intelligibly presenting them.

In the development of this pilot assistance system, we had to note following matters. Namely, we considered the simplicity of employment and maintenance, and the cost of devices. Moreover, since scenery consists of the achromatic color objects such as sand and rubble by many cases at place of the rescue operation, the measurement and the calibration of the system should be appropriately executed without the color information. Furthermore, since the vision system equipped on the gripper, the system should be robust to the vibration caused by the movement of the arm. Thus, we had developed a fast executable stereovision system robust against vibration of the arm without using the color information and easy to understand for the pilot. The system consists of a parallel stereo camera equipped on the gripper joint of the right arm and a laptop PC, and they were installed into the Enryus.

The remainder of this chapter is divided into the following topics: Section 2 describes the structure of the pilot assistance system. Section 3 describes a calibration method of the system robust for environmental change. Section 4 details interfaces of the system for displaying the results of target tracking and distance measurement, and finally Section 5 discusses the overall implementation and successful application of the robot.
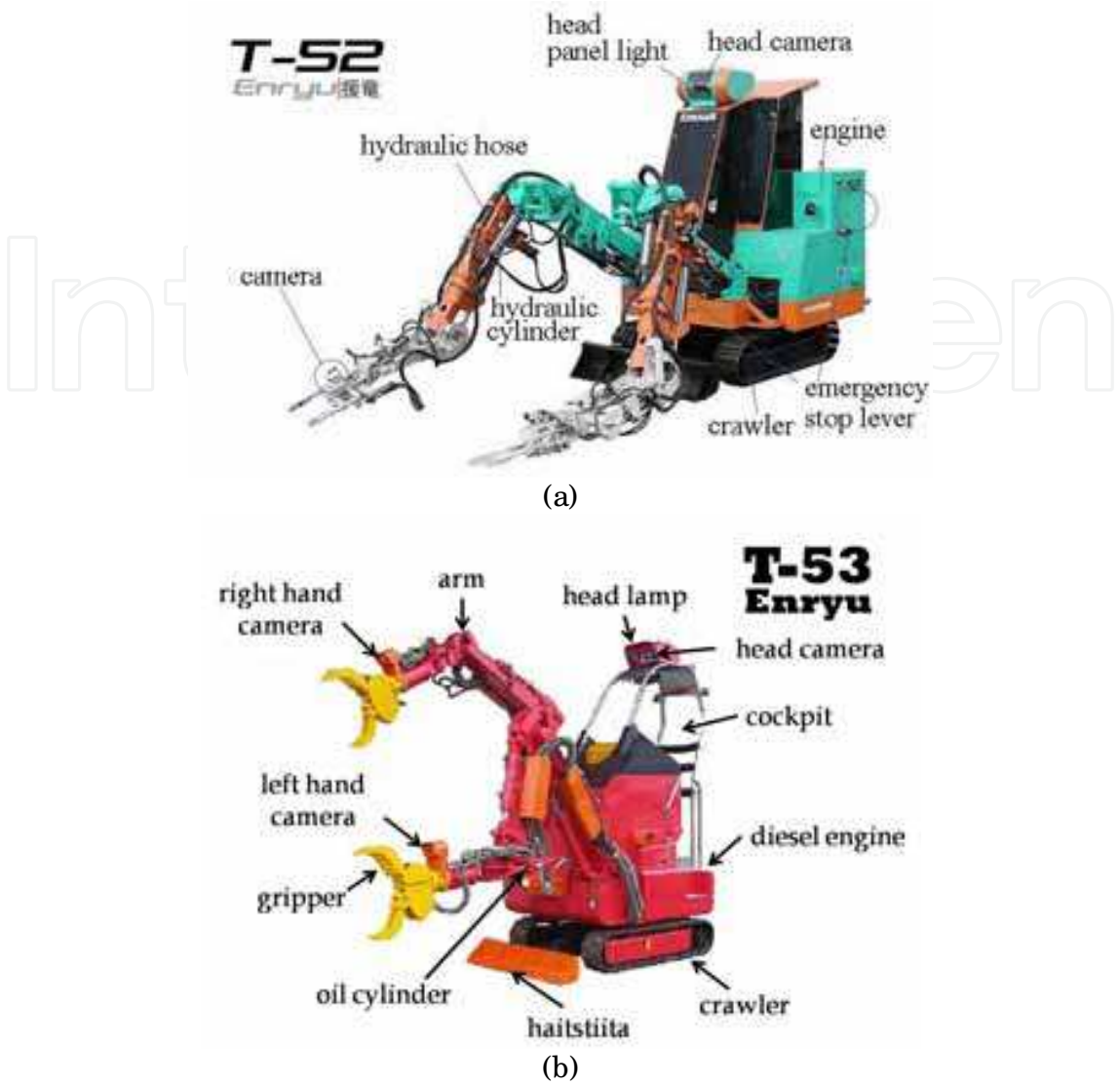
(a)

(b)

Fig. 1. Rescue robots: (a) T-52 Enryu, height: 3.45 [m], width: 2.4 [m], length: 3.5 [m] (maximum length: 10 [m]), weight: 5 [t]; and (b) T-53 Enryu, height: 2.3 [m], width: 1.4 [m] (maximum length: 7.23[m]), length: 2.32 [m], weight: 2.9 [t].

(a) T-52   (b) T-53

Fig. 2. Appearances of the activity of the robots in disaster district.

2. Parallel stereo camera system

Two kinds of camera systems and those applications were developed. One is a suite of fixed parallel cameras and a 3D motion sensor (Fig. 3(a)), and another is a suite of parallel cameras and a camera mount for adjusting the installation angles and positions of them (Fig. 3 (b)). The former called "A-system" was composed to verify what visual application is effective to support of the pilot; the latter called "B-system" was developed to examine installation position and effective calibration method on the actual robot arm.
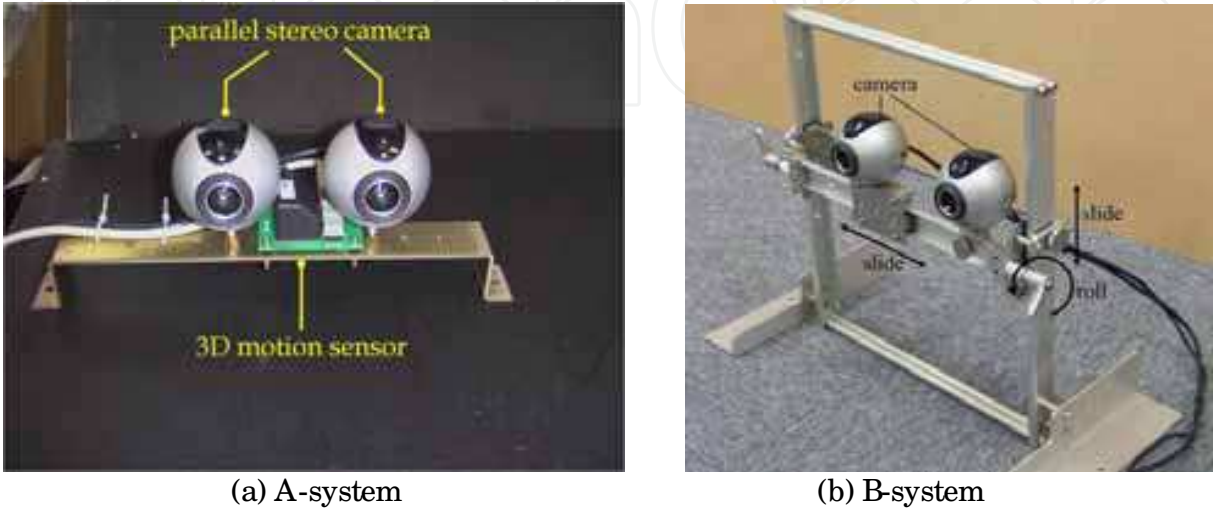

(a) A-system   (b) B-system

Fig. 3. Developed parallel stereo camera systems: (a) A-system consists of fixed parallel USB cameras and a motion sensor; (b) B-system consists of parallel USB cameras and a camera mount for adjusting the installation angles and positions.

We selected the inexpensive USB cameras (QV-400R, USB CCD camera) and motion sensor in order to simplify the replacement of them even if breaking down by severe work in outdoor environment. This suite of camera system is mounted on a joint of gripper as shown in Fig. 4 (a) and (b), and they are connected to a laptop PC installed in the robot. The obtained images are transmitted to a remote control cockpit via wireless LAN for supplying the information around the right gripper and the image processing results (Fig. 5).

(a)   (b)

Fig. 4. Installation of the system: (a) the system mounted on the right gripper of T-52; (b) the system mounted on the right gripper of T-53.
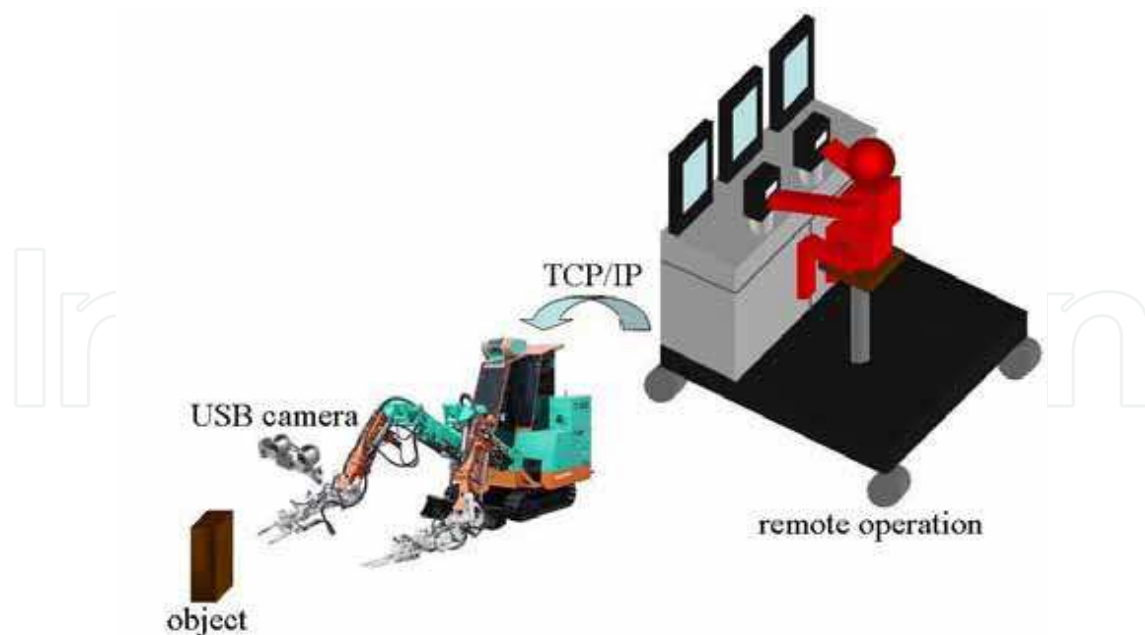
Fig. 5. Whole structure of the system. The pilot operates the robot remotely by watching plural monitors and using joysticks etc.

3. Calibration of the camera systems robust for environmental change

3.1 Camera model

The USB cameras are modeled by using the basic pinhole camera model (e.g. Tsai et al. 1987; Hartley et al. 2003) and the geometry is shown in Fig. 6. Let $\Sigma_w = \{X, Y, Z\}$ be the camera coordinate system, and $\Sigma_c = \{x, y\}$ be the image plane coordinate system. A mapping a point $\boldsymbol{X} = (X, Y, Z)^T \in \Sigma_w$ to a point $\boldsymbol{x} = (x, y)^T \in \Sigma_c$ is represented as

$$\tilde{\boldsymbol{x}} = \boldsymbol{K} \begin{bmatrix} \boldsymbol{R} | \boldsymbol{t} \end{bmatrix} \tilde{\boldsymbol{X}} \triangleq \boldsymbol{P} \tilde{\boldsymbol{X}}, \tag{1}$$

where,

$$\tilde{\boldsymbol{x}} \triangleq \begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix}, \quad \tilde{\boldsymbol{X}} \triangleq \begin{bmatrix} \boldsymbol{X} \\ 1 \end{bmatrix}, \quad \boldsymbol{K} = \begin{bmatrix} \alpha & -\alpha \cot \theta & u_0 \\ 0 & \beta/\sin \theta & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{R} = \begin{bmatrix} \boldsymbol{r}_1^T \\ \boldsymbol{r}_2^T \\ \boldsymbol{r}_3^T \end{bmatrix}, \boldsymbol{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}, \tag{2}$$

and $[u_0, v_0]^T$ is the camera center. The number of pixels per unit distance in image coordinates are m_x and m_y in the x and y directions, f is focal length, and $\alpha = fm_x$ and $\beta = fm_y$ represent the focal length of the camera in terms of pixel dimensions in x and y directions. Moreover the parameter θ is the angle between the vertical and horizontal axes in the image plane, namely it is referred to as the skew parameter.
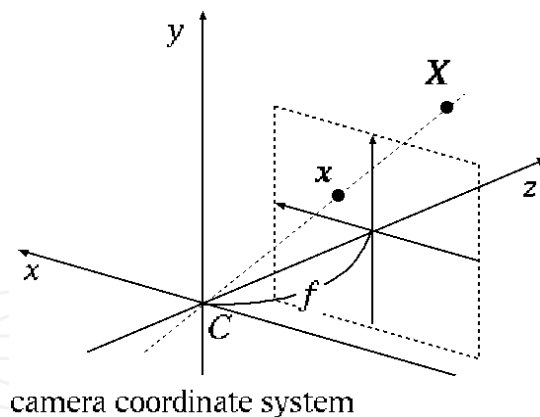
Fig. 6. Pinhole camera geometry.

3.2 Calculation of camera parameters

Next, in this research, Eqn. (1) is expanded to two linear equations as follows,

$$\begin{cases} \boldsymbol{p}_1^T \tilde{\boldsymbol{X}}_i - x_i \boldsymbol{p}_3^T \tilde{\boldsymbol{X}}_i + p_{14} - x_i p_{34} = 0, \\ \boldsymbol{p}_2^T \tilde{\boldsymbol{X}}_i - y_i \boldsymbol{p}_3^T \tilde{\boldsymbol{X}}_i + p_{24} - y_i p_{34} = 0, \end{cases} \tag{3}$$

where

$$\boldsymbol{P} \triangleq \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \triangleq \begin{bmatrix} \boldsymbol{p}_1^T & p_{14} \\ \boldsymbol{p}_2^T & p_{24} \\ \boldsymbol{p}_3^T & p_{34} \end{bmatrix}. \tag{4}$$

Combine the points $\tilde{\boldsymbol{X}}_i = [X_i, Y_i, Z_i, 1]^T$ $(i = 1, 2, \cdots, n)$ and the points $\boldsymbol{x}_i = [x_i, y_i]^T \in \Sigma_c$ correspond to them, it becomes

$$\boldsymbol{E}\boldsymbol{p} = \boldsymbol{0} \tag{5}$$

$$\boldsymbol{E} \triangleq \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1 X_1 & -x_1 Y_1 & -x_1 Z_1 & -x_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1 X_1 & -y_1 Y_1 & -y_1 Z_1 & -y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & -x_n X_n & -x_n Y_n & -x_n Z_n & -x_n \\ 0 & 0 & 0 & 0 & X_n & Y_n & Z_n & 1 & -y_n X_n & -y_n Y_n & -y_n Z_n & -y_n \end{bmatrix}, \tag{6}$$

$$\boldsymbol{p} \triangleq \left[\boldsymbol{p}_1^T, p_{14}, \boldsymbol{p}_2^T, p_{24}, \boldsymbol{p}_3^T, p_{34} \right]^T. \tag{7}$$

The projection matrix $\boldsymbol{E}$ is including 11 independent variables. In general, the solution of $\boldsymbol{p}$ in Eqn. (7) is obtained as an eigenvector corresponds to the minimum eigen value which derived by singular value decomposition of the covariance matrix $\boldsymbol{E}^T \boldsymbol{E}$. Here, the elements of $\boldsymbol{p}$ are normalized to be $\|p_{34}\| = 1$. Moreover,

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{p}_1^T & p_{14} \\ \boldsymbol{p}_2^T & p_{24} \\ \boldsymbol{p}_3^T & p_{34} \end{bmatrix} = \begin{bmatrix} \alpha \boldsymbol{r}_1^T - \alpha(\cot\theta)\boldsymbol{r}_2^T + u_0 \boldsymbol{r}_3^T & \alpha t_x - \alpha(\cot\theta)t_y + u_0 t_z \\ (\beta/\sin\theta)\boldsymbol{r}_2^T + v_0 \boldsymbol{r}_3^T & (\beta/\sin\theta)t_y + v_0 t_z \\ \boldsymbol{r}_3^T & t_z \end{bmatrix}, \tag{8}$$

where,

$$
\begin{aligned}
u_0 &= \boldsymbol{p}_1^T \boldsymbol{p}_3, \\
v_0 &= \boldsymbol{p}_2^T \boldsymbol{p}_3, \\
\theta &= \cos^{-1}\left(-\frac{(\boldsymbol{p}_1 \times \boldsymbol{p}_3)^T (\boldsymbol{p}_2 \times \boldsymbol{p}_3)}{\|\boldsymbol{p}_1 \times \boldsymbol{p}_3\| \|\boldsymbol{p}_2 \times \boldsymbol{p}_3\|} \right), \\
\alpha &= \sqrt{\boldsymbol{p}_1^T \boldsymbol{p}_1 - u_0^2} \sin\theta, \\
\beta &= \sqrt{\boldsymbol{p}_2^T \boldsymbol{p}_2 - v_0^2} \sin\theta, \\
\boldsymbol{r}_1 &= \frac{1}{\alpha}\left(\boldsymbol{p}_1 + (\boldsymbol{p}_2 - v_0 \boldsymbol{p}_3)\frac{\alpha}{\beta}\cos\theta - u_0 \boldsymbol{p}_3 \right), \\
\boldsymbol{r}_2 &= \frac{\sin\theta}{\beta}(\boldsymbol{p}_2 - v_0 \boldsymbol{p}_3), \\
\boldsymbol{r}_3 &= \boldsymbol{p}_3,
\end{aligned}
\tag{9}
$$

$$
\begin{aligned}
t_x &= \frac{1}{\alpha}\left(p_{14} + (p_{24} - v_0 p_{34})\frac{\alpha}{\beta}\cos\theta - u_0 p_{34} \right), \\
t_y &= \frac{\sin\theta}{\beta}(p_{24} - v_0 p_{34}), \\
t_z &= p_{34}.
\end{aligned}
\tag{10}
$$

3.3 measurement of 3D position

Next, let the projected points of $\boldsymbol{X} \in \Sigma_w$ on the left and the right camera image planes be $\boldsymbol{x}^{\text{left}} \triangleq \begin{bmatrix} x^{\text{left}}, y^{\text{left}} \end{bmatrix}^T$ and $\boldsymbol{x}^{\text{right}} \triangleq \begin{bmatrix} x^{\text{right}}, y^{\text{right}} \end{bmatrix}^T$. Moreover, let $\boldsymbol{P}^{\text{left}}$ and $\boldsymbol{P}^{\text{right}}$ be projection matrices, the following equations are derived,

$$\tilde{\boldsymbol{x}}^{\text{left}} = \boldsymbol{P}^{\text{left}} \tilde{\boldsymbol{X}}, \qquad \tilde{\boldsymbol{x}}^{\text{right}} = \boldsymbol{P}^{\text{right}} \tilde{\boldsymbol{X}}. \tag{11}$$

When the above mentioned expression are brought together about the variable $\boldsymbol{P}$, the following expression is obtained.

$$\boldsymbol{B}\tilde{\boldsymbol{X}} = \boldsymbol{b}, \tag{12}$$

where,

$$\boldsymbol{B} \triangleq \begin{bmatrix} x^{\text{left}}p_{31}^{\text{left}} - p_{11}^{\text{left}} & x^{\text{left}}p_{32}^{\text{left}} - p_{12}^{\text{left}} & x^{\text{left}}p_{33}^{\text{left}} - p_{13}^{\text{left}} \\ y^{\text{left}}p_{31}^{\text{left}} - p_{21}^{\text{left}} & y^{\text{left}}p_{32}^{\text{left}} - p_{22}^{\text{left}} & y^{\text{left}}p_{33}^{\text{left}} - p_{23}^{\text{left}} \\ x^{\text{right}}p_{31}^{\text{right}} - p_{11}^{\text{right}} & x^{\text{right}}p_{32}^{\text{right}} - p_{12}^{\text{right}} & x^{\text{right}}p_{33}^{\text{right}} - p_{13}^{\text{right}} \\ y^{\text{right}}p_{31}^{\text{right}} - p_{21}^{\text{right}} & y^{\text{right}}p_{32}^{\text{right}} - p_{22}^{\text{right}} & y^{\text{right}}p_{33}^{\text{right}} - p_{23}^{\text{right}} \end{bmatrix}, \quad (13)$$

$$\boldsymbol{b} \triangleq \begin{bmatrix} p_{14}^{\text{left}} - x^{\text{left}}p_{34}^{\text{left}}, p_{24}^{\text{left}} - x^{\text{left}}p_{34}^{\text{left}}, p_{14}^{\text{right}} - y^{\text{right}}p_{34}^{\text{right}}, p_{24}^{\text{right}} - y^{\text{right}}p_{34}^{\text{right}} \end{bmatrix}^{T}. \quad (14)$$

The least squares solution $\tilde{\boldsymbol{X}}$ of Eqn. (12) is solved by the pseudo-inverse matrix $\boldsymbol{B}^{+}$ of $\boldsymbol{B}$ as follows,

$$\tilde{\boldsymbol{X}} = \boldsymbol{B}^{+}\boldsymbol{b}. \quad (15)$$

3.4 procedure of actual camera calibration

The distance measurement with enough high accuracy is necessary in the range of motion of the arm. Namely, the important measurement distances are as follows: about 1500 [mm] for rough scene understanding; about 1000 [mm] for searching of grasping point on the target object; about 500 [mm] for judgment to hold the target object by the gripper. Therefore, we developed a calibration rig where 30 target patterns were arranged as shown in Fig. 7. This calibration rig is set up to become a constant position to the installed cameras on the robot arm by using several tripods, and when the robot starts, it is used according to the following procedures if calibration is necessary. A comparatively high accuracy calibration can be achieved in the depth direction by using this calibration rig.
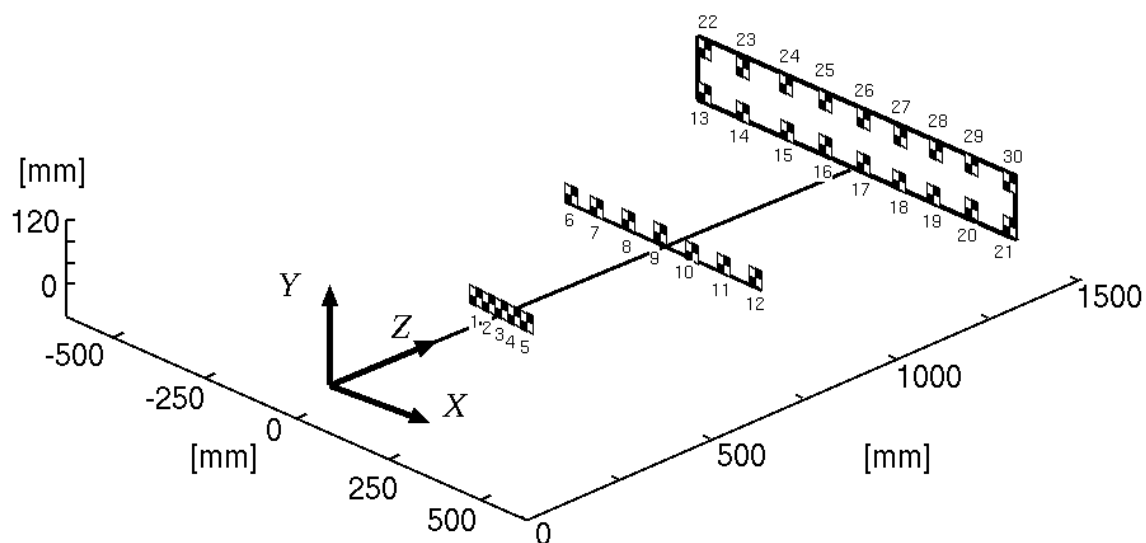

Fig. 7. Configuration of a calibration rig.

The camera calibration by using the rig might be executed in outdoor; therefore the detection of the target patterns by steady accuracy in the various outdoor environments is necessary. In this research, we employ the LoG (Laplacian of Gaussian) filter (Marr, 1982) to search for the position of the target patterns from camera images including noise, brightness

change, geometric transformation, etc. with stability and fast computing. It is well known that the LoG filter is isotropic band pass filter capable of reducing low and high frequency noise and completely removing both constant and linear-varying intensity. The LoG filter is represented by

$$\nabla^2 G(\boldsymbol{u}) \triangleq \frac{-1}{\pi \sigma^4} \left(1 - \frac{\|\boldsymbol{u}\|^2}{2\sigma^2} \right) e^{-\frac{\|\boldsymbol{u}\|^2}{2\sigma^2}}, \tag{16}$$

where $\boldsymbol{u} = (u^{(*)}, v^{(*)})^T$ (* is left or right) represents a unit vector on the projection plane. Namely, this filter eliminates high frequency noise and local brightness change by Gaussian, and eliminates low frequency brightness change by Laplacian. Fig. 8 shows the example of applying this filter to the image of the calibration rig. The target patterns are detected by the subpixel matching with template image after this filtering is applied in each, and the camera calibration is executed by the above mentioned procedure.

(a) left camera image   (b) right camera image

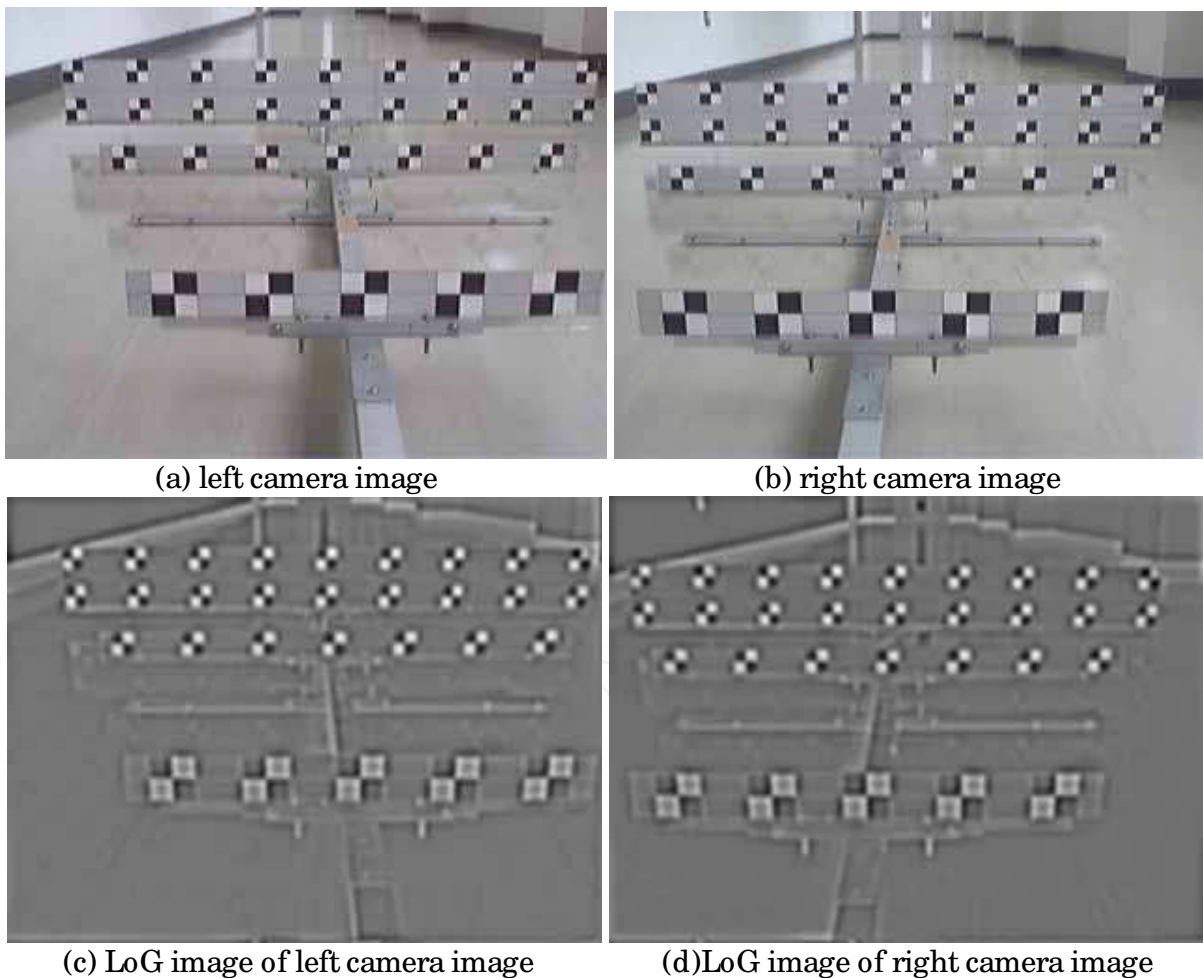(c) LoG image of left camera image   (d)LoG image of right camera image

Fig. 8. Captured images and LoG filtering images of the calibration rig.

The installation position and angles of the cameras were decided by the trial and error in consideration of the structure of the gripper of T-52, and then, the calibration accuracy according to the procedure is shown in Fig. 9. It is understand from this figure that the error

of the calibration was within 100 [mm] in the range of motion of the arm, and this camera calibration procedure was able to decrease the error of the target pattern extraction and to achieve the highly accurate calibration under the outdoor environment. Here, these errors include the setting error of the rig. To cancel these errors, we are developing now the method of extraction of characteristic feature points of the gripper for automatic calibration. The camera parameters obtained by calibrations above are used to calculate the distance to the object in the following camera systems.
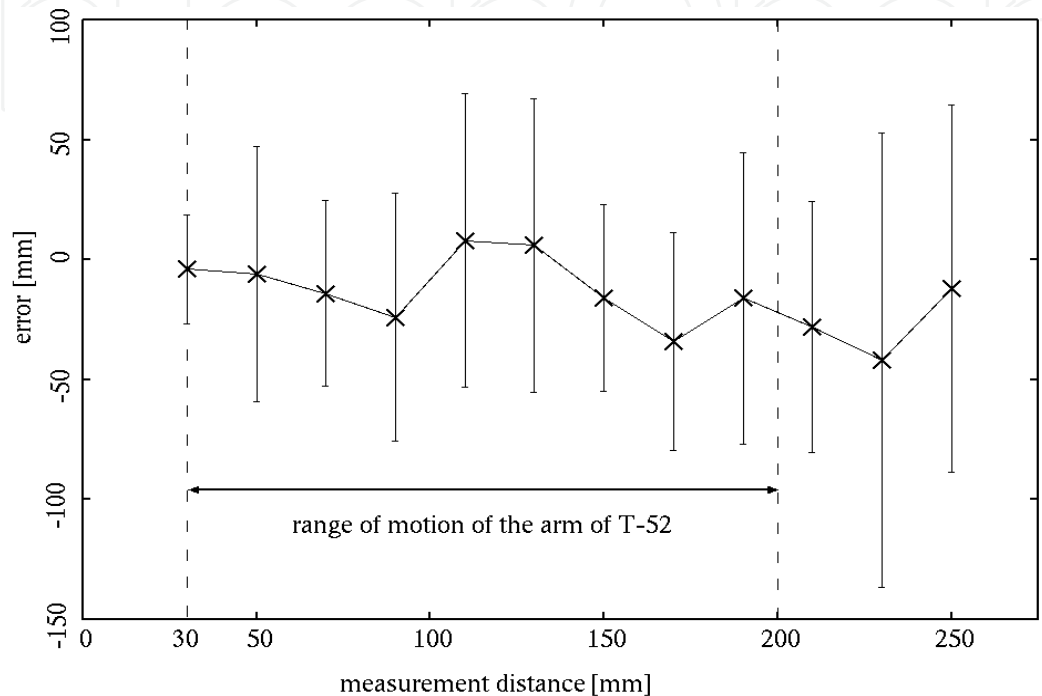

Fig. 9. Calibration accuracy vs. measurement distance.

4. System interfaces

Here, two kinds of developed interfaces are described: one is a monitor system that provides the sensor information efficiently for the pilot, and another is an interaction display system that tracks and displays the distances of the places that are set by the pilot. The image processing algorithms of them are different; the former is designed to present overall information of surroundings: the latter is designed to provide specific information desired by the pilot and to execute the processing in video rate. In the following, these systems are explained respectively.

4.1 Display of overall information on scene

The pilot understands a situation around the robot from information of the installed various sensors and operates it remotely. Although it is desirable that a lot of information is presented to the pilot, on the other hand, it is desirable for the system to be intuitive on the presentation of information, and to be designed not to accumulate tiredness in the pilot for efficiency operation. Therefore, we developed an intuitive interface system that indicates posture of the arm to the ground and distance surroundings of a target object. Fig. 10 shows an example of a monitor image presented to the pilot in the remote place by this system.
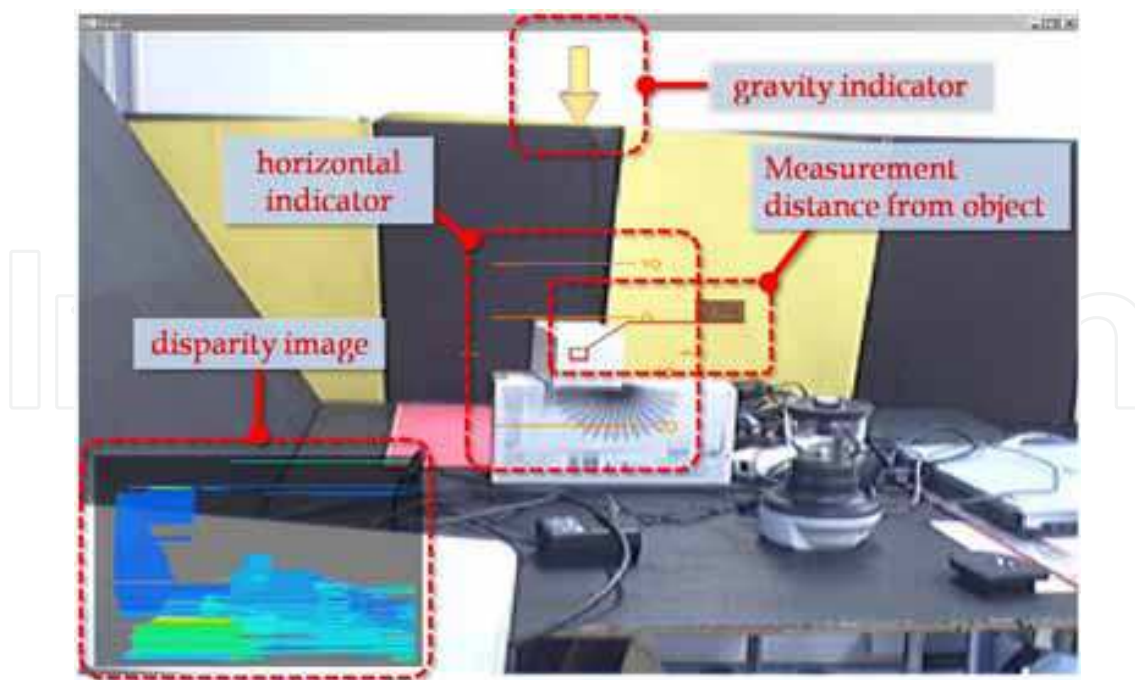
Fig. 10. Monitor image presented to the pilot. The system generates a gravity indicator, a horizontal indicator, a disparity image window, and a distance indicator.

4.1.1 Gravity indicator and disparity image

A gravity indicator that shows the direction of ground in the camera system was arranged at the upper part on the screen, and a horizontal indicator was arranged at the center part on the screen. These are displayed based on information obtained with the motion sensor: the rotation angles to the world coordinate system are calculated by integrating the acceleration x-y-z axes in the camera coordinate system, and they are re-displayed in the video rate, i.e. 30 [fps]. Moreover, a disparity image generated by a dynamic programming method (Birchfield et al. 1988) is displayed at the lower right of the screen, and the distance from the object is shown at the center part. Since the computational complexity of these processing based on the stereo vision is comparatively high, these indicators are re-displayed in the rate of 3-5 [fps]. The pilot can acquire various information by these displays without greatly changing the glance on the monitor. Examples of transition of display according to motion of the camera system are shown in Fig. 11. We see from Fig. 11 that the disparity images had been including the noise caused by matching error of the dynamic programming, and the property of the algorithm will be considered in the following.

4.1.2 Generation of disparity image in outdoor environments

The Birchfield's method is an algorithm to generate the disparity image efficiently by using a rectified parallel stereo camera (i.e. the image planes of two cameras are coplanar and their scanlines are parallel to the baseline). Moreover, the adjustments of five parameters, namely maximum disparity, occlusion penalty, match reward, reliability threshold, and reliability buffer factor are required for the execution of this algorithm. It is necessary to adjust these parameters in consideration of distance from target object, complexity of the scene, etc. by trial and error method after the installation of the camera system. Furthermore, the synchronization of the white balance and the shutter timing and the stable brightness images of the scene are needed for execution of this algorithm. The experimental results of

this algorithm on some conditions are shown in Fig. 12. Although the disparity image was generable efficiently when a steady brightness images were able to be acquired in the indoor environment, much noise was caused in the disparity images generated from the images with "black outs or white outs" occurred due to shortages of dynamic range or with complex background. Namely, it has been understood that the condition concerning the environment might be limited for the use though this system efficiently provides the pilot various information.
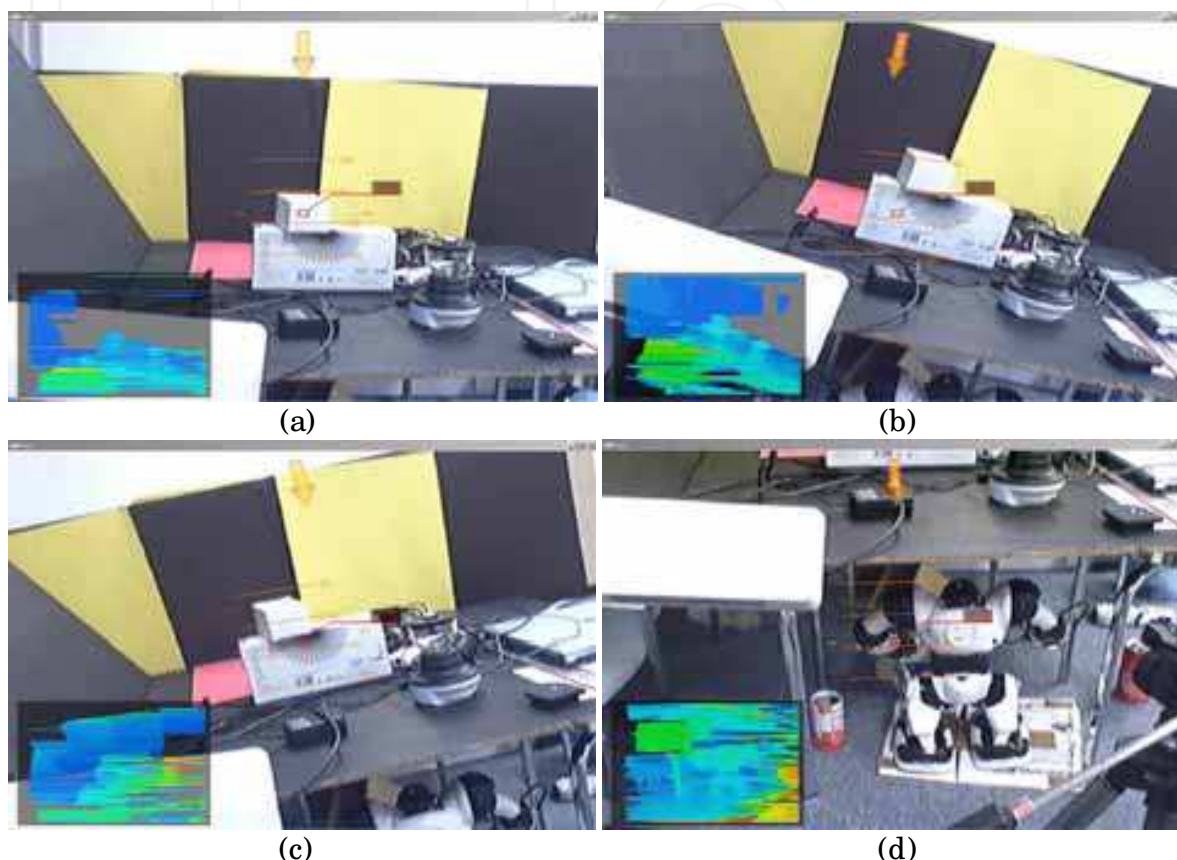

(a)   (b)

(c)   (d)

Fig. 11. Example of transition of display according to motion of the camera system.

4.2 Target tracker and distance indicator

Next we aimed to construct an interface executable at high speed and robust against the change of the environmental condition. The outline of an operating flow of the developed system is as follows. First of all, the pilot touches the monitor to specify the target points for tracking, and small windows (e.g. 9×9 pixels) are generated around the set positions on the monitor. Next, the tracking and the distance measurement of the set points are executed at once, and the results of them are reflected on the monitor in real time and the distance is displayed on the window frame. When the pilot touches the tracking window again, the set point is released. Moreover, the color of the window frame shows the following meanings: "green color" represents the success of the tracking and the measurement; "blue color" represents that the tracking point exists in the distance where it can be grasped by the gripper, i.e. the point exists near less than 1 [m]; "yellow color" represents the failure of the stereo matching and the displayed distance is untrustworthy. As mentioned above, this system can be used by an easy operation of the touch panel. The procedure of the processing of this system is described as follows.

(a) left camera image   (b) right camera image   (c) disparity image

(d) left camera image   (e) right camera image   (f) disparity image

(g) left camera image   (h) right camera image   (i) disparity image
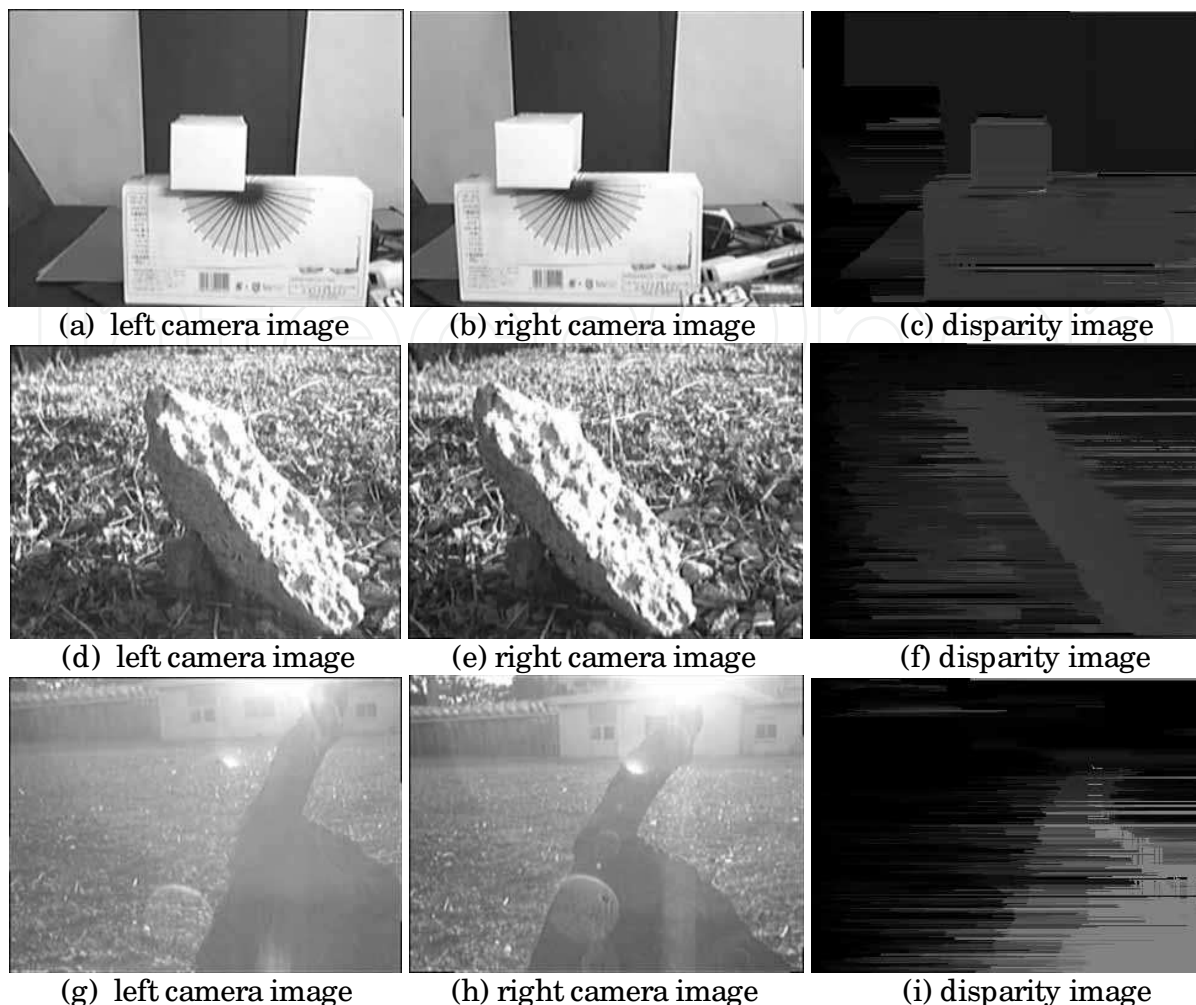
Fig. 12. Experimental results of generating of disparity images on several conditions: (a), (b), and (c) show results of an indoor environment; (d), (e), and (f) show results of an outdoor environment; and (g), (h), and (i) show results of a backlighting condition.

4.2.1 Detection algorithm

Considering of the following items is necessary for the tracking and the distance measurement of specified points under the outdoor environment.

- Keep the image processing speed at about 15-30 [fps].
- Ambient light changing.
- Camera shakes from vibration of the arms by oil pressure drive and mis-registration of the cameras.
- Achromatic color and complex scenes at the disaster scene.

Therefore, we developed an algorithm for simultaneous execution of the tracking and the stereo matching by using the Harris corner detector (Harris et al. 1988) and a fast feature tracking algorithm (Bouguet et al. 2000). The Bouguet's algorithm is a pyramidal implementation of the Lucas Kanade feature tracker description (Lucas et al. 1981), and it generates the optical flow between two images for the given set of points with sub-pixel accuracy. The proposed image processing algorithm is executed according to the following procedures.

4.2.2 Flow of target tracking and measurement

Let I_t and J_t be images captured by the left and the right camera at time t, and let $I_t(\boldsymbol{u})$ ($\boldsymbol{u} \equiv [x,y]^T$) be a pixel on I_t.

1. The pilot sets the target point $I_t(\boldsymbol{u}_n)$ on I_t, where n represents the index of target.
2. Generate a tracing window

$$W_{n,t}^I \triangleq \left\{ I_t(\boldsymbol{u}_n) + \boldsymbol{w} \right\},$$

   where, $\boldsymbol{w} = [\pm w, \pm w]^T$ (e.g. $w = 4[\text{pixel}]$).

3. Find the corner points in $W_{n,t}^I$ by employing the Harris corner detector. If more than three points are not found around $I_t(\boldsymbol{u}_n)$, then the set point is canceled.

4. Generate a search window

$$S_{n,t}^J \triangleq \left\{ J_t(\boldsymbol{u}_n + \boldsymbol{b}) + \boldsymbol{s} \right\}$$

   on J_t, where $\boldsymbol{b} = [b_x, 0]^T$, $b_x = \text{const.}$ is determined based on the camera parameters, and $\boldsymbol{s} = [\pm s, \pm s]^T$ (e.g. $s = 10[\text{pixel}]$) represents the size of the search window.

5. Find

$$W_{n,t}^J \triangleq \left\{ J_t(\boldsymbol{u}_n + \boldsymbol{d}^J) + \boldsymbol{w} \right\} \in S_{n,t}^J,$$

   corresponding to $W_{n,t}^I$ on J_t by employing the Bouguet's algorithm (Fig. 13(b):(a-0)), where $\boldsymbol{d}^J \triangleq [d_x^J, d_y^J]^T$.

6. Calculate the distance to the point $I_t(\boldsymbol{u}_n)$ by using $\boldsymbol{d}^J$ and the calibrated camera parameters. If $d_y^J > d_{\text{threshold}}$, the result is judged as untrustworthy.

7. Find $I_{t+1}(\boldsymbol{u}_n + \boldsymbol{d}^I)$, i.e. the feature point in the next frame image corresponding to feature points of $I_t(\boldsymbol{u}_n)$ by employing the Bouguet's algorithm for the feature tracking (Fig. 13(b):(b-0)), and go to the 2nd step (Fig. 13(b):(a-1)).

The distance measurement and the target tracking are executed simultaneously by iteration of this procedure. Each window parameter is decided based on the range of motion of the arm and the camera parameters etc. beforehand. The experimental results of this algorithm implemented to the interface are shown in Fig. 14. Since this interface provides the pilot information only on the target points based on Bouguet's algorithm, it was possible to adapt to the tracking of fast movement object and the oscillation of the camera system.
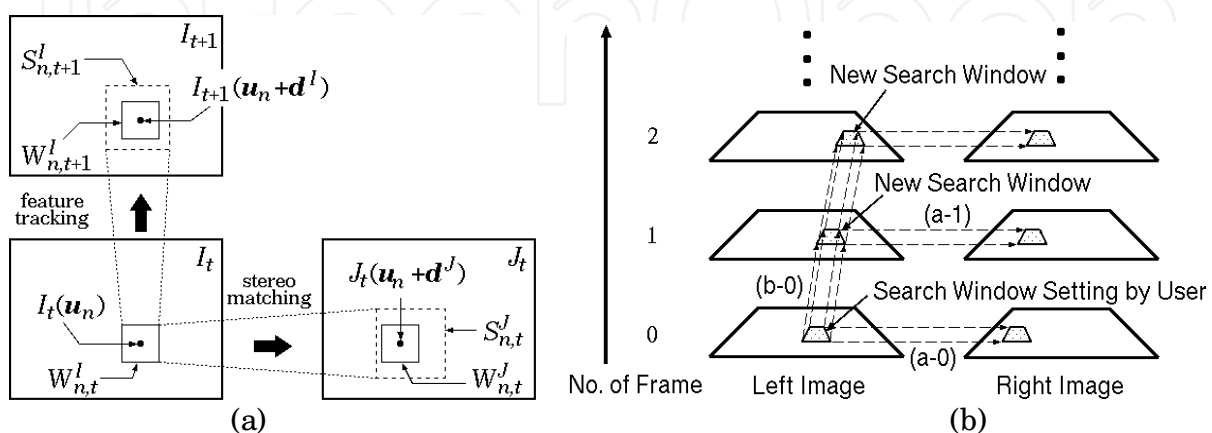

Fig. 13. Outline of the target tracking and the feature tracking: (a) relation of the target points, the tracking windows, and the search windows; (b) relation of the windows on the time flow.

(a)   (b)

(c)   (d)

(e)   (f)

Fig. 14. Execution results of the proposed algorithm for simultaneous execution of the tracking and the stereo matching. Two small windows at the center part of images are shown the tracking windows, and the windows at the bottom of the images are shown parameters of the tracking points. The target points were not lost though the camera system was intensely shaken for five seconds. Although the stereo matching failed on two frames (the color of target window changed to yellow as shown in (d)), the distance measurement succeeded in almost all frames.

4.2.3 Experiment with robot

The result of the experiment conducted by installing the developed interface in the robot is shown in Fig. 15. These images were captured by the left camera of the B-system mounted on the joint of the gripper of T-52, and a rectangular lumber was set in front of the gripper. The tracking windows at the upper part of the image were specified by the pilot and a figure in the window shows the order of specifying the feature point (first and second points had

been canceled). Moreover, the figure at upper part of the tracking window shows the measured distance, and the lower squares of the images and its inner figures show the order of the feature points and 3D position in a robot arm coordinate system, respectively. The color of the window changes according to the distance and the state of measurement as mentioned in **4.2**. It was found from the experiment that the interface was able to provide the pilot with the measurement distance of the tracking points with stability, and the operation to grasp the rectangular lumber was executed smoothly.

(a)

(b)

(c)

(d)

(e)

(f)
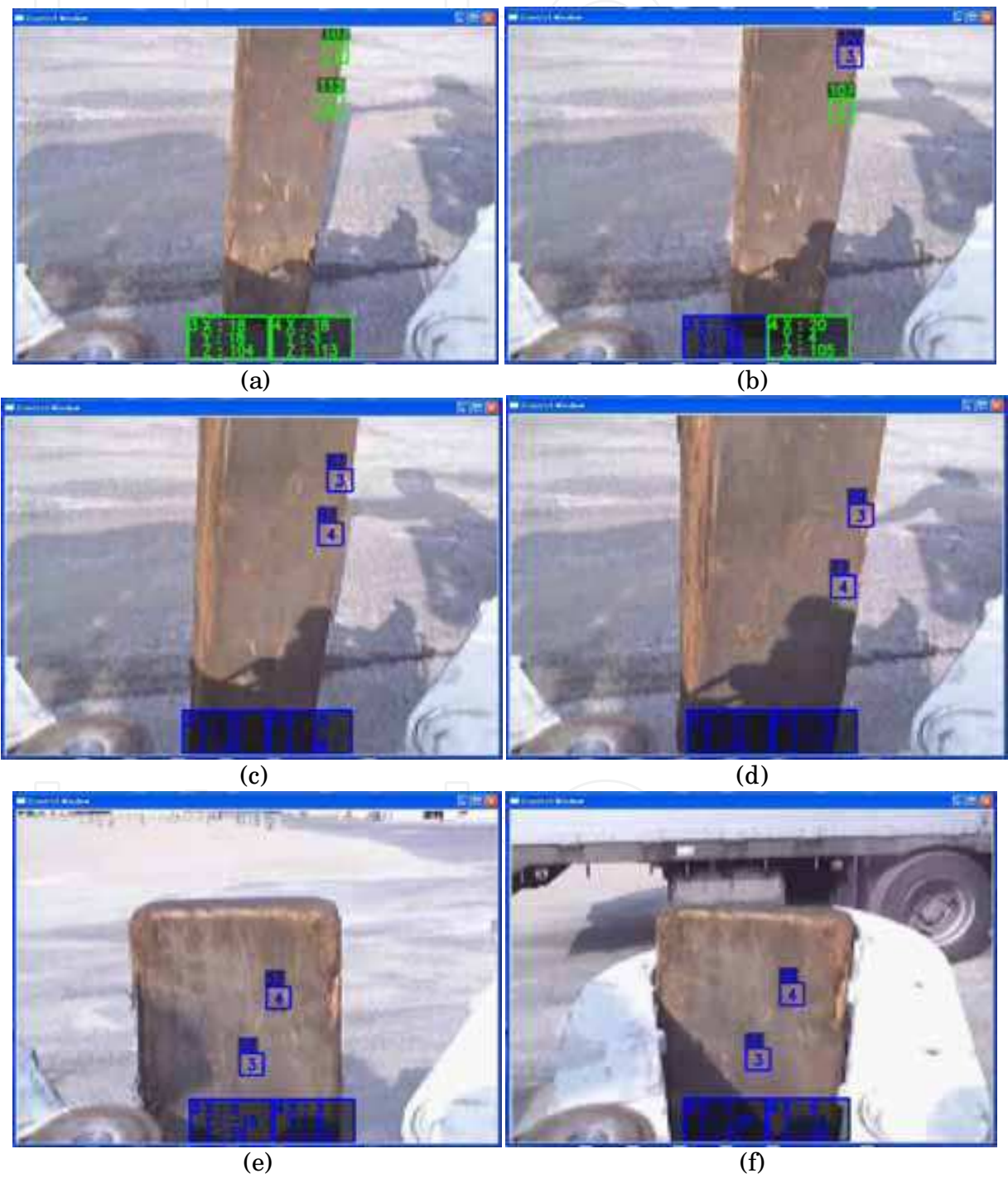
Fig. 15. Results of an experiment of an operation to grasp a rectangular lumber by using B-system installed in the joint of gripper of T-52. These images were captured by the left camera.

5. Conclusion

In this research, we developed two kinds of pilot assistance systems for T-52 and T-53 working in a disaster district. One is an interface that provides the information from several sensors efficiently for the pilot. Another is an interface that tracks and displays the distances of the points that are set by the pilot. This system can tell the pilot the distance by the change in the color that understands intuitively easily, and the processing speed is also fast enough. The pilot switches the two kinds of proposed interfaces according to the situation. The former is suitable for searching of a target object by the camera installed in the robot arm, and the latter is suitable for the operation that grasps a specific position of the target object by the arm. Then, the basic experiments for the construction of each system were conducted, and the effectiveness of them was confirmed.

From now on, a further various tests and examinations will be necessary to use of these system and interface efficiently in the disaster scene. Moreover, more interface development research according to disaster scene to achieve the many kinds of work by the rescue robot will be necessary. We are continuing the research for the solution of these research topics now, furthermore have started the development for automatic work system of the robots.

6. Acknowledgements

7. References

Birchfield, S. & Tomasi C. (1988). Depth Discontinuities by Pixel-to-Pixel Stereo, *Proc. of the 6th IEEE International Conference on Computer Vision*, pp. 1073–1080.

Bouguet, J. Y. (2000). Pyramidal Implementation of the Lucas Kanade Feature Tracker: Description of the algorithm, *Technical Report, Intel Corporation, Microprocessor Research Labs 2000, OpenCV documentation*, 2680.

Harris, C. & Stephens., M. (1988). A Combined Corner and Edge Detector, *Proc. of Alvey Vision Conf.,* pp. 147–151.

Lucas, B. D. & Kanade, T. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision, *Proc. of the 7th International Joint Conference on Artificial Intelligence*, pp. 674–679.

Marr, D. (1982). Vision: A Computational Investigation into the Human Presentation and Processing of Visual Information, *W. H. Freedman Company*.

Tsai, Y. Roger. (1987). Synopsis Recent Progress on Camera Calibration for 3D Machine Vision, The Robotcs Review 1, Oussama Khatib, John J. Craig and Tomas Lozano-Perez, pp.147–159, MIT Press.

Yokokohji, Y. (2006). ENRYU T-52 Advance -A Teleoperated Heavy-Duty Robot for Assisting Rescue Activities", *Journal of Robotics Society of Japan*, Vol.24, No.2, pp.113. (in Japanese)

Nishida, T., Koga, H., Fuchikawa, Y., Kitayama, Y., Kurogi, S., & Arimura, Y. (2006). Development of Pilot Assistance System with Stereo Vision for Robot Manipulation, *Proc. of SICE-ICCAS 2006*, pp. 2675-2680.

Hartley, R. & Zisserman, A. (2003). Multiple View Geometry in Computer Vision Second Edition, *Cambridge Univercity Press*.

Computer Vision

Edited by Xiong Zhihui

This book presents research trends on computer vision, especially on application of robotics, and on advanced approachs for computer vision (such as omnidirectional vision). Among them, research on RFID technology integrating stereo vision to localize an indoor mobile robot is included in this book. Besides, this book includes many research on omnidirectional vision, and the combination of omnidirectional vision with robotics. This book features representative work on the computer vision, and it puts more focus on robotics vision and omnidirectioal vision. The intended audience is anyone who wishes to become familiar with the latest research work on computer vision, especially its applications on robots. The contents of this book allow the reader to know more technical aspects and applications of computer vision. Researchers and instructors will benefit from this book.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Takeshi Nishida, Shuichi Kurogi, Koichi Yamanaka, Wataru Kogushi and Yuichi Arimura (2008). Development of Pilot Assistance System with Stereo Vision for Robot Manipulation, Computer Vision, Xiong Zhihui (Ed.), ISBN: 978-953-7619-21-3, InTech, Available from:
http://www.intechopen.com/books/computer_vision/development_of_pilot_assistance_system_with_stereo_visi on_for_robot_manipulation

INTECH
open science | open minds