# 12

# Development of a Sensor System for an Outdoor Service Robot

Takeshi Nishida[1], Masayuki Obata[2],
Hidekazu Miyagawa[2] and Fujio Ohkawa[1]
[1]*Kyushu Institute of Technology,*
[2]*YASKAWA INFORMATION SYSTEMS Corporation*
*Japan*

## 1. Introduction

In general, service robots are equipped with multiple types of sensor for environmental recognition and to avoid the measurement error that occurs by various measurement noises. Especially for the robots that work in outdoor environments, cameras and LRFs (Laser Rangefinders) are the most useful sensor devices, and they have been installed into many prototype service robots. Robots can acquire texture, color, shadow, etc. of objects or a scene via cameras, and execute various tasks, e.g. landmark recognition, face recognition, target tracking etc. based on those information. Moreover, the stereovision composed by using two or more cameras can acquire 3D information on the scene. However, the distance measurement of the objects with difficulty of decision of correspondence of feature points, such as walls without texture and shadow, is difficult. Furthermore, when the strength of environmental light change that exceeds the dynamic range of the camera image sensor occurs, the measurement accuracy greatly falls (DeSouza et al., 2002). On the other hand, LRF is a device which uses a laser beam in order to determine the distance to a reflective object, and then the distance with comparatively high accuracy can be measured even in the situation in which the measurement with the camera becomes unstable. Therefore, the LRF is frequently used for localization, map building, and running route inspection of autonomous mobile robots. However, the calculation algorithm to recognize the target object by using the LRF is very complex, the calculation cost is also high, and have following disadvantages:

1. The range data sometimes involve lack of data called black spots around the corner or the edge of the objects.
2. The range data involves quantization errors owing to measurement resolution (e.g. 10 [mm]).
3. The number of data points in a range data set is large.
4. Texture and color information etc. on the object cannot be acquired.

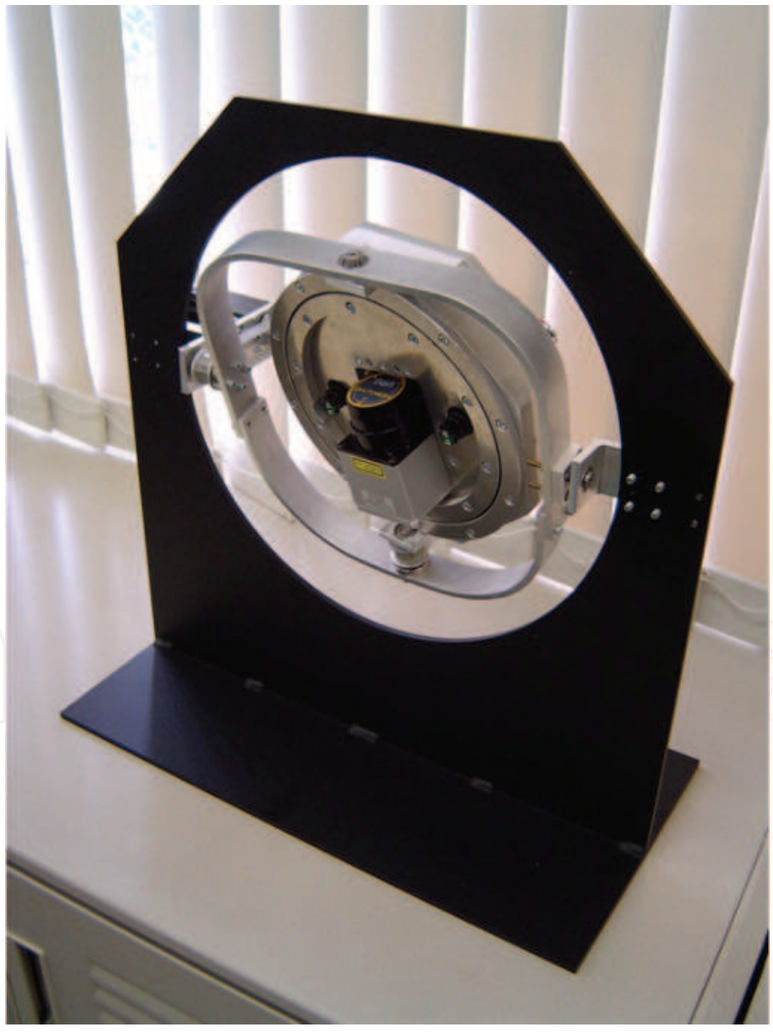
Therefore, we have developed a novel sensor system that consists of two cameras and a LRF; various types of measurement and recognition of targets are possible according to those combinations. Moreover, the sensor system has both advantages of camera and LRF, and has the robustness against environment variations. In this chapter, we show the

methods of target recognition and 3D posture measurement using the sensor system after giving explanation about the construction of hardware. Furthermore we show the results of several outdoor experiments of the robot equipped with the sensor system.

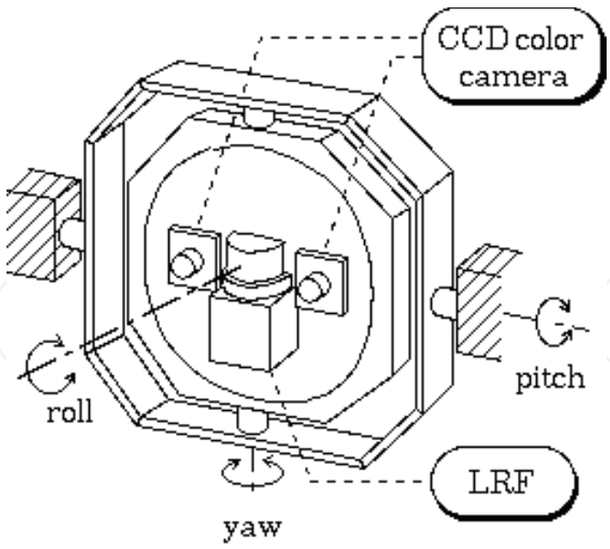
## 2. Structure of sensor system

### 2.1 Devices

The sensor system (Fig. 1) consists of two cameras, a LRF, and three stepping motors. The maximum measurement distance of the LRF (URG-04LX) is 4 [m], and the measurement error margin is less than 1 [%]. The infrared laser is radially irradiated from the center part of the LRF, its range of the measurement is forward 170 [deg] (the maximum range is 240 [deg]), and the angle resolution is about 0.36 [deg]. The CCD color cameras capture images in 24-bit color at $640 \times 480$ resolution at rates to 30 [fps]. Moreover, these installation positions have been designed so that those optical axes and the measurement plane of LRF are parallel.



(a) Sensor system




(b) Configuration of the sensor system

Fig. 1. Sensor system developed for an outdoor robot.

### 2.2 Coordinate systems

The definition of each coordinate system and the relations between them are shown here. At first, Table 1 shows the definition of each coordinate system, and the relations between the cameras and the LRF are shown in Fig. 2.

| | |
|---|---|
| World coordinate system | $\Sigma_w = \left\{ {}^w x, {}^w y, {}^w z \right\}$ |
| Robot coordinate system | $\Sigma_r = \left\{ {}^r x, {}^r y, {}^r z \right\}$ |
| LRF coordinate system | $\Sigma_l = \left\{ {}^l x, {}^l y, {}^l z \right\}$ |
| Left camera coordinate system | $\Sigma_{cl} = \left\{ {}^{cl} x, {}^{cl} y, {}^{cl} z \right\}$ |
| Right camera coordinate system | $\Sigma_{cr} = \left\{ {}^{cr} x, {}^{cr} y, {}^{cr} z \right\}$ |
| Left camera screen coordinate system | $\Sigma_{sl} = \left\{ {}^{sl} x, {}^{sl} y \right\}$ |
| Right camera screen coordinate system | $\Sigma_{sr} = \left\{ {}^{sr} x, {}^{sr} y \right\}$ |
| Right hand coordinate system | $\Sigma_{hr} = \left\{ {}^{hr} x, {}^{hr} y, {}^{hr} z \right\}$ |
| Right hand-eye screen coordinate system | $\Sigma_{hsr} = \left\{ {}^{hsr} x, {}^{hsr} y \right\}$ |

Table 1. Coordinate systems

Let $\boldsymbol{R}_X(\cdot)$, $\boldsymbol{R}_Y(\cdot)$, and $\boldsymbol{R}_Z(\cdot)$ be the rotation matrices for each axis of the robot coordinate system. The homogeneous coordinate transformation matrix from the sensor coordinate system $\Sigma_r$ to the LRF coordinate system $\Sigma_l$ is represented as follows,

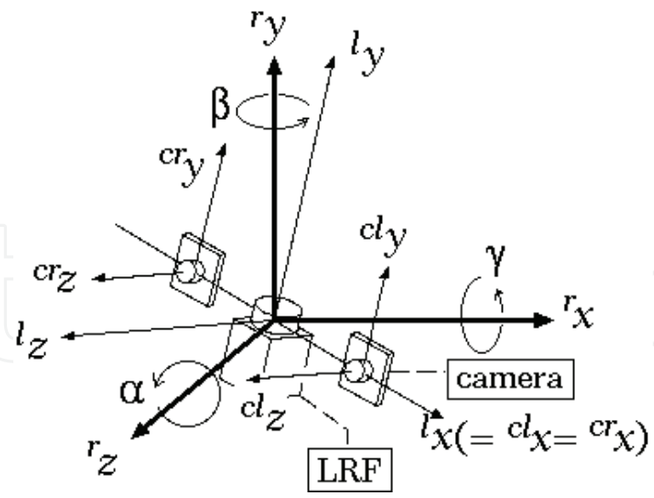



Fig. 2. Relations of coordinate systems

$$ {}^{r}_{l}\boldsymbol{H} \triangleq \begin{pmatrix} {}^{r}_{l}\boldsymbol{R} & {}^{r}_{l}\boldsymbol{p} \\ \boldsymbol{0} & 1 \end{pmatrix} \in \mathbb{R}^{4\times 4}, \tag{1} $$

where,

$$ {}^{r}_{l}\boldsymbol{R} \triangleq \boldsymbol{R}_Z(\alpha)\cdot\boldsymbol{R}_Y(\beta)\cdot\boldsymbol{R}_X(\gamma), \tag{2} $$

$$ \boldsymbol{R}_X(\gamma) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{pmatrix}, \quad \boldsymbol{R}_Y(\beta) = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix}, \quad \boldsymbol{R}_Z(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}, \tag{3} $$

and ${}^{r}_{l}\boldsymbol{p}$ represents a translation vector from the origin of $\Sigma_r$ to the origin $\Sigma_l$. From these relations, the mapping from a measurement point ${}^{l}\boldsymbol{x}_{ij} \triangleq \left({}^{l}x_{ij},\ {}^{l}y_{ij},\ {}^{l}z_{ij}\right)^T \in \Sigma_l$ to the point ${}^{r}\boldsymbol{x}_{ij} \triangleq \left({}^{r}x_{ij},\ {}^{r}y_{ij},\ {}^{r}z_{ij}\right)^T \in \Sigma_r$ is described by

$$ \begin{pmatrix} {}^{r}\boldsymbol{x}_{ij} \\ 1 \end{pmatrix} = {}^{l}_{r}\boldsymbol{H} \begin{pmatrix} {}^{l}\boldsymbol{x}_{ij} \\ 1 \end{pmatrix} = {}^{r}_{l}\boldsymbol{H}^{-1} \begin{pmatrix} {}^{l}\boldsymbol{x}_{ij} \\ 1 \end{pmatrix}, \tag{4} $$

where, $i = 1,\cdots I, j = 1\cdots J$ and

$$ {}^{r}_{l}\boldsymbol{H}^{-1} = \begin{pmatrix} {}^{r}_{l}\boldsymbol{R}^T & -{}^{r}_{l}\boldsymbol{R}^T\,{}^{r}_{l}\boldsymbol{p} \\ \boldsymbol{0} & 1 \end{pmatrix}. \tag{5} $$

Next, the homogeneous coordinate transformation matrices from the LRF coordinate system $\Sigma_l$ to the left and the right camera coordinate system $\Sigma_{cl}$ and $\Sigma_{cr}$ are represented as follows,

$$ {}^{l}_{cl}\boldsymbol{H} = \begin{pmatrix} {}^{l}_{cl}\boldsymbol{R} & {}^{l}_{cl}\boldsymbol{p} \\ \boldsymbol{0} & 1 \end{pmatrix}, \quad {}^{l}_{cr}\boldsymbol{H} = \begin{pmatrix} {}^{l}_{cr}\boldsymbol{R} & {}^{l}_{cr}\boldsymbol{p} \\ \boldsymbol{0} & 1 \end{pmatrix}, \tag{6} $$

where, ${}^{l}_{cl}\boldsymbol{p}$ (or ${}^{l}_{cr}\boldsymbol{p}$) is translation vector from the origin of $\Sigma_l$ to the origin of $\Sigma_{cl}$ (or $\Sigma_{cr}$). Therefore, the posture of the left (right) camera against $\Sigma_r$ is described by

$$ {}^{r}_{cl}\boldsymbol{H} = {}^{r}_{l}\boldsymbol{H}\,{}^{l}_{cl}\boldsymbol{H}, \quad ({}^{r}_{cr}\boldsymbol{H} = {}^{r}_{l}\boldsymbol{H}\,{}^{l}_{cr}\boldsymbol{H}). \tag{7} $$

Here, the sensor system is designed to be ${}^{l}_{cl}\boldsymbol{R} \approx \boldsymbol{I}_{3\times3}$, ${}^{l}_{cr}\boldsymbol{R} \approx \boldsymbol{I}_{3\times3}$ (i.e. ${}^{l}_{r}\boldsymbol{R} = {}^{cl}_{r}\boldsymbol{R} = {}^{cr}_{r}\boldsymbol{R}$), and ${}^{l}_{cl}\boldsymbol{p} = -{}^{l}_{cr}\boldsymbol{p}$ for easiness. Furthermore, the mapped point ${}^{sl}\boldsymbol{x}_{ij}$ of ${}^{l}\boldsymbol{x}_{ij} \in \Sigma_l$ on $\Sigma_{sl}$ is represented as follows,

$$ {}^{r}_{cl}\boldsymbol{H}\begin{pmatrix} {}^{r}\boldsymbol{x}_{ij} \\ 1 \end{pmatrix} = \begin{pmatrix} {}^{r}_{l}\boldsymbol{R}\cdot\left({}^{r}\boldsymbol{x}_{ij} + {}^{l}_{cl}\boldsymbol{p}\right) + {}^{r}_{l}\boldsymbol{p} \\ 1 \end{pmatrix} \triangleq \begin{pmatrix} {}^{cl}\boldsymbol{x}_{ij} \\ 1 \end{pmatrix}, \tag{8} $$

$$ {}^{cl}_{sl}\boldsymbol{f}\left({}^{cl}\boldsymbol{x}_{ij}\right) \triangleq \begin{pmatrix} -f_{cl}\dfrac{{}^{cl}x_{ij}}{{}^{cl}z_{ij}}, & f_{cl}\dfrac{{}^{cl}y_{ij}}{{}^{cl}z_{ij}} \end{pmatrix}^{T} \triangleq {}^{sl}\boldsymbol{x}_{ij}, \tag{9} $$

where, $f_{cl}$ is the focal length of the left camera. Moreover, to convert the unit of ${}^{sl}\boldsymbol{x}_{ij}$ from the distance into the pixel, the following transformation is defined.

$$ F\left({}^{sl}\boldsymbol{x}_{ij}\right) \triangleq \begin{pmatrix} s_x\cdot\,{}^{sl}x_{ij} \\ s_y\cdot\,{}^{sl}y_{ij} \end{pmatrix} \triangleq {}^{sl}\boldsymbol{x}_{mn}, \tag{10} $$

where, $M = \{m \,|\, m \in \mathbb{Z}_+, m \le |M|\}$ and $N = \{n \,|\, n \in \mathbb{Z}_+, n \le |N|\}$ are width and height pixel indexes of the $|M|\times|N|$ image, and $(s_x, s_y)$ are scale factors to transform them from the measurement distance to the pixel number. Hence, the measurement data points of the LRF are mapped into the left camera screen pixels by Eqn. (8), (9), (10), and the measurement data between devices is mutually mapped by the above relational expressions. In addition, notes of this system are enumerated as follows. Firstly, pinhole camera model is assumed and lens distortion of the camera is disregarded. Secondly, the values of ${}^{r}_{l}\boldsymbol{p}$, ${}^{l}_{cl}\boldsymbol{p}$, ${}^{l}_{cr}\boldsymbol{p}$ are already-known according to a prior calibration. Thirdly, if the system configuration with high accuracy is difficult, the calibration of ${}^{cl}_{r}\boldsymbol{R}$, ${}^{cr}_{r}\boldsymbol{R}$ are also necessary. Lastly, the rotation angles that provides for ${}^{l}_{r}\boldsymbol{R}$ can be measured by counting the pulse input of the stepping motors.

## 3. Object detection and target tracking

The applications for the object detection and the target tracking were developed on the assumption that this sensor system will be installed in the autonomous mobile robot. Therefore, to execute those tasks, high speed execution and robustness are demanded for the processing of the data obtained from the sensor system. Although the objects detection and the target tracking by the camera image processing are executable at high speed, they are influenced easily from the change in environmental light. On the other hand the calculation cost for the measurement and the object detection by the LRF is comparatively high; however, the measurement is robust under the influence of environmental light. Therefore, these devices of the sensor system are combined to complement each other for high accuracy detection and measurement. Namely, the target is measured with only one device when a limited measurement for fast computation is required, and is measured with multiple devices when highly accurate processing is required.

In this section, we will show applications of the sensor system, and one of the flowchart is shown in Fig. 3. In this application, at first, a target object is detected and tracked by camera image processing, and after that, the range data of the object obtained by rotating of the LRF analyzed. By such measurement flow, the object detection is quickly achieved by the camera image processing, and the analysis in detail of the target object is done by using the LRF data not influenced easily from an environmental change. In the following, we consider the procedure for detection and measurement of plastic bottles for a concrete example.

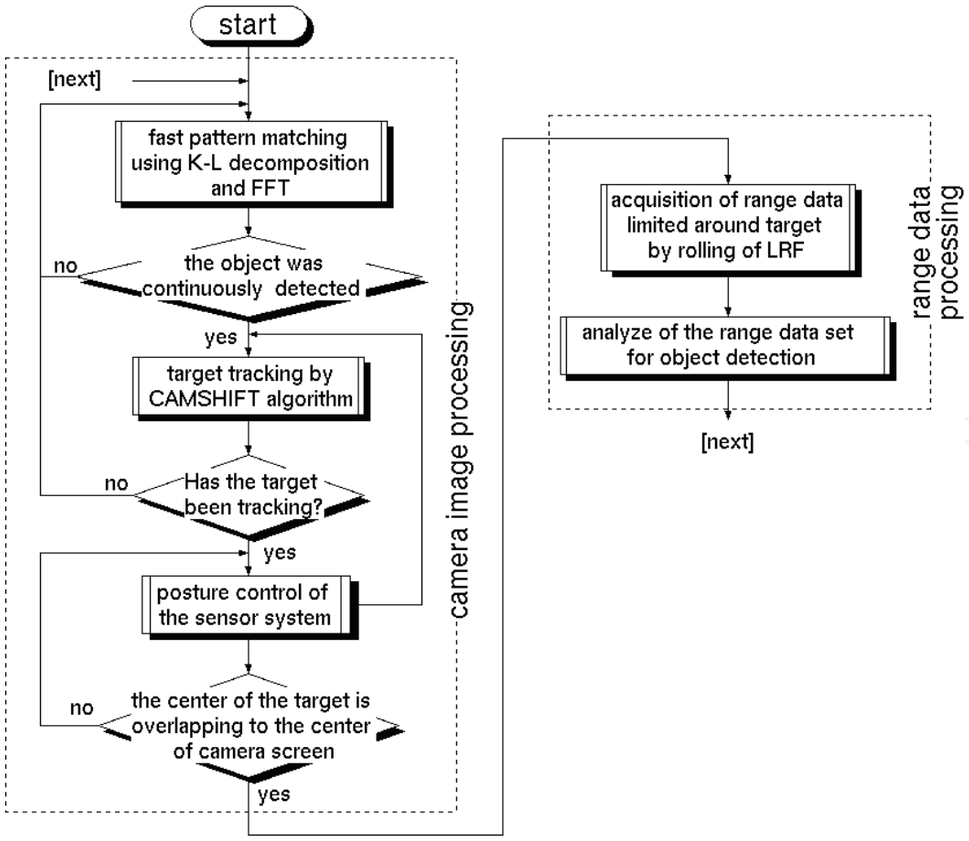



Fig. 3. Flowchart of an application of the sensor system.

### 3.1 Camera image processing

Since the appearance of trash changes according to position and posture, robustness to distortion and transformation of shape is necessary for the trash detection method. Therefore, from the point of view for robustness and fast execution, we employed the eigen space method (Uenohara & Kanade, 1997) for the object detection, and employed the CAMSHIFT method (Gray, 1998) for the target tracking (see (Fuchikawa et al., 2005) for detail). These methods are briefly explained as follows.

#### 3.1.1 Karhunen-Loeve expansion

Let $\boldsymbol{q}_l(\boldsymbol{x}) \triangleq \boldsymbol{q}_l(x,y),\ (l=1,\cdots,L)$ be a set of template images consisting of $r \times s$ pixels. Moreover let $\boldsymbol{q}_l$ be its $rs\ (=r\times s)$ dimensional vector in scan line order. The singular value decomposition (SVD) of a matrix $\boldsymbol{Q}=(\boldsymbol{q}_1-\bar{\boldsymbol{q}},\cdots,\boldsymbol{q}_l-\bar{\boldsymbol{q}},\cdots,\boldsymbol{q}_L-\bar{\boldsymbol{q}})\in\mathbb{R}^{rs\times L}$ is given by

$$\boldsymbol{Q}=\boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T, \tag{11}$$

where $\bar{\boldsymbol{q}}$ is average vector of $\boldsymbol{q}_l$, $\boldsymbol{U}=(\boldsymbol{u}_1,\cdots,\boldsymbol{u}_L)\in\mathbb{R}^{rs\times L}$ and $\boldsymbol{V}=(\boldsymbol{v}_1,\cdots,\boldsymbol{v}_L)\in\mathbb{R}^{L\times L}$ are orthogonal matrices, $\boldsymbol{D}=\mathrm{diag}(d_1,d_2,\cdots,d_L)\in\mathbb{R}^{L\times L}$ is a diagonal matrices, and $d_1\geq d_2\geq\cdots\geq d_L$.

There are some methods for detecting objects by using eigenvectors (eigen images). We assume that a robot works under outdoor environment, so robustness for changes of lighting is required. Since normalized cross-correlation is insensitive to the variation of intensity of the background, we employ a method obtaining normalized correlation approximately by using eigenvectors. Suppose we have major $K$ eigenvectors $\boldsymbol{u}_i$ and the following vector

$$\boldsymbol{u}_0=\frac{\bar{\boldsymbol{q}}-\boldsymbol{U}'\boldsymbol{a}^c}{\left\|\bar{\boldsymbol{q}}-\boldsymbol{U}'\boldsymbol{a}^c\right\|}, \tag{12}$$

where the coefficient vector $\boldsymbol{a}^c$ and $\boldsymbol{U}'$ are given by

$$\boldsymbol{a}^c=\boldsymbol{U}'^T\bar{\boldsymbol{q}}, \tag{13}$$

$$\boldsymbol{U}'=(\boldsymbol{u}_1,\cdots,\boldsymbol{u}_K)\in\mathbb{R}^{rs\times K}. \tag{14}$$

The number of eigenvector $K$ affects the processing speed and reliability of result. Here, the order $K$ is decided from the cumulative proportion

$$\mu^{(K)}=\frac{\sum_{k=1}^{K}d_k}{\sum_{l=1}^{L-1}d_l}, \tag{15}$$

and experimental recognition rate. Finally, the matrix composed of $\boldsymbol{u}_0$ and $\boldsymbol{U}'$ as follows

$$\bar{\boldsymbol{U}}\triangleq(\boldsymbol{u}_0,\boldsymbol{u}_1,\cdots,\boldsymbol{u}_K) \tag{16}$$

is used in the recognition processes. We use images of 12 kinds of plastic bottles as template images, and assume that the range of object detection is about 500 [mm] to 2000 [mm] from a robot, the images were captured as shown in Fig. 4. Moreover, we set the height of a camera to 400 [mm], the depression angle $\theta_p = 10;\ 20;\ 30;\ 40$ [deg], the rotation angle of a bottle $\theta_r = 0; 10; \cdots; 350$ [deg] (every 10 [deg]). Namely, we collected 432 template images with $150 \times 110$ pixels (Fig. 5 (a)). The centers of appearance of the plastic bottles and the centers of the images were adjusted to be overlapped, and the intensities of the background of the images adjusted to be zero. Moreover, for determining *K*, we conducted the object detection experiments by using several outdoor images that contains plastic bottles, examples of the experimental results are shown in Fig. 6. We decided *K* = 20 from these experiments, and the cumulative proportion was $\mu^{(20)} = 0.418$. Since it is advisable that the details such as difference of labels were disregarded for the generality of the object detection processing, the cumulative proportion was not set so high. The reconstructed images by using the dimension *K* are shown in Fig. 5(b), and we can see from these figures that the eigenvectors and coefficients kept features of plastic bottles.

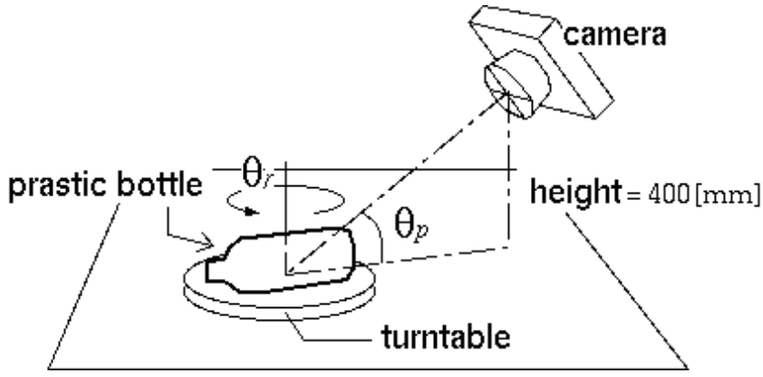
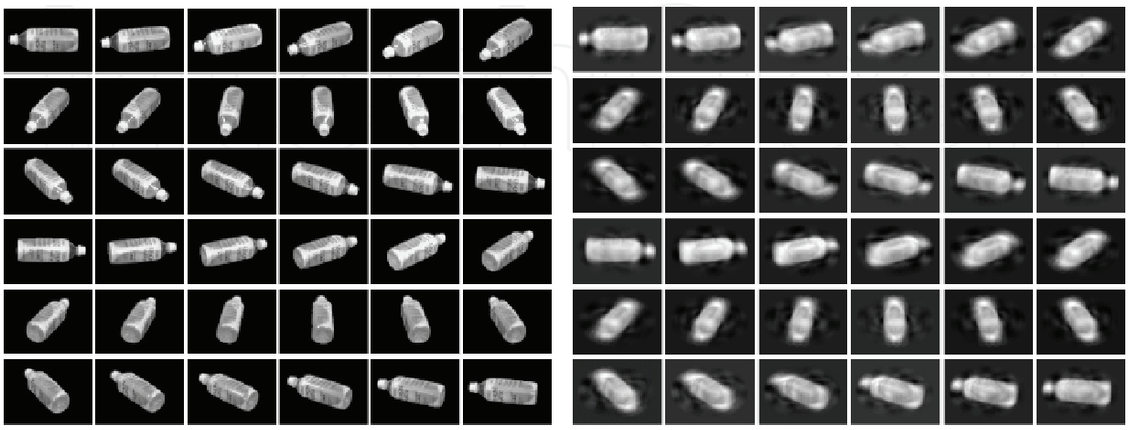



Fig. 4. Shooting conditions of template images.



(a) (b)

Fig. 5. Example of plastic bottle images: (a) 36 of 432 are shown; (b) reconstructed images.

Fig. 6. Examples of results of object detection experiments under outdoor environments.

### 3.1.2 Template matching in vector subspace

Let $p(\boldsymbol{x}) \in \mathbb{R}^{r \times s}$ be an input image extracted around the pixel $\boldsymbol{x}$ from a camera image, and $\boldsymbol{p} \in \mathbb{R}^{rs}$ be its vector. First, $\boldsymbol{p}$ is normalized and transformed to be zero mean vectors $\overline{\boldsymbol{p}}$, i.e. the average of vector elements is zero. The pattern $\tilde{\boldsymbol{q}}$ constructed by using the vector subspace is given by

$$\tilde{\boldsymbol{q}} = \sum_{i=0}^{K} (\boldsymbol{u}_i^T \overline{\boldsymbol{p}}) \boldsymbol{u}_i . \cdot \tag{17}$$

Furthermore, the normalized cross-correlation of $\overline{\boldsymbol{p}}$ and $\tilde{\boldsymbol{q}}$ is calculated by

$$R = \frac{\overline{\boldsymbol{p}}^T \tilde{\boldsymbol{q}}}{\|\overline{\boldsymbol{p}}\|} = \frac{\sqrt{\sum_{i=0}^{K} (\boldsymbol{u}_i^T \overline{\boldsymbol{p}})^2}}{\|\overline{\boldsymbol{p}}\|}. \tag{18}$$

Hence, we can obtain the normalized correlation $R$ approximately by Eq. (18), and the computational cost is reduced greatly than calculating directly with original template images. Moreover, Eq. (18) can be calculated by using FFT (Fast Fourier Transform) efficiently. Furthermore, we introduce a value $R_{\text{th}} = 0.7$ as a threshold value of $R$ to adjust the false detection rate. Therefore, the positions $\boldsymbol{x}_d$ $(d = 1, \cdots)$ where the normalized correlation values are larger than $R_{\text{th}}$ are searched, and they are decided as targets.

### 3.1.3 Target tracking

The detected targets are tracked by means of CAMSHIFT algorithm (Bradski, 1998). It is a nonparametric algorithm that investigates gradients of the hue values, obtained by HSV transformation, around the targets and tracks the peak points of these gradients. Moreover, this algorithm has robustness for the occlusion and the appearance changing of the object,

the computational cost is comparatively low because the target tracking regions are limited around ROI in last frame, and it is executable by the video rate.

First, the point $\boldsymbol{x}_d$ is set as an initial point of center of the target region that has been continuously and stably detected by the eigen space method mentioned above. When the plural objects are detected, the object nearest the robot, i.e., the object with the shortest distance from the center of the lower side of the image is set as a target. Next, the target is tracked by executing the CAMSHIFT algorithm. When losing sight of the target by pedestrians or moving of the robot equipped with the sensor system, the object is detected by using the eigen space method again. The experimental results of these procedures are shown in Fig. 7. The frame rates of these processing for $512 \times 512$ input image were about 5 [fps] for the eigen space method and about 20 [fps] for the CAMSHIFT algorithm by a computer equipped with a Pentium4 3GHz.

#### 3.1.4 Measurement of the target by stereo vision

This sensor system has a 3D measurement function based on a literature (Birchfield et al., 1999) on stereovision. Moreover, various parameters of the cameras in the sensor system were designed to meet the following requirements. Namely, the measurement is in error by less than 10 [mm] when the distance of the sensor system is within 2 [m] from the object. On the other hand, the measurement range of the LRF is from 60 [mm] to 4095 [mm], and its measurement is in error by 10 [mm] to 40 [mm]. Thus, the measurement by the stereovision is effective when the distance from the object is within 2 [m], and the LRF is effective for a longer distance. However, the measurement by the stereovision is influenced by environment light strongly, and the accurate 3D measurement is difficult by vibration while running of the robot. Therefore, in the following experiments, 3D measurements were executed by using LRF. From now on, an ingenious stabilization method of the measurement of stereovision will be needed to construct high accuracy and robust measurement method for autonomous robot under outdoor environment.

### 3.2 Measurement of target by LRF

A detailed three dimensional measurement is necessary for manipulation of the target by the robot. Here, the method of three dimensional measurement by the LRF without affected by environmental light is described. It is necessary to extract the data set of the target object from measurement data set for the analysis of the object. We employ PSF (Plane Segment Finder) method (Okada et al., 2001) or RHT (Randomized Hough Transform) (Xu et al. 1993, Ding et al. 2005) to extract and eliminate planes such as floor surface from the range data set. A concrete procedure is shown as follows.

First, let ${}^r\boldsymbol{x}_{ij} \triangleq ({}^rx_{ij}, {}^ry_{ij}, {}^rz_{ij})^T$ be 3D points measured the LRF and the indices of these points $i = 1, \cdots, I$ and $j = 1, \cdots, J$ represents the measurement order. The relations of data points are shown in Fig. 8(a), and Fig. 9 shows an example of measurement range data set of a plastic bottle. As the figure indicates, only the surrounding of the target is measured in high density by the rotating of the LRF. Next, to extract and remove the floor surface from this measurement data, the above mentioned PSF method is applied to this data according to the following procedures.

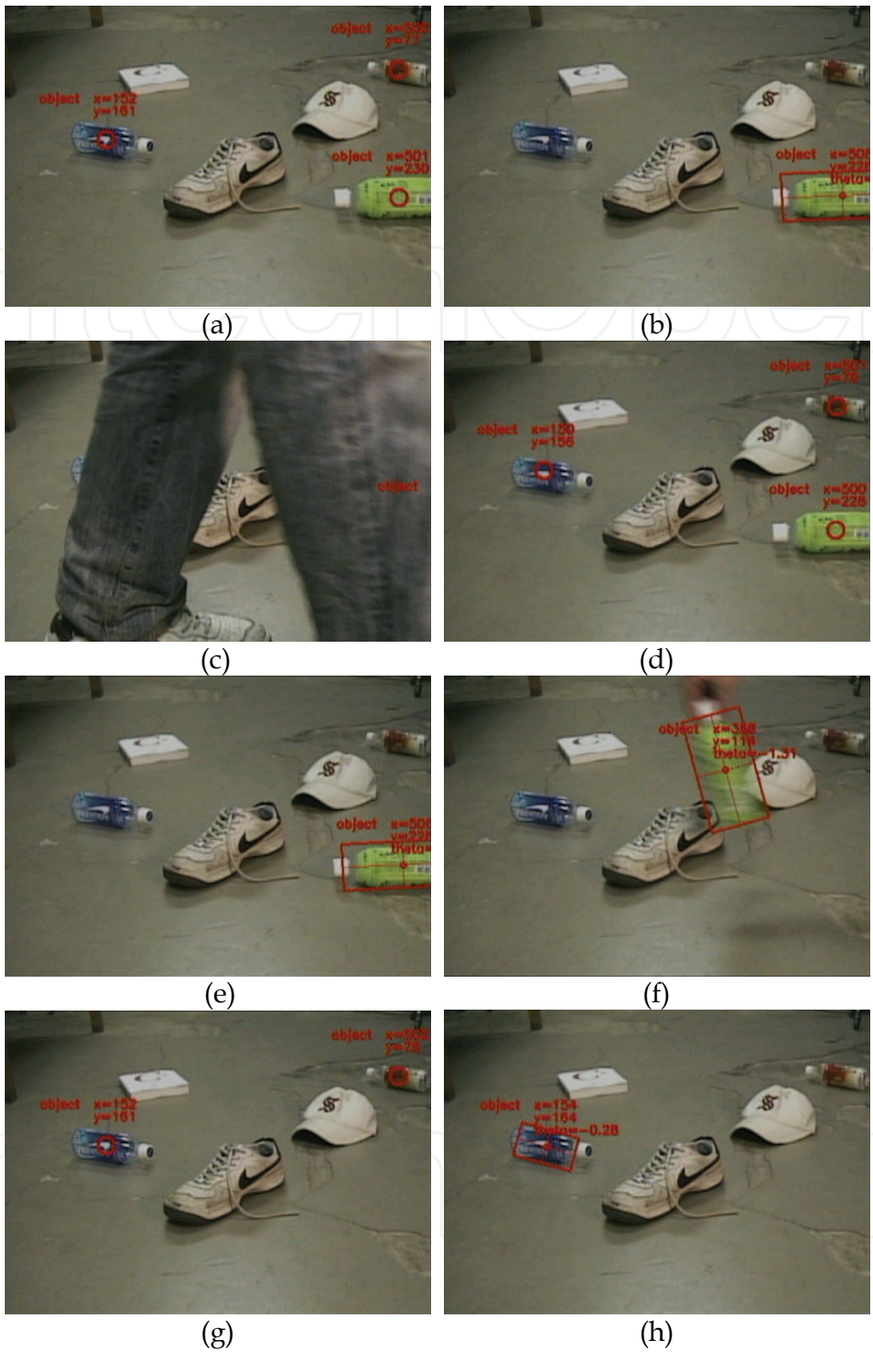



Fig. 7. Example of image processing: (a) the objects had been detected by the eigen space method; (b) the nearest object had been continuously and stably detected and it had been set as a target for tracking; (c) the target had been lost sight of by a pedestrian; (d) the objects were detected again; (e) the target was tracked by CAMSHIFT algorithm; (f) the target was correctly tracked though the object moved at high speed; (g) the target was excluded from the scene and the objects were detected again; (h) the target was tracked by CAMSHIFT algorithm.

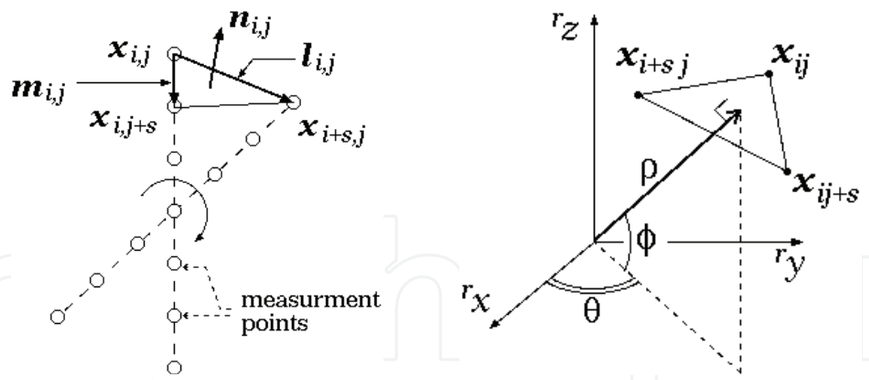



Fig. 8. Relations of 3D data points and a plane constructed by them: (a) relation of the measurement points; (b) relations of the coordinate system and Hough parameters.

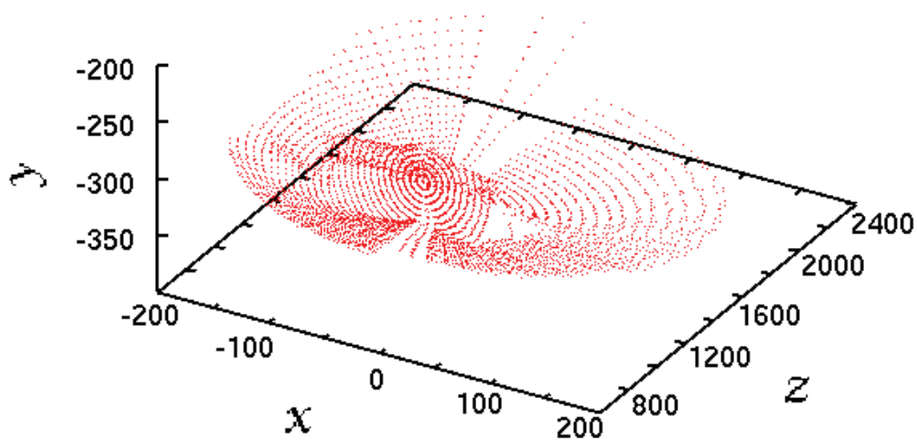



Fig. 9. Range data set of a detected plastic bottle.

Namely, the plane parameters $\theta_{ij}\,(0 \le \theta_{ij} < \pi)$ and $\phi_{ij}\,(0 \le \phi_{ij} < \pi)$ shown in Fig. 8(b) are calculated from the normal vector of the plane including the points $(x_{ij}, x_{i+s,j}, x_{i,j+s})$ as follows,

$$\theta_{ij} = -\tan^{-1}\left(n_{ij}^{y} / n_{ij}^{x}\right) \tag{19}$$

$$\phi_{ij} = -\tan^{-1}\left(n_{ij}^{z} \Big/ \left(n_{ij}^{x2} + n_{ij}^{y2}\right)\right)^{\frac{1}{2}} \tag{20}$$

where,

$$\boldsymbol{n}_{ij} = \frac{\boldsymbol{l}_{ij} \times \boldsymbol{m}_{ij}}{\left\| \boldsymbol{l}_{ij} \times \boldsymbol{m}_{ij} \right\|} \triangleq \left(n_{ij}^{x},\ n_{ij}^{y},\ n_{ij}^{z}\right)^{T}, \tag{21}$$

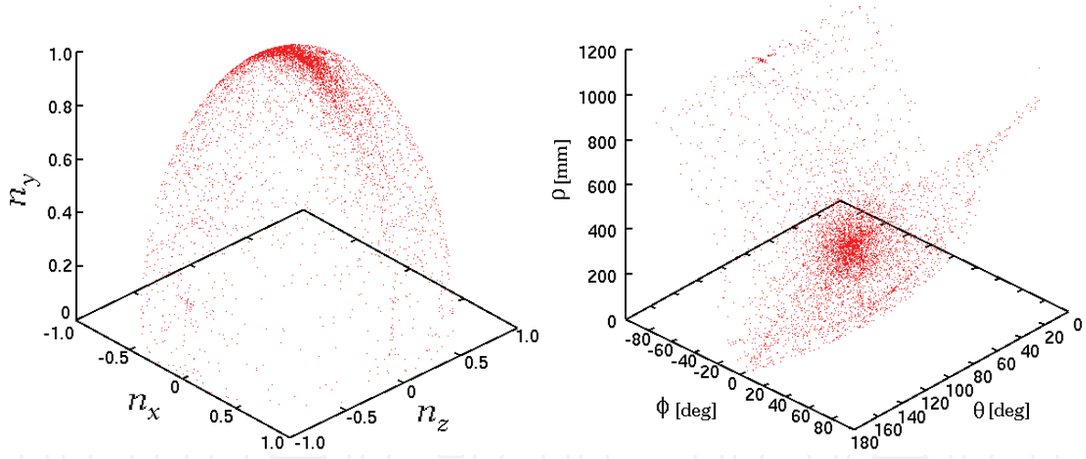
$$\boldsymbol{l}_{ij} = \boldsymbol{x}_{i+s,j} - \boldsymbol{x}_{ij}, \tag{22}$$

$$\boldsymbol{m}_{ij} = \boldsymbol{x}_{i,j+s} - \boldsymbol{x}_{ij}\,, \tag{23}$$

and $s$ represents an interval of index of data point. Here, we set $s = 2$ and calculated normal vectors $\boldsymbol{n}_{ij}$, and the distribution of them are shown in Fig. 10(a). We see from this figure that the normal vectors were distributed around $\boldsymbol{n}_{ij} = (0,\ 1.0,\ 0)^T$, however the specification of the peak of the distribution is difficult. Therefore, the values of $(\theta_{ij}, \phi_{ij}, \rho_{ij})$ in Hough space are calculated from these normal vectors. Namely, the values of $\rho_{ij}$ are obtained by the following expression.

$$\rho_{ij} = (x_{ij}\cos\theta_{ij} + y_{ij}\sin\theta_{ij})\cos\phi_{ij} + z_{ij}\sin\phi_{ij} \tag{24}$$

Since the variance of $\rho_{ij}$ caused by the measurement noise is relatively large, it is also difficult to specify the peak (Fig. 10(b)). Therefore, the values of $(\theta_{ij}, \phi_{ij})$ are normalized and mapped into $\theta - \phi$ space. Moreover two dimension histogram of 36×36 division is generated and the peak cell of $(\theta_{ij},\ \phi_{ij})$ is searched. In the histogram shown in Fig. 11(a), it is found that the peak value $C_{\theta_{pF},\phi_{pF}}$ exists at $\left(\theta_{pF},\ \phi_{pF}\right) = (90,0)$.




(a) (b)

Fig. 10. Distribution of vectors; (a) normal vectors; (b) $\left(\rho_{ij},\ \theta_{ij},\ \phi_{ij}\right)$ vectors.

Next, the following processing is executed for only the measurement points with the value within a certain range $\Delta_C$ (=50 [deg]) around the peak value. The Hough parameters of these measurement points are replaced with the peak value, i.e.,

$$\theta_{ij} := \theta_{pF} \quad \text{if } (\theta_{pF} - \Delta_C/2) \le \theta_{ij} \le (\theta_{pF} + \Delta_C/2), \tag{25}$$

$$\phi_{ij} := \phi_{pF} \quad \text{if } (\phi_{pF} - \Delta_C/2) \le \phi_{ij} \le (\phi_{pF} + \Delta_C/2), \tag{26}$$

and the following values are calculated.

$$\rho'_{ij} = (x_{ij}\cos\theta_{pF} - y_{ij}\sin\theta_{pF})\cos\phi_{pF} + z_{ij}\sin\phi_{pF}. \quad (27)$$

The measurement points with $\rho'_{ij}$ close to the peak value $\hat{\rho}^{\max}_{pF}$ can be estimated to be on the same plane. Namely, the measurement points on the floor surface can be excluded by eliminating these measurement points. The histogram of the values $\rho'_{ij}$ is shown in Fig. 11(b). We employed the discriminant analysis method to separate the histogram, and then it was separated at 300 [mm]. The separated points with $\rho'_{ij} \geq 300[\mathrm{mm}]$, i.e. the measurement points on the floor surface are shown in Fig. 12(a), and other measurement points are shown in Fig. 12(b). It was found from these results that planes such as floor surface can be excluded from the target data set by this procedure. As a consequence, the robot can know size, position, and posture of the target object by the extracted range data.

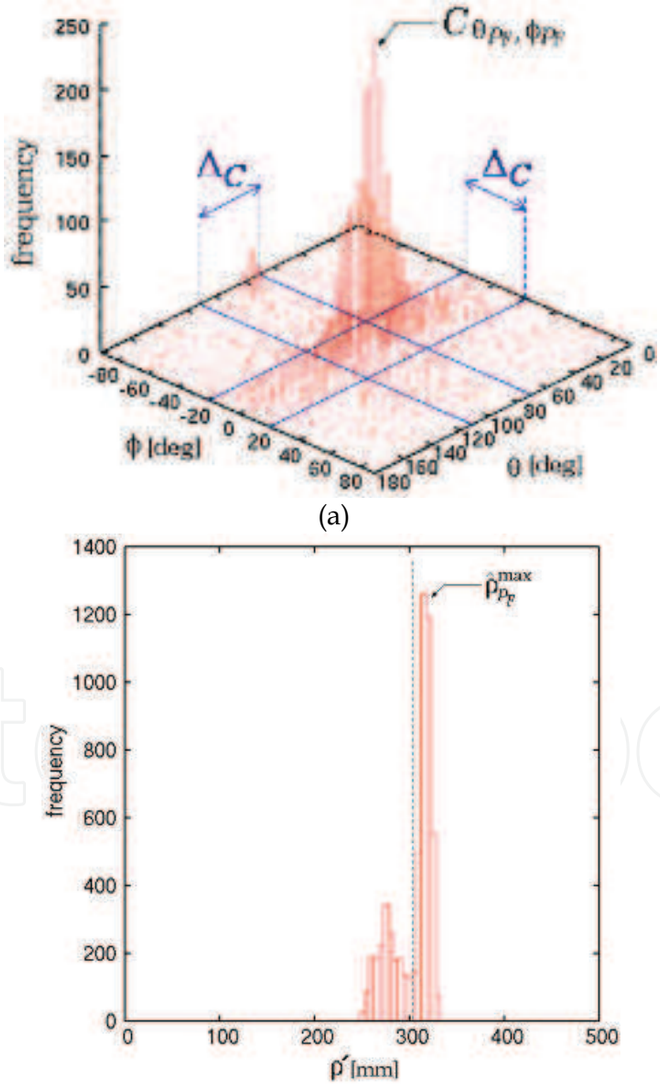
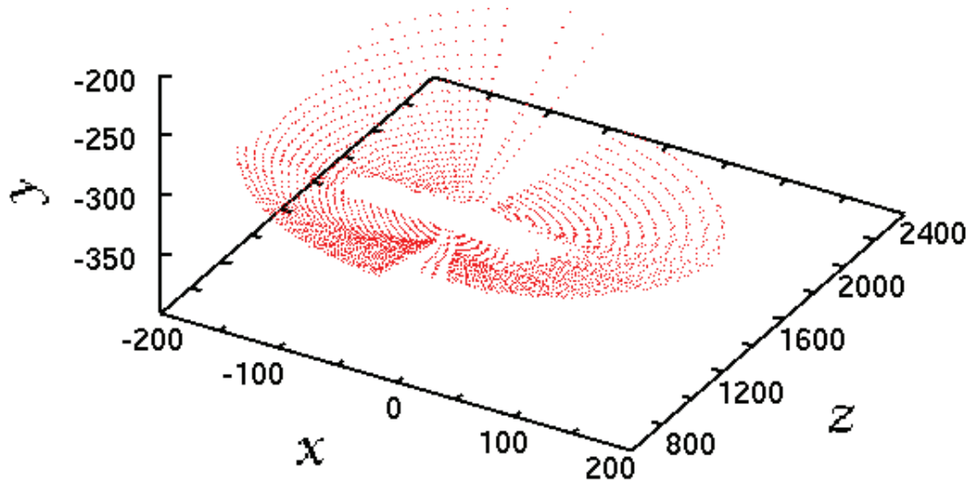


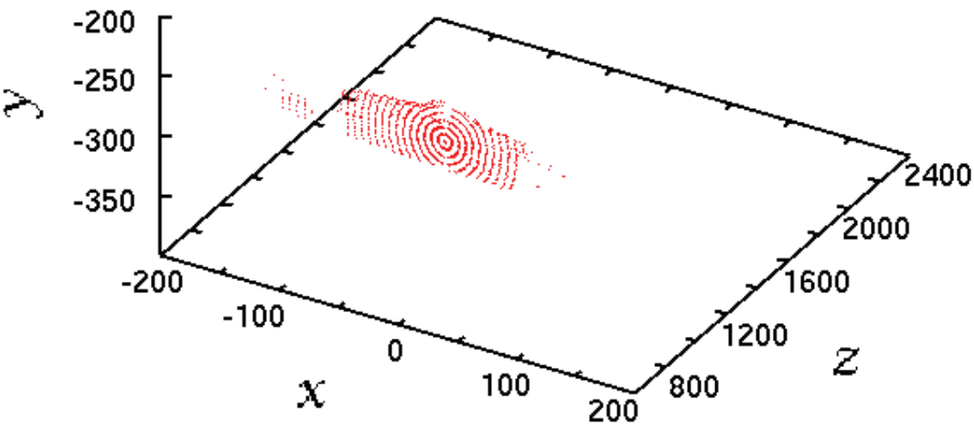

Fig. 11. Histograms: (a) normalized Hough parameter $(\theta_{ij}, \phi_{ij})$; (b) values of $\rho'_{ij}$.

### 3.3 Simple measurement by one camera and LRF

This sensor system can simply track and measure the distance of the object by using only a camera and the LRF. Namely, the object is detected and tracked by using only one camera image, and the posture of the sensor system is controlled to overlap the center of the object and the center of the image. At this time, since the pixels of the camera image corresponding to the measurement points of the LRF can be easily found from Eqn. (8), (9), (10), the distance to the object can be known by using a part of the measurement result of the LRF. An example of processing image of this application is shown in Fig. 13. In the rectangular at the upper right of the image, the distance to the object (D. to Object), the pixel position of center of the object (C. of pix.), the posture of appearance of the object (Theta), the size of the object (Area), and the posture angle of the sensor system (Yaw, Pitch, Roll) are displayed. Moreover, it was confirmed that these information about the target object was calculated in the video rate (30 [fps]).




(a)




(b)

Fig. 12. Example of measurement of a plastic bottle: (a) results of extraction of the floor surface; (b) range data set of the plastic bottle.

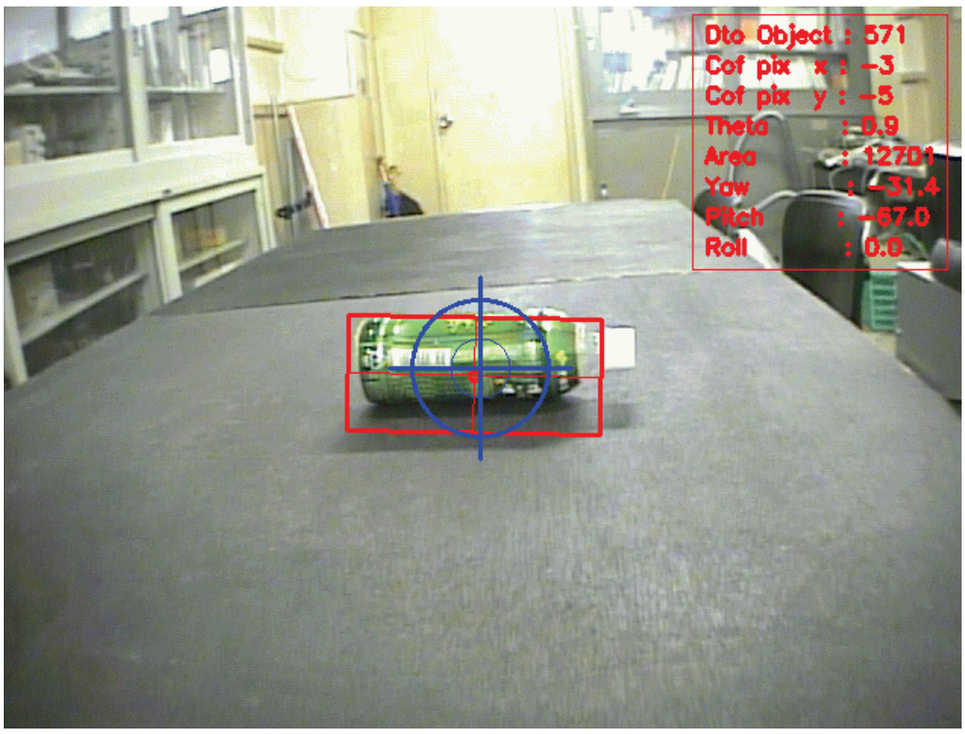


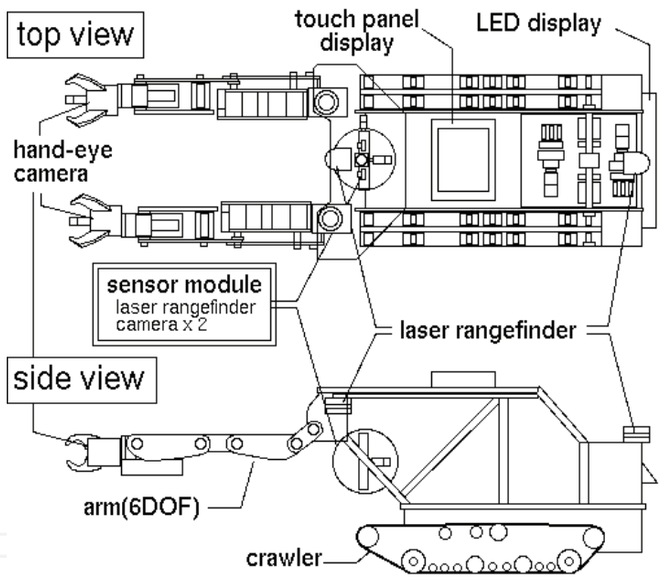

Fig. 13. Example of the detection and the measurement of a plastic bottle. The scope mark shows the center of the image, and the square around the plastic bottle shows a detected area of the object.

## 4. Installation for outdoor service robot OSR-02

Since November 28th of 2003, the Japanese government has designated the Fukuoka Prefecture as the one of the special zones for robot research and development. There researchers are allowed to operate their robots in public areas by applying for police permission. Thus, for example, in shopping areas in the city, we can take advantages of executing realistic experiments of outdoor service robots for cleaning up streets, welcoming customers, guiding them to a certain place, providing information about the shopping area, and so on. Under such circumstances, we have started developing the outdoor service robots called OSR-01 (Obata et al., 2006) and OSR-02 (Nishida et al., 2006) intended for cleaning up shopping streets by means of the separated collection of discarded trash, such as plastic, glass, and steel, etc. on shopping streets.

The outdoor service robot which we call OSR-02 (Fig. 14) had been developed intended for cleaning up urban areas by means of collecting discarded trash such as plastic bottles, cans, plastic bags and so on. So far, several type robots for the cleaning work have been developed; however, they only vacuum on the specified route. Since OSRs recognize and collect trash or target objects with the manipulator in shopping street etc., the technology that composes the OSRs is greatly different. Although many operations have been programmed for the above task, we, in this section, mainly show the sensor system and the manipulation system of OSR-02 for trash detection and distance measurement of the target after giving a briefing about the hardware of the robot.

(a)




(b)

Fig. 14. Outdoor service robot OSR-02: (a) whole image of OSR-02; (b) configuration.

### 4.1 Structure

OSR-02 has two manipulators and crawlers. The length of the manipulators is 850[mm] and they have five degrees of freedom and a hand. The specification of OSR-02 is shown in Table 2 and its electrical structure is shown in Fig. 15. Moreover, five computers for control of the crawler, the manipulators, and the sensor system are installed in OSR-02, and they communicate via Ethernet. Furthermore, in order to recognize the surrounding environment, and not to harm surrounding persons, several sensors had been installed in the robot.

| | | |
|---|---|---|
| D.O.F. of manipulator | | 5 (x2) |
| hand | | 1 (x2) |
| drive wheel | | 2 |
| dimension | height | 600[mm] |
| | width | 500[mm] |
| | depth | 800[mm] |
| manipulator length | | 850[mm] |
| weight | | 90[kg] |
| running speed | | 2.0 [km/h] |
| sensor | LRF | 3 |
| | color CCD camera | 4 |
| | force sensor (hand) | 2 |
| | magnetic sensor (hand) | 2 |

Table 2. Specification of OSR-02

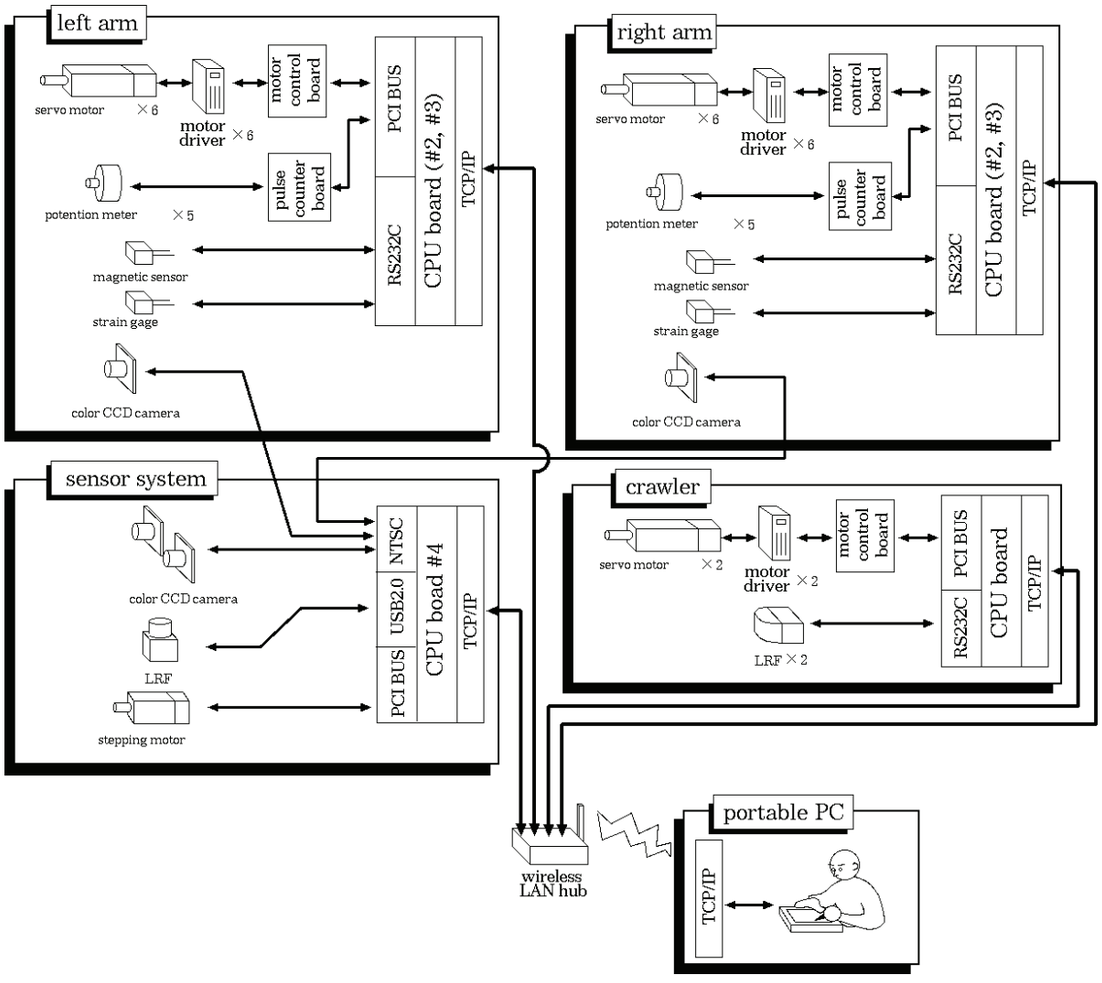
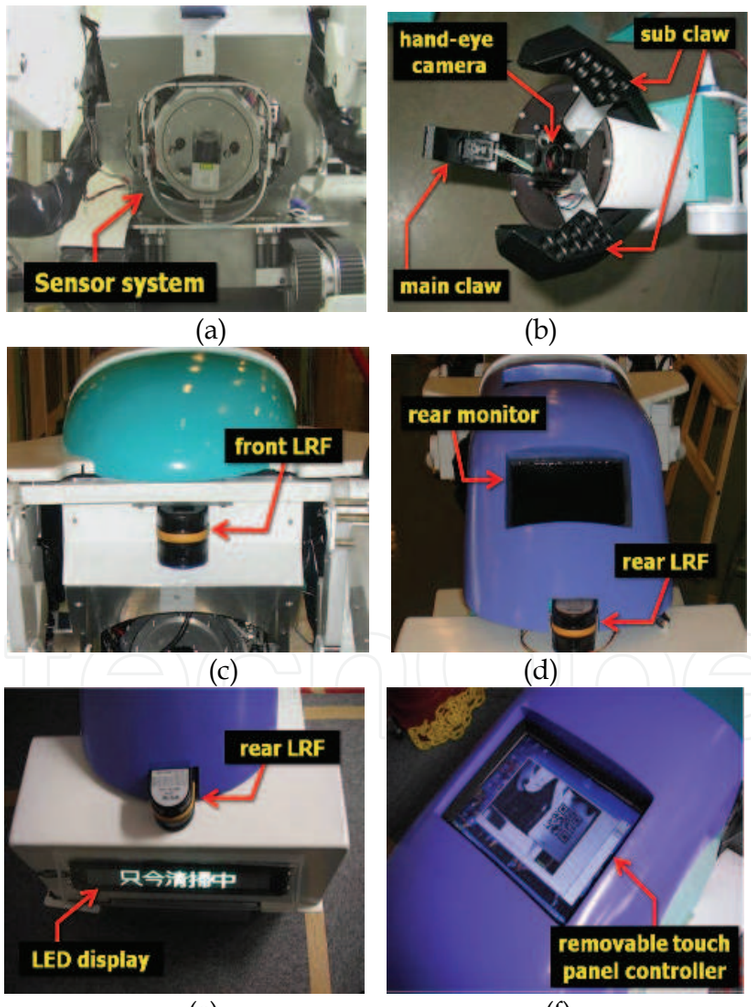



Fig. 15. Electrical structure of OSR-02

### 4.2 Equipped devices

The main equipped devices of OSR-02 are shown in Fig.16. Fig. 16(a) shows the installed sensor system, and it is used for the running route inspection and the trash detection. The right hand is shown in Fig. 16(b), and a hand-eye camera is installed in the center position of claws. The image obtained from the camera is used for the hand-eye system driven for the grasping of target. Moreover, a magnetic sensor and a strain gage have been installed into a claw of each hand as shown in Fig.17. The sorted collection of the grasping objects and the damage prevention of the hands by controlling of the grasping power are possible by using these sensors. The rear and the front LRF for the obstacle detection of the robot are shown in Fig. 16(c) and (d), and the horizontal scanning range of each LRF is 180 [deg] by 10 [Hz]. The results of the image processing in the robot can be displayed on the rear monitor shown in Fig. 16(d). The publicity LED panel shown in Fig. 16(e) displays the content of the job of the robot, the advertisement of the shopping street, and so on. It is possible to give the command to OSR-02 by the touch panel of removable laptop PC via wireless LAN (Fig. 16(f)).




(a) (b) (c) (d) (e) (f)

Fig. 16. Equipped devices of OSR-02: (a) installed sensor system; (b) right hand; (c) front LRF; (d) rear monitor and rear LRF; (e) LED display; (f) removable touch panel controller.

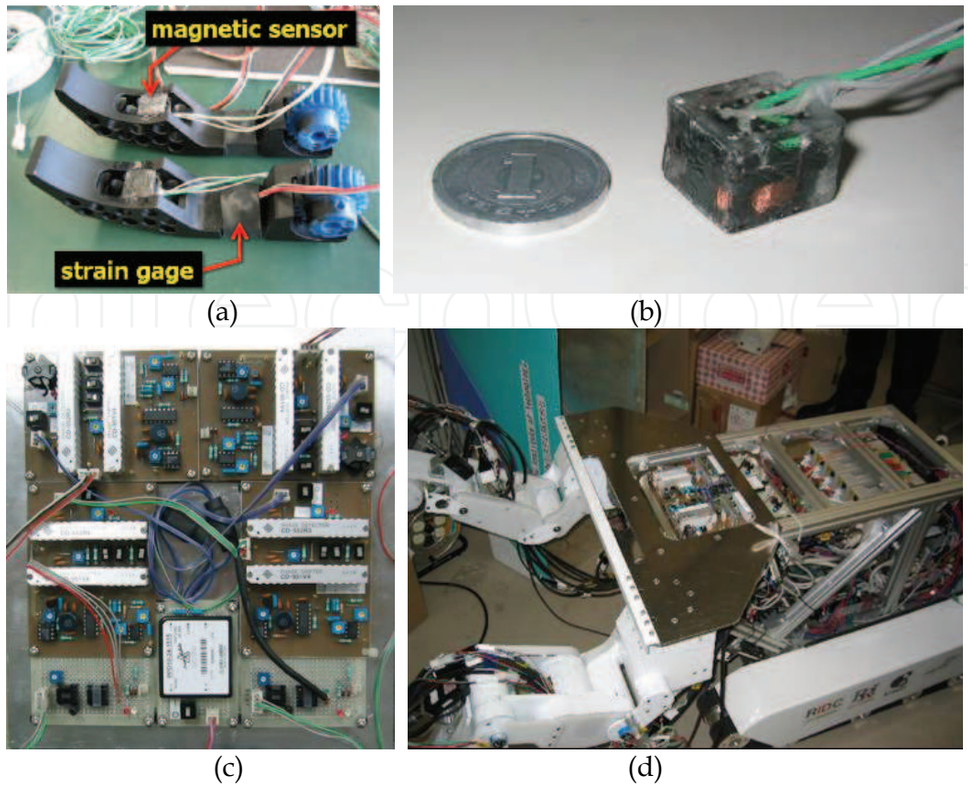



Fig. 17. Magnetic sensor for trash separation: (a) main claws of each hand with magnetic sensor and strain gage; (b) magnetic sensor; (c) sensor circuit; (d) the circuit built into the robot.

### 4.3 Posture control of sensor system

We constructed a visual feedback running control system by using the sensor system for OSR-02 to approach the target object. It is necessary to control the posture of the sensor system so as not to lose sight of the target until the robot approaches it. When the target is detected at $(x_d(t),\ y_d(t))^T \in \sum_{sl}$ on the left camera screen at time $t$, the yawing angle $\theta(t)$ and the pitching angle $\phi(t)$ of the sensor system are controlled as follows,

$$\begin{pmatrix} \dot{\theta}(t) \\ \dot{\phi}(t) \end{pmatrix} = \begin{pmatrix} -k_{01} & 0 \\ 0 & -k_{11} \end{pmatrix} \begin{pmatrix} \theta(t) \\ \phi(t) \end{pmatrix} + \begin{pmatrix} -k_{02} & 0 \\ 0 & k_{12} \end{pmatrix} \begin{pmatrix} x_d(t) \\ y_d(t) \end{pmatrix}, \tag{28}$$

where, gain coefficients $k_{01}; k_{02}; k_{11}; k_{12} > 0$ are design parameters, and the bias between the rotation center of the sensor system and the optical center of the cameras are disregarded.

### 4.4 Running control system

The relation of the coordinate systems on the robot and the target position is shown in Fig. 18(a). The angular velocity input $\boldsymbol{w}(t) \triangleq (w_l(t),\ w_r(t))^T$ for approaching the target is controlled as follows,

$$\boldsymbol{w} = \boldsymbol{R}_g \boldsymbol{v}, \tag{29}$$

where,

$$\boldsymbol{R}_g = \frac{1}{r_w}\begin{pmatrix} 1 & l \\ 1 & -l \end{pmatrix}, \quad \boldsymbol{v} = \begin{pmatrix} v_g \\ w_g \end{pmatrix}, \tag{30}$$

$r_w$ is radius of drive wheel in the crawler, and $l$ is distance between the drive wheel and the center of rotation of the robot. Moreover, $v_g$ and $w_g$ are the velocity and the angular velocity of the center of rotation of the robot, respectively. The velocity input in the robot coordinate system toward the target position is decided as follows,

$$\boldsymbol{v} = \boldsymbol{K} \cdot {}^r\boldsymbol{d}, \tag{31}$$

where ${}^r\boldsymbol{d} \triangleq ({}^r d_x, {}^r d_z)^T$ is the position of the target object in which height is disregarded. Moreover, $\boldsymbol{K}$ is represented as follows,

$$\boldsymbol{K} = \begin{pmatrix} k_r & 0 \\ 0 & k_\theta \end{pmatrix}, \tag{32}$$

where, $k_r$ and $k_\theta$ are gain coefficients about distance from the target object and about yawing angle $\theta$ of the sensor system, respectively. The values of $k_r$ and $k_\theta$ are adjusted according to the situation of the road. Therefore, the velocity of each wheel are determined by Eqn. (29) and Eqn. (31) as follows,

$$\boldsymbol{w} = \boldsymbol{R}_g \boldsymbol{K}\, {}^r\boldsymbol{d} \tag{33}$$

where the value of $\boldsymbol{w}$ does not exceed the maximum velocity set beforehand. A picture of the rear monitor of the robot at the time of running by the above visual feedback control system is shown in Fig. 18(b).

### 4.5 Target manipulation by hand-eye system

After the robot approached to a constant distance from the target, it raises the arm and takes up an initial posture to grasp the target (Fig. 19(a)). When the manipulator is in the initial posture, the robot views the target by the hand-eye camera and measures grasping position and posture of it. The relation between the hand-eye camera and the target object is shown in Fig. 19(b). The hue value obtained by the sensor system in a prior running sequence is used to identify the target region on the right hand-eye screen. Namely, the region with the specific hue value tracked in a prior sequence is searched on the hand-eye camera image, and only the region with the maximum size is set as the target region. Then the minimum square that suits this region, the center position $\boldsymbol{x}_e = (x_e, y_e)^T \in \Sigma_{hsr}$ and the posture angle $\theta_h$ of the target object are calculated. An example of the result of right hand-eye image processing is shown in Fig. 19(c). Here, the height $h_h$ of the hand in the initial posture is always constant, and the height $h_t$ of the target object can be estimated from the kind of

trash (e.g. $h_h = 65[\mathrm{mm}]$ for a plastic bottle). Therefore, the grasping position $\boldsymbol{x}_t \in \Sigma_{hr}$ of target can be calculated as follows,

$$\boldsymbol{x}_t \triangleq \begin{pmatrix} x_t \\ y_t \\ z_t \end{pmatrix} = \begin{pmatrix} -\alpha x_e z_t / f_h \\ \alpha y_e z_t / f_h \\ h_h - h_t \end{pmatrix}, \tag{34}$$

where $f_h$ is focal length of the hand-eye camera and $\alpha$ is a coefficient to convert the unit from the pixel into the distance. Here, the hands had been designed so that the right hand coordinate system $\sum_{hr}$ overlaps with the hand-eye coordinate system $\sum_{hsr}$ as shown in Fig. 19(c), namely, rotation and bias of them are able to be disregarded. Next, the hand and claws are controlled for grasping and collection of the target object by a feed forward control system, therefore the joint angles of the right arm $\theta_{hrj}\ (j = 1, \cdots, 5)$ shown in Fig. 19(d) are controlled to be $\boldsymbol{x}_t \to \boldsymbol{0}$ and $\theta_{hr5} \to \theta_h$. Moreover, appropriate holding of the target is confirmed by the strain gauge, and the above-mentioned sequence is repeated again when the robot failed in the holding.

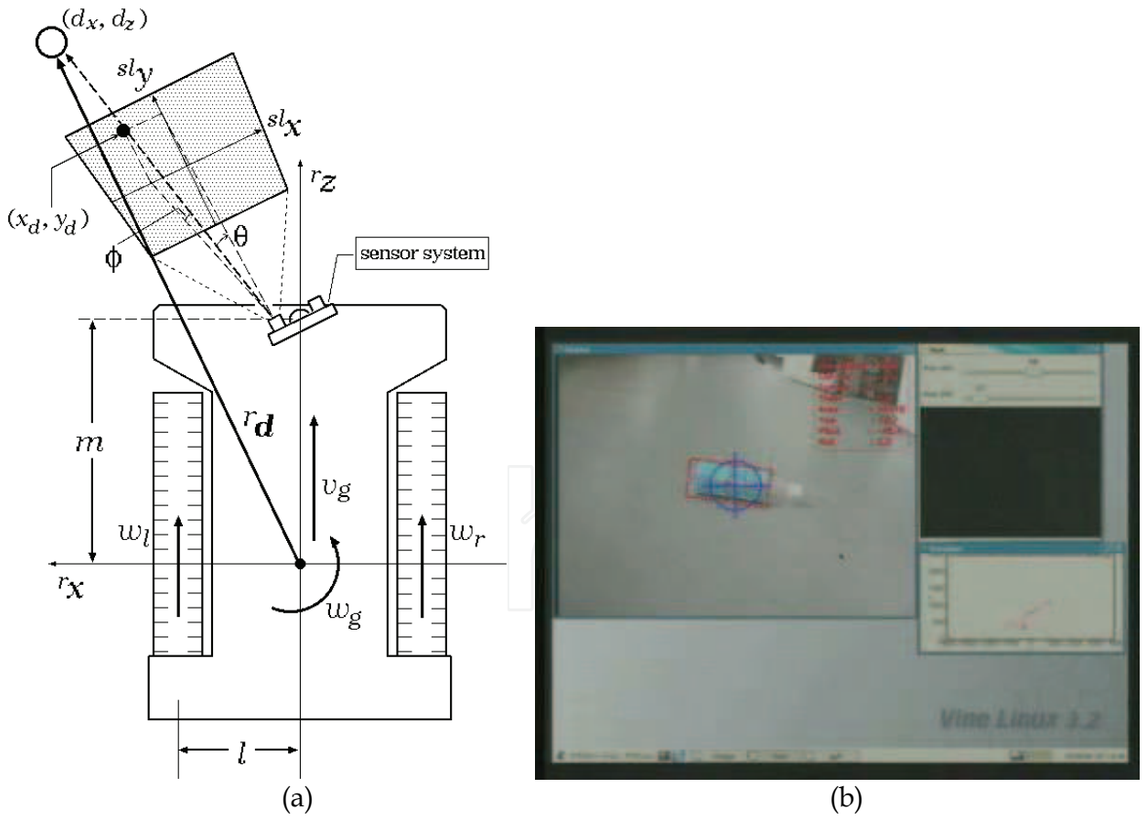



Fig. 18. Running control system of OSR-02: (a) relation of the target position and the robot coordinate system; (b) a picture of the rear monitor during approaching to the target object, upper left window is a left camera image, upper right window is a right hand-eye camera image, and lower right window is a measurement data of the LRF.

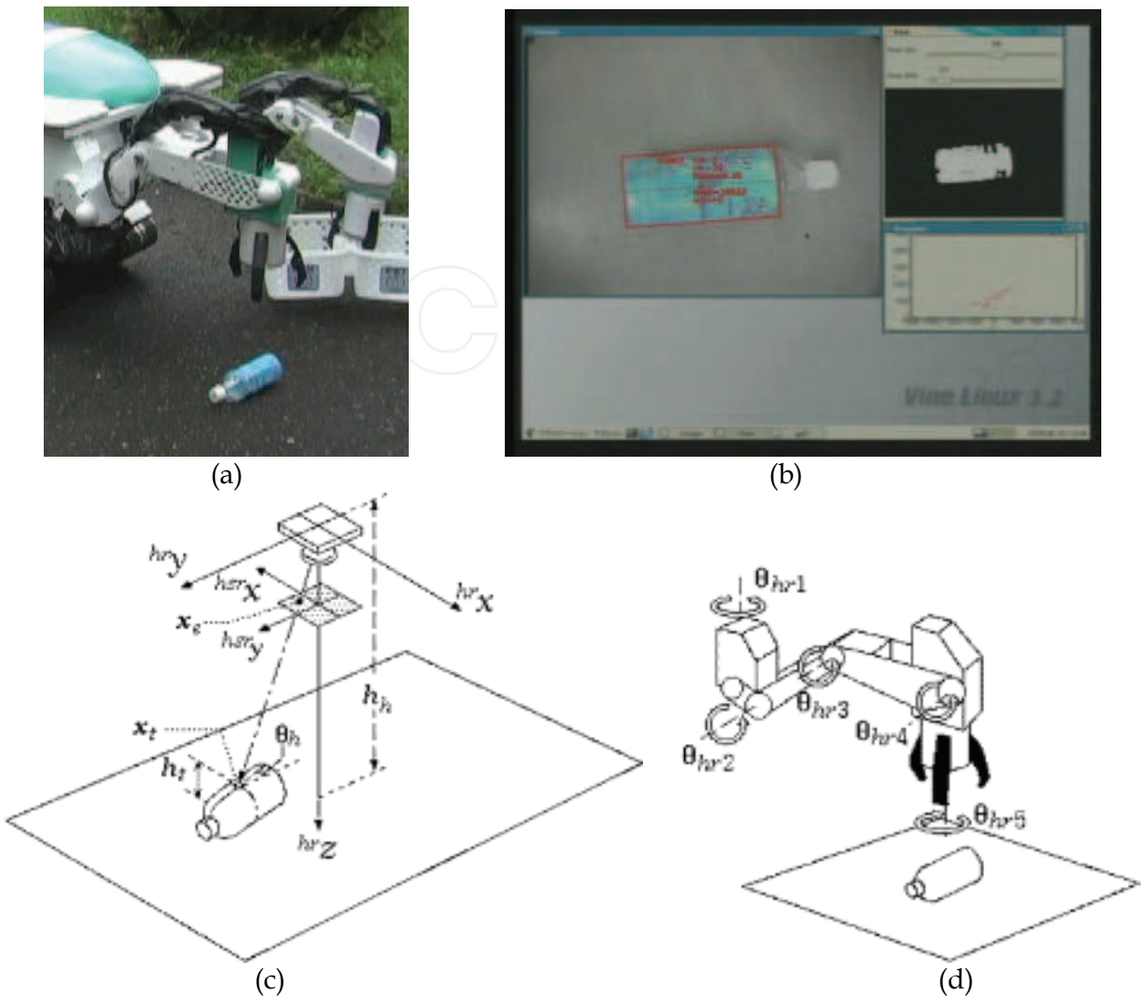



Fig. 19. Hand-eye system: (a) initial posture for grasping the target; (b) a picture of the rear monitor during the grasping action, the information about the target is displayed by the red lines in the upper left window, and the region with the target hue is shown in the upper right window; (c) relation of the hand-eye camera screen coordinate system $\Sigma_{hsr}$ and the hand coordinate system $\Sigma_{hr}$; (d) joint angles of the right arm.

## 5 Experiments

### 5.1 Trash collection task

Trash collection task of OSR-02 was experimented under outdoor environments. Plastic bottles and cans were put forward of about 2 [m] of the robot, and the collection experiments were conducted. In this experiment, the robot had been programmed to classify the target by the right hand and to store it into the trash box for the separate collection held by the left hand. The weather on the day was a drizzle and the results of the experiment are shown in Fig. 20. It was confirmed that the above mentioned various measurement methods, the control systems, and the operation sequence of them are able to be behaved appropriately by these experiments.

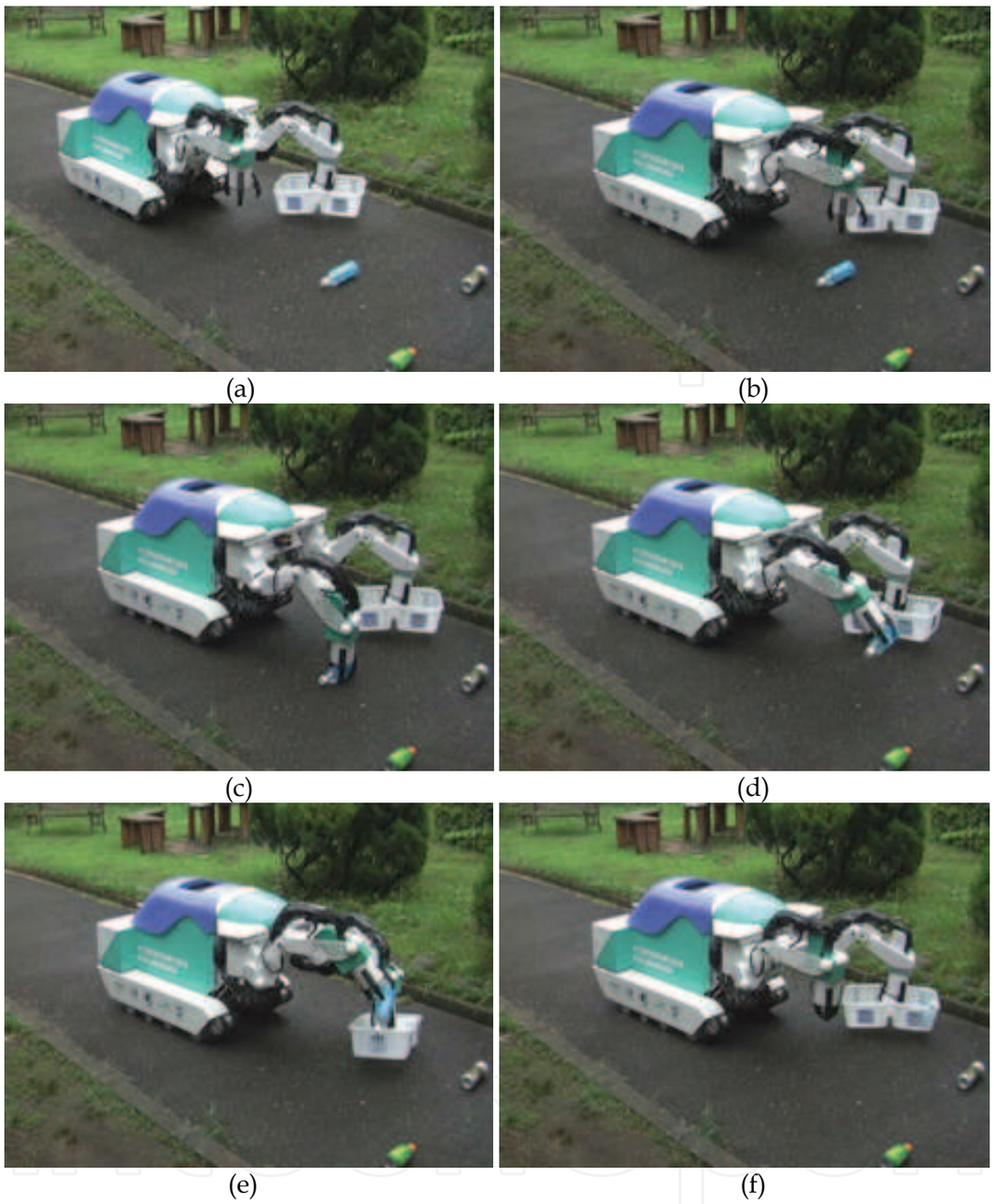


Fig. 20. Trash collection work flow of OSR-02: (a) a plastic bottle had been detected; (b) initial posture for grasping; (c) grasping the target; (d) lift operation; (e) the robot puts the grasped object into the trash box for plastic based on the classification result; (f) searching for next trash.

### 5.2 Experiments in public space

We demonstrated of OSR-02 several times at a shopping arcade and a playground in Kitakyushu-city (Fig.21). In the demonstrations, the robot went round the shopping arcade and the playground, and collected trash. We confirmed that OSR-02 was able to detect the

plastic bottle, glass bottle, and cans on the street, to approach, to collect accurately, and to avoid the pedestrian. Furthermore, it was found from these experiments that the calculation speed of the developed vision system was fast enough for the collection of trash.


Fig. 21. Demonstrations of OSR-02: (a) at a shopping street; (b) at a playground.

## 6. Conclusion

In this research, we have proposed a novel sensor system and its applications, and have shown its effectiveness by the fundamental experiments. Moreover, we installed the sensor system into the outdoor service robot OSR-02 and conducted experiments at the shopping street and the playground. Then, it was found from these experiments that the calculation speed of the developed vision system was fast enough for appropriate running and the collection of trash.

In the design of the autonomous robot, the relation between the electricity consumption and the performance is always considered, and the required measurement accuracy differs by tasks. For example, the measurement should be high speed while the robot is moving, and it should be high accuracy while the robot stops and is manipulating the object. This sensor system can achieve such demand of the measurement according to kind of task by switching the measurement method and devices though it is composed of basic sensor devices.

## 7. Acknowledgements

This work was supported in part by Robotics Industry Development Council (RIDC). We would like to thank Takashi Kondo, Yasuhiro Fuchikawa and Shuich Kurogi of Kyushu Institute of Technology; Shuji Itoh and Norio Hiratsuka of YASKAWA INFORMATION SYSTEMS Co.; Yasuhiro Watanabe and Fumitaka Koga of Mechanics & Electronics Research Institute Fukuoka Industrial Technology Center; Toshinori Suehiro of Fukuoka Industry, Science & Technology Foundation; Yoshinori Kawamura of YASKAWA ELECTRONICS; and Yoshimitsu Kihara of Kihara Iron Works.

## 8. References

Birchfield, S. & Tomasi, C. (1999). Depth Discontinuities by Pixel-to-Pixel Stereo, *International Journal of Computer Vision*, Vol. 35, No. 3, pp. 269-293.

Bradski, R., G., (1998). Computer vision face tracking as a component of a perceptual user interface, *In Workshop on Applications of Computer Vision*, pp. 214–219.

DeSouza, N., G. & Kak, C. (2002). A. Vision for Mobile Robot Navigation: A Survey, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 24, No.2, pp. 237–267.

Ding, Y.; Ping, X.; Hu, M. & Wang, D. (2005), Range Image Segmentation Based on Randomized Hough Transform, *Pattern Recognition Letters*, No. 26, pp. 2033-2041.

Fuchikawa, Y.; Nishida, T.; Kurogi, S.; Kondo, T.; Ohkawa, F.; Suehiro, T.; Watanabe, Y.; Kawamura, Y.; Obata, M.; Miyagawa, H. & Kihara, Y. (2005). Development of a Vision System for an Outdoor Service Robot to Collect Trash on Streets, *Proc. of CGIM05*, pp.100-105.

Hihnel, D.; Burgard, W. & Thrun, S. (2003). Learning compact 3D models of indoor and outdoor environments with a mobile robot, *Robotics and Autonomous Systems*, Vol. 44, pp. 15–27.

Kondo, T.; Nishida, T.; Obata, M. & Ohkawa, F. (2005). A Research Report about The Outdoor Service Robot OSR-01, *Proc. of The 1st International Conference on Design Engineering and Science Vienna*, pp. 271-275.

Xu, L. & Oja, E. (1993). Randomized HoughTransform (RHT): Basic Mechanisms, Algorithms and Complexities, *Computer Vision, Graphics, and Image Processing: Image Understanding*, Vol.57, No.2, pp. 131-154.

Nishida, T.; Takemura, Y.; Fuchikawa, Y.; Kurogi, S.; Ito, S.; Obata, M.; Hiratsuka, N.; Miyagawa, H.; Watanabe, Y.; Koga, F.; Suehiro, T.; Kawamura, Y.; Kihara, Y.; Kondo, T. & Ohkawa, F. (2006). Development of Outdoor Service Robots, *Proc. of SICE-ICASE International Joint Conference*, pp. 2052-2057.

Nishida, T.; Takemura, Y.; Fuchikawa, Y.; Kurogi, S.; Ito, S.; Obata, M.; Hiratsuka, N.; Miyagawa, H.; Watanabe, F.; Suehiro, T.; Kawamura, Y. & Ohkawa, F. (2006). Development of Outdoor Service Robots, *Proc. of SICE-ICASE International Joint Conference*, pp. 2687-2691.

Obata, M.; Nishida, T.; Miyagawa, H.; Kondo, T. & Ohkawa, F. (2006). Development of Outdoor Service Robot to Collect Trash on Streets, *IEEJ Trans. EIS*, Vol. 126, No. 7, pp.840-848.

Okada, K.; Kagami, S.; Inaba, M. & Inoue, H. (2001). Plane Segment Finder: Algorithm, Implementation and Applications, *Proc. of Int. Conf. on Robotics and Automation*, pp. 2120-2125.

Surmann, H.; Lingemann, K.; Nuchter, A. & Hertzberg, J. (2001). A 3D laser range finder for autonomous mobile robots, *Proc. the 32nd ISR*, pp. 153–158.

Surmann, H.; Nuchter, A. & Hertzberg, J. (2003). An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor environments, *Robotics and Autonomous Systems*, Vol. 45, pp. 181–198.

Uenohara, M. & Kanade, T. (1997). Use of Fourier and Karhunen-Loeve Decomposition for Fast Pattern Matching With a Large Set of Templates, *IEEE trans. on pattern analysis and machine intelligence*, Vol. 19, No. 8, pp. 891–898.


**Advances in Service Robotics**

Edited by Ho Seok Ahn

ISBN 978-953-7619-02-2
Hard cover, 342 pages
**Publisher** InTech
**Published online** 01, July, 2008
**Published in print edition** July, 2008

This book consists of 18 chapters about current research results of service robots. Topics covered include various kinds of service robots, development environments, architectures of service robots, Human-Robot Interaction, networks of service robots and basic researches such as SLAM, sensor network, etc. This book has some examples of the research activities on Service Robotics going on around the globe, but many chapters in this book concern advanced research on this area and cover interesting topics. Therefore I hope that all who read this book will find lots of helpful information and be interested in Service Robotics. I am really appreciative of all authors who have invested a great deal of time to write such interesting and high quality chapters.



**InTech Europe**
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

**InTech China**
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen