Advanced Robotics, Vol. 16, No. 8, pp. 701–719 (2002)

Also available online - www.vsppub.com

Full paper

On-line extraction of stable visual landmarks for a mobile robot with stereo vision

INHYUK MOON [1,*], JUN MIURA [2] and YOSHIAKI SHIRAI [2]

[1] *Department of Medical Engineering, Yonsei University, Wonju, Kangwon 220-071, Korea*
[2] *Department of Computer-Controlled Mechanical Systems, Osaka University, Suita, Osaka 565-0871, Japan*

Received 18 October 2001; accepted 27 February 2002

Abstract—This paper proposes a method to extract on-line stable visual landmarks from sensory data obtained by stereo vision so as to adapt to changes of environment. Given a two-dimensional obstacle map, the robot first extracts vertical line segments which are distinct and inside planar surfaces not near boundary edges as they are expected to be observed reliably from various viewpoints. However, the extracted feature information such as position and length include uncertainty due to errors of vision and motion. The robot then reduces the uncertainty by matching the planar surface containing the features to the map. These processes are performed on-line in order to adapt to actual changes of lighting and the scene depending on the robot's view. Experimental results in real scenes show the validity of the proposed method.

Keywords: Stable visual landmark; on-line extraction; uncertainty; mobile robot; stereo vision.

1. INTRODUCTION

Localization is one of the fundamental functions for a mobile robot [1]. Since the uncertainty in localization only by dead reckoning monotonically increases as a mobile robot moves, the robot should sense external information for localization. Landmark-based localization methods are often used because they have less cost for observing external information. Artificial landmarks such as sign patterns [2] have often been used for performing reliable observation. However, we do not want to arrange such artificial landmarks in the environment. Boundary edges of known obstacles are also used for localization [3], but these may be *unstable* for visual observation from various viewpoints.

Some research uses a learned feature map for navigation. In the learning phase, the robot first selects natural features such as points and lines [4] or regions of

*To whom correspondence should be addressed. E-mail: ihmoon@dragon.yonsei.ac.kr

high edge density [5] from the observed images, and registers them in a map. In the execution phase, the robot's state, i.e. position and orientation, is estimated by matching the observed features to the map. However, this method needs two phases (leaning and execution), and the navigation may be unstable when the robot's viewpoints and the lighting condition are changed from the learning phase.

Little *et al.* [6] proposed a method for selecting point features on vertical planar surfaces as stable visual landmarks. The basic idea is that the scene structure around the features is not changed although the robot's viewpoints are changed. However, the point is known as a weak feature for the visual observation, because the appearance of such features is easily affected by noise. Moreover, the study did not show how to estimate the feature positions and how to register them in the map as landmarks.

There is a research that uses planar surfaces as landmarks in order to achieve reliable observation [7]. The robot selects a planar surface from given a CAD-based three-dimensional (3D) map, and it observes the planar surface using a two-dimensional (2D) laser range sensor. However, it is inefficient to extract the planar surface because such a range sensor has high sensing costs for scanning 3D data. Recently, localization methods that do not use explicit landmarks were proposed [8]. The robot gathers 2D range data around its motion environment using a range sensor and then it matches the data to an occupancy grid map for localization. However, when it tries to detect obstacles lower than the position of the range sensor or sense an object with slender legs such as a table or a chair, the robot may be unable to localize due to the fixed height of the range sensor.

In this paper, we propose a method to extract *stable visual features* and to register them as *stable visual landmarks* that are reliably detectable under changes of lighting and robot's viewpoint. We deal with the case that the robot has an environment map to describe 2D obstacle positions except for feature information to be used as landmarks. Since inner features of a planar surface are less affected by changes of lighting and robot's viewpoints as compared with the boundary edges [6], we extract *vertical line segments* which are inside the planar surface as the stable visual landmarks.

The robot behavior is simple. When the robot moves to a destination, it performs the *landmark-based non-stop navigation* [9] using the extracted stable visual landmarks. If the robot has no available landmarks for the landmark-based navigation, it stops its motion and extracts new stable vertical line segments as landmarks. As a result, the robot can achieve more efficient visual navigation because the extraction process requiring high computational costs is performed selectively in the whole of the navigation task.

The extraction of stable visual landmarks is performed as follows (see also Fig. 1). The robot first selects a *reference plane* from a given map. Then, the robot controls the viewing direction of the stereo camera mounted on a rotational table to the reference plane and inputs stereo images. We apply two types of stereo matching, i.e. area- and segment-based methods, to the input stereo images. In order to

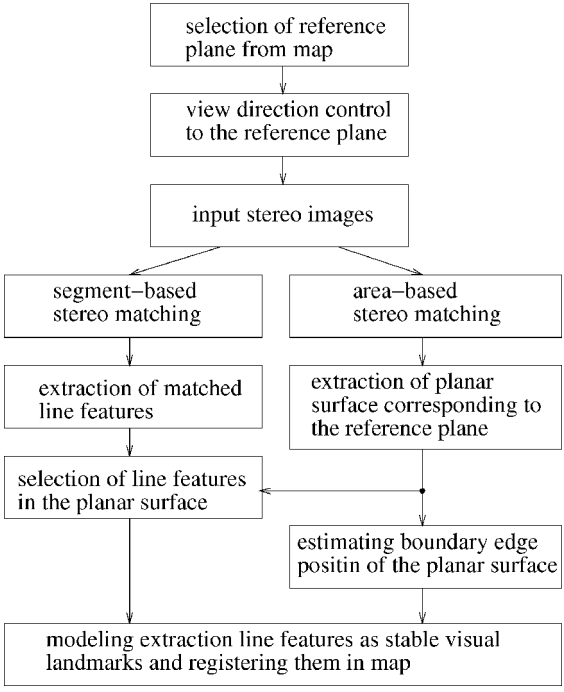


Figure 1. Processing flow for extracting stable visual landmarks.

extract vertical line segments, image processing techniques such as the Sobel edge detection, thinning and labeling are first applied to the input images, and then segment-based stereo matching is performed. As a result, we can select matched vertical line features. By applying the area-based stereo method to the input stereo images, we obtain a disparity image. A planar surface corresponding to the given reference plane is detected from the disparity image. From the two results, we select a set of line features on the planar surface by the AND operation. However, the extracted features and the planar surface include uncertainties due to errors of vision and motion [10]. The feature positions are modeled by a probabilistic distribution and estimated by using the extended Kalman filter (EKF) [11, 12]. Lastly, we match the extracted planar surface to the given map in order to estimate the feature positions in the map. This extraction process is performed on-line so as to adapt to changes of the environment and the robot's viewpoints.

The rest of this paper is organized as follows. Section 2 shows how to extract vertical line features and planar surfaces from stereo images. Section 3 presents a method for estimating stable visual landmarks. In this section, we also show how to use the EKF for combining the two results obtained by the two stereo methods. In Sections 4 and 5, we give experimental results from various environments using a three-wheeled mobile robot as shown in Fig. 2. From the experimental results, we show the validity of the proposed method. Lastly, we conclude this study and address some future work.

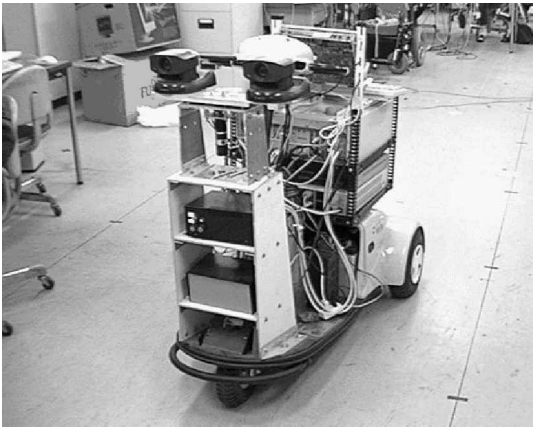


Figure 2. The mobile robot with the stereo vision system.

2. EXTRACTING VERTICAL LINE FEATURES AND PLANAR SURFACES

2.1. Extraction of vertical line segments

Vertical line segments are extracted from stereo images by image processing techniques such as smoothing, edge detection, labeling and thinning. By applying the Sobel operator to 3 × 3 image centered the pixel E as shown in Fig. 3, we can obtain the *edge intensity* I and *orientation* θ of the pixel E as follows:

$$I = \|\mathrm{d}x\| + \|\mathrm{d}y\|, \tag{1}$$

$$\theta = \tan^{-1}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right), \tag{2}$$

where

$$\mathrm{d}x = (C + 2F + I) - (A + 2D + G), \tag{3}$$

$$\mathrm{d}y = (G + 2H + I) - (A + 2B + C). \tag{4}$$

A series of adjacent pixels with the same orientation consists of a line segment with the same label. In this study we only use the vertical line segments with the orientation 2 or 6.

In order to obtain stable features for visual observation, the robot rejects short and low-contrast segments from the extracted line segments. Assuming that the positional order of the vertical line segments in the planar surfaces does not change, the remained segments in the stereo images are matched by the segment-based *DP* (dynamic programming) stereo matching method using the weighted graph composed by line similarities [13]. We define the similarity of each pair, $w(l, r)$, as the mean value of the orientation similarity (*OS*), the length similarity (*LS*) and the

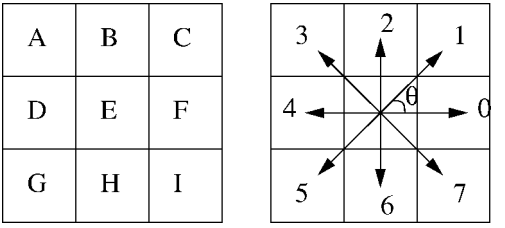


Figure 3. Eight orientations of pixel E.

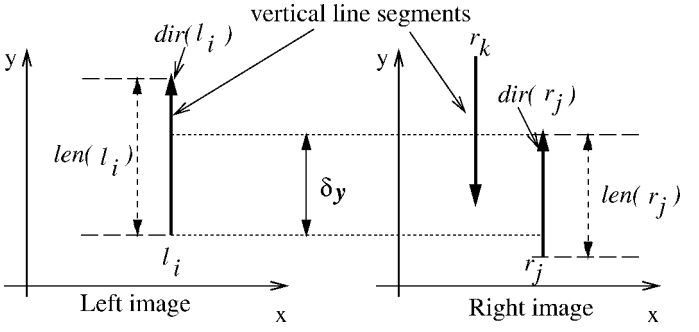


Figure 4. Similarity of stereo segments.

overlap length (*OL*) (see also Fig. 4):

$$w(l_i, r_j) = \frac{1}{3}\{OS(l_i, r_j) + LS(l_i, r_j) + OL(l_i, r_j)\}, \tag{5}$$

where

$$\begin{aligned} OS(l_i, r_j) &= \begin{cases} 1 & \text{if } \mathrm{dir}(l_i) = \mathrm{dir}(r_j), \\ 0 & \text{else}, \end{cases} \\ LS(l_i, r_j) &= \frac{\min\{\mathrm{len}(l_i), \mathrm{len}(r_j)\}}{\max\{\mathrm{len}(l_i), \mathrm{len}(r_j)\}}, \\ OL(l_i, r_j) &= \frac{1}{2}\left\{\frac{\delta y}{\mathrm{len}(l_i)} + \frac{\delta y}{\mathrm{len}(r_j)}\right\}. \end{aligned} \tag{6}$$

$\mathrm{dir}(l_i)$ and $\mathrm{len}(l_i)$ in (6) denote the orientation and the length of the segment l_i, respectively. Using the weighed graph constructed by combining all stereo segments, a set of pairs with the maximum cumulated weight from the start to the terminal node is selected as the matched segments, but remaining segments are rejected as candidates for the stable landmarks.

Figure 5 shows the results of *DP*-based stereo segment matching. The matched pairs are indexed with the same number. We can also see two mismatched segments numbered 21 and 22.

Figure 5. Results of DP-based stereo segments matching: (a) left image and (b) right image.

2.2. Extraction of disparity image

We use *SAD* (Sum of Absolute Differences) [14] as the area-based stereo matching method in order to extract the disparity image. In this study the planar surface is extracted from the disparity image. The *SAD* value for the evaluation of matching reliability is expressed by the following equation:

$$SAD = \sum_{i,j \in W} |I_{\mathrm{L}}(i, j) - I_{\mathrm{R}}(i + d, j)|, \tag{7}$$

where W denotes the size of area, and $I_{\mathrm{L}}(i, j)$ and $I_{\mathrm{R}}(i, j)$ are intensity at position (i, j) in the stereo images. d, called the *disparity*, is the difference between the center positions of the matching areas, which can be transformed into the range information. If the *SAD* value is smaller than the threshold value, the stereo matching between two areas is completed.

Since *SAD* needs higher computational costs than *DP*-based segment matching, the robot only performs the landmark extraction process when it has no landmark information or it cannot observe the landmarks because of positional limitations such as the short distance or large panning angle to the landmark. Figure 6b shows the disparity image obtained by applying *SAD* to the stereo images.

2.3. Uncertainty of stereo observation and disparity

Figure 7 shows the geometry of the stereo observation. The robot can predict the mean position of the reference plane in the stereo images. Considering the errors of motion and vision, the robot can calculate the positional uncertainty of the reference plane in the stereo images. Thus we can select valid disparity data corresponding to the reference plane from the disparity image. By comparing the disparity values of the matched segments with the valid disparity data, the robot can reject false matched segments.

Figure 6. Matching results by *SAD*: (a) input left image and (b) disparity image obtained by *SAD*. The brightness shows nearness from the robot's viewpoint.

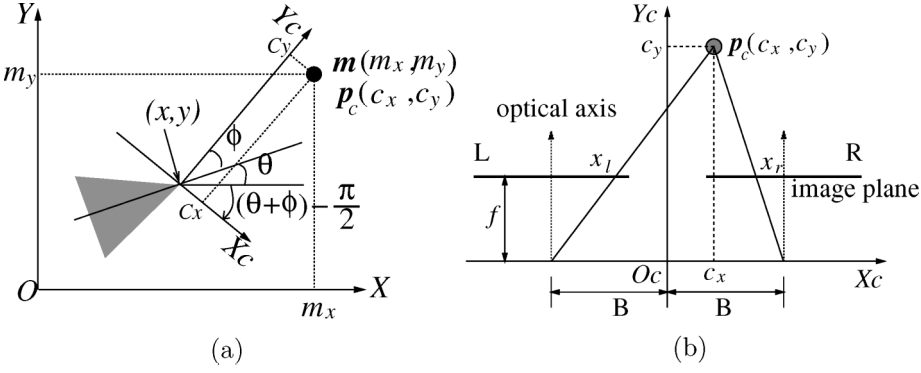


Figure 7. Geometry of the stereo vision system: (a) the robot coordinates system and (b) geometry of the stereo observation in the camera coordinates system.

Let the robot state be $\boldsymbol{x} = [x\ y\ \theta\ \phi]^T$ that denotes the position of the front wheel, x and y, the orientation of the robot, θ, and the view direction ϕ. Given a point $\boldsymbol{m} = [m_x\ m_y]^T$ located in the X–Y plane of the world coordinates system, we can obtain the position $\boldsymbol{p}_c = [c_x\ c_y]^T$ in the camera coordinates system as follows (see Fig. 7a):

$$\boldsymbol{p}_c = \boldsymbol{R}^{-1}_{[(\theta+\phi)-\frac{\pi}{2}]}[\boldsymbol{m} - \boldsymbol{x}_p], \tag{8}$$

where $\boldsymbol{R}_{[\rho]}$ and $\boldsymbol{x}_p = [x\ y]^T$ denote the rotation matrix with angle ρ and the position of the front wheel, respectively. Using the geometry of stereo observation as shown in Fig. 7b, $\boldsymbol{p}_c$ is projected into stereo images at position $\boldsymbol{q} = [x_l\ x_r]^T$, which is

obtained as follows:

$$\boldsymbol{q} = \begin{bmatrix} f\dfrac{c_x + B}{c_y} \\ f\dfrac{c_x - B}{c_y} \end{bmatrix}, \tag{9}$$

where f and B denote the focal length and the distance from a focal point to the origin of the camera coordinates system, respectively. Substituting c_x and c_y obtained from (8) for (9), we can rewrite (9) as follows:

$$\boldsymbol{q} = \begin{bmatrix} f\dfrac{\sin(\theta+\phi)\cdot(m_x - x) - \cos(\theta+\phi)\cdot(m_y - y) + B}{\cos(\theta+\phi)\cdot(m_x - x) + \sin(\theta+\phi)\cdot(m_y - y)} \\ f\dfrac{\sin(\theta+\phi)\cdot(m_x - x) - \cos(\theta+\phi)\cdot(m_y - y) - B}{\cos(\theta+\phi)\cdot(m_x - x) + \sin(\theta+\phi)\cdot(m_y - y)} \end{bmatrix} = \boldsymbol{H}(\boldsymbol{x}, \boldsymbol{m}). \tag{10}$$

By linearizing (10) using the Taylor series expansion around the mean $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{m}}$, the covariance matrix of the projection position, $\boldsymbol{\Sigma}_q$, is obtained as follows:

$$\boldsymbol{\Sigma}_q = \frac{\partial \boldsymbol{H}}{\partial \boldsymbol{x}} \boldsymbol{\Sigma}_x \frac{\partial \boldsymbol{H}}{\partial \boldsymbol{x}}^T + \frac{\partial \boldsymbol{H}}{\partial \boldsymbol{m}} \boldsymbol{\Sigma}_m \frac{\partial \boldsymbol{H}}{\partial \boldsymbol{m}}^T, \tag{11}$$

where $\boldsymbol{\Sigma}_x$ and $\boldsymbol{\Sigma}_m$ denote the covariance of the robot state and the point. If the point $\boldsymbol{m}$ has no uncertainty, $\boldsymbol{\Sigma}_m$ is zero.

As mentioned above, the disparity is the difference of the projected positions x_l and x_r in the stereo images.

$$d = x_\mathrm{l} - x_\mathrm{r}. \tag{12}$$

Substituting x_l and x_r obtained from (10) for (12), we can rewrite (12) to the following non-linear equation.

$$d = \frac{2fB}{\cos(\theta+\phi)\cdot(m_x - x) + \sin(\theta+\phi)\cdot(m_y - y)} = D(\boldsymbol{x}, \boldsymbol{m}). \tag{13}$$

Using the same linearization method, the variance of disparity, σ_d^2, is obtained as follows:

$$\sigma_\mathrm{d}^2 = \frac{\partial D}{\partial \boldsymbol{x}} \boldsymbol{\Sigma}_x \frac{\partial D}{\partial \boldsymbol{x}}^T + \frac{\partial D}{\partial \boldsymbol{m}} \boldsymbol{\Sigma}_m \frac{\partial D}{\partial \boldsymbol{m}}^T. \tag{14}$$

If the point $\boldsymbol{m}$ has no uncertainty, the covariance $\boldsymbol{\Sigma}_m$ is also ignored.

In this study, we define a *valid region* for a Gaussian random variable x with mean μ and variance σ^2 as follows:

$$[\mu - 3\sigma, \mu + 3\sigma]. \tag{15}$$

2.4. Extraction of a planar surface

Given a reference plane, the robot calculates the mean projection position corresponding to the given reference plane in the stereo images using (10) and (11), and the valid region of the projection position is determined from (15). The valid disparity region is also obtained by using (13)–(15). Figure 9 shows extraction results from Fig. 6b when the partition and the white board are given as the reference planes.

The disparity data can be transformed to a position vector $\boldsymbol{p}_c$ in the camera coordinates system by the following equation:

$$\boldsymbol{p}_c = \begin{bmatrix} x_c \\ y_c \end{bmatrix} = \begin{bmatrix} \dfrac{2Bx_l}{d} - B \\ \dfrac{2Bf}{d} \end{bmatrix} = \begin{bmatrix} \dfrac{B(x_l + x_r)}{x_l - x_r} \\ \dfrac{2Bf}{x_l - x_r} \end{bmatrix}, \tag{16}$$

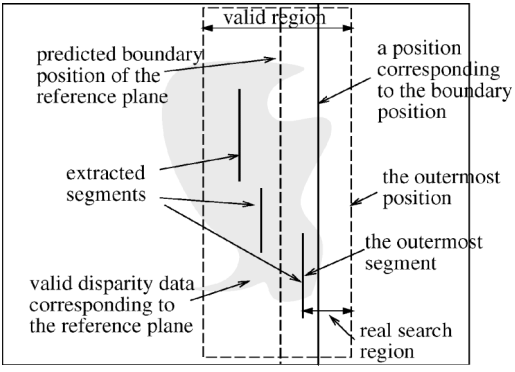


Figure 8. Search for the boundary edge position.

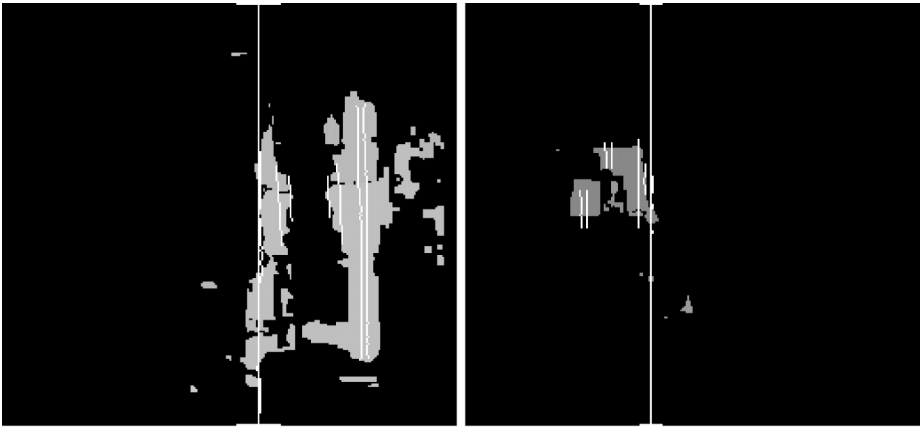


Figure 9. Extracted reference planes, line features belonging each reference plane and boundary edge positions: (a) partition and (b) white board.

where d denotes the disparity. Using the least-squares method based on the χ^2 fitting as shown (17), we can fit the N transformed position vectors (x_{c_i}, y_{c_i}) to the straight line model $y = a + bx$ in the camera coordinates system:

$$\chi^2 = \sum_{i=1}^{N} \frac{y_i - a - bx_i}{\sigma_i}. \tag{17}$$

In (17) σ_i denotes the standard deviation of y_i. Assuming $\sigma_i = 1$, we can obtain the line parameter $\boldsymbol{l} = [a\ b]^T$ and its covariance matrix $\boldsymbol{\Sigma}_l$ [15].

2.5. Estimation of boundary edges

The boundary edge position of the planar surface is estimated by fitting the boundary data to a vertical line. The robot first calculates the valid search region by (10), (11) and (15). Assuming that the boundary edge of the planar surface is straight and vertical, we search for the edge position with the minimum variance until the search position reaches the outermost line segment in the valid region because the boundary edge position must be located outside of the outermost line segment in the valid region (see Fig. 8). Long vertical lines in Fig. 9 show the estimated boundary edges of the given reference planes.

3. ESTIMATION OF STABLE VISUAL LANDMARKS

3.1. Combining two observations using the EKF

From the non-linear equation (16) we can calculate the matched segment position $\boldsymbol{p}_c = [x_c\ y_c]^T$ and the positional uncertainty $\boldsymbol{\Sigma}_{p_c}$ in the camera coordinates system, and the line $y = a + bx$ corresponding to the reference plane in the camera coordinates system is obtained as described in Section 2.4. However, the feature position may deviate from the extracted planar surface $y = a + bx$ by the uncertainties of vision and motion. Thus, we estimate the feature position and the positional uncertainty using the EKF. Assuming that the feature is located on the planar surfaces, the segment position is satisfied with the following constraint equation:

$$y_c - (a + bx_c) = 0. \tag{18}$$

Equation (18) is rewritten to the following equation:

$$Q(\boldsymbol{p}_c, \boldsymbol{l}) = 0, \tag{19}$$

where $\boldsymbol{l} = [a\ b]^T$ is the line parameter of the extracted planar surface corresponding to the reference plane. Using the Taylor series expansion around the mean $\tilde{\boldsymbol{p}}_c$ and $\tilde{\boldsymbol{l}}$, we can derive a new linear observation equation from (19) as follows [12]:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{p}_c + \boldsymbol{v}, \tag{20}$$

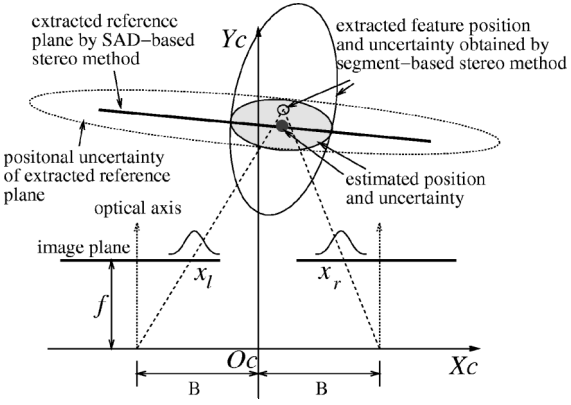


Figure 10. Combining two observations by the EKF.

where:

$$\boldsymbol{y} = -Q(\tilde{\boldsymbol{p}}_{\mathrm{c}}, \tilde{\boldsymbol{l}}) + \frac{\partial Q}{\partial \boldsymbol{p}_{\mathrm{c}}}\tilde{\boldsymbol{p}}_{\mathrm{c}},$$

$$\boldsymbol{H} = \frac{\partial Q}{\partial \boldsymbol{p}_{\mathrm{c}}},$$

$$\boldsymbol{v} = \frac{\partial Q}{\partial \boldsymbol{l}}(\boldsymbol{l} - \tilde{\boldsymbol{l}}).$$

The Kalman filter is an optimal estimation method based on prediction and observation of the linear system. We first set the feature position and positional uncertainty obtained by segment-based stereo matching to the predicted position and uncertainty, $\tilde{\boldsymbol{p}}_{\mathrm{c}}$ and $\boldsymbol{\Sigma}_{\tilde{p}_{\mathrm{c}}}$, respectively (see Fig. 10). The observation result obtained by (20) is then combined with the predicted information by the Kalman filter as follows:

$$\hat{\boldsymbol{p}}_{\mathrm{c}} = \tilde{\boldsymbol{p}}_{\mathrm{c}} + \boldsymbol{K}[\boldsymbol{y} - \boldsymbol{H}\tilde{\boldsymbol{p}}_{\mathrm{c}}], \tag{21}$$

$$\boldsymbol{\Sigma}_{\hat{p}_{\mathrm{c}}} = [\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H}]\boldsymbol{\Sigma}_{\tilde{p}_{\mathrm{c}}}, \tag{22}$$

where $\hat{\boldsymbol{p}}_{\mathrm{c}}$ and $\boldsymbol{\Sigma}_{\hat{p}_{\mathrm{c}}}$ denote the estimated position and uncertainty. $\boldsymbol{K}$ is called the Kalman gain obtained by the following equation:

$$\boldsymbol{K} = \boldsymbol{\Sigma}_{\tilde{p}_{\mathrm{c}}}\boldsymbol{H}^T[\boldsymbol{H}\boldsymbol{\Sigma}_{\tilde{p}_{\mathrm{c}}}\boldsymbol{H}^T + \boldsymbol{\Sigma}_V]^{-1}, \tag{23}$$

where $\boldsymbol{\Sigma}_V$ is the covariance matrix of the observation error $\boldsymbol{v}$ [9]. Using the EKF, we can estimate the position and uncertainty of the extracted features and boundary edges.

3.2. Modeling of stable visual landmarks

Using the known boundary edge positions of the reference plane, we can localize the feature position into the world coordinates system (see Fig. 11).

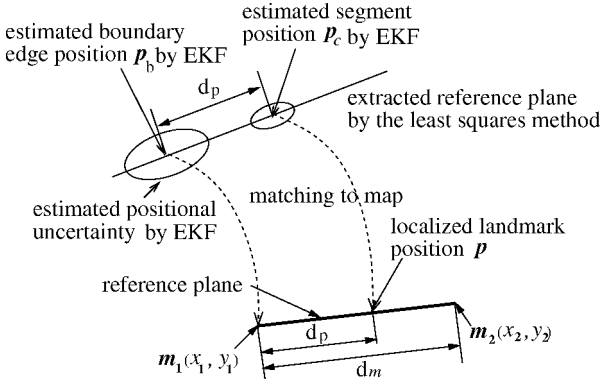


Figure 11. Localization of the extracted feature position by matching to the reference plane.

Let $\boldsymbol{p}_{\mathrm{b}}$ and $\boldsymbol{p}_{\mathrm{c}}$ be the positions of the extracted boundary edge and feature in the camera coordinates system. The boundary edge positions of the reference plane, $\boldsymbol{m}_1 = [x_1\ y_1]^T$ and $\boldsymbol{m}_2 = [x_2\ y_2]^T$, in the world coordinates system are given. Therefore the feature position $\boldsymbol{p}_{\mathrm{c}}$ in the world coordinates system is calculated by the following equation,

$$\begin{aligned}\boldsymbol{p} &= \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} \dfrac{d_p}{d_m}(x_2 - x_1) \\ \dfrac{d_p}{d_m}(y_2 - y_1) \end{bmatrix} \\ &= \boldsymbol{M}(\boldsymbol{m}_1, \boldsymbol{m}_2, \boldsymbol{p}_{\mathrm{b}}, \boldsymbol{p}_{\mathrm{c}}),\end{aligned} \tag{24}$$

where d_p and d_m denote the distance from the boundary edge to the feature and the width of the given reference plane, respectively. Using the linearization method, the covariance matrix $\boldsymbol{\Sigma}_p$ is also obtained. Since the map has no uncertainty, $\boldsymbol{\Sigma}_{m_1}$ and $\boldsymbol{\Sigma}_{m_2}$ are zero, and $\boldsymbol{\Sigma}_p$ can be expressed as follows:

$$\boldsymbol{\Sigma}_p = \frac{\partial \boldsymbol{M}}{\partial \boldsymbol{p}_{\mathrm{b}}} \boldsymbol{\Sigma}_{p_{\mathrm{b}}} \frac{\partial \boldsymbol{M}}{\partial \boldsymbol{p}_{\mathrm{b}}}^T + \frac{\partial \boldsymbol{M}}{\partial \boldsymbol{p}_{\mathrm{c}}} \boldsymbol{\Sigma}_{p_{\mathrm{c}}} \frac{\partial \boldsymbol{M}}{\partial \boldsymbol{p}_{\mathrm{c}}}^T, \tag{25}$$

where $\boldsymbol{\Sigma}_{p_{\mathrm{b}}}$ and $\boldsymbol{\Sigma}_{p_{\mathrm{c}}}$ denote the positional uncertainties of the extracted boundary edge and feature, respectively. The rank of the covariance matrix $\boldsymbol{\Sigma}_p$ in (25) is 1. This means that the uncertainty of the extracted landmarks is constrained on the reference plane.

3.3. Selection of stable visual landmarks

A feature such as the boundary edge is unstable for visual observation because such a feature is easily affected by changes of the background scene. In this study, the robot rejects features near the boundary edge as candidates of the stable landmarks. It also rejects a pair of extracted features with low length similarity (*LS*). We set the threshold value of the *LS* to 0.5.

4. EXPERIMENTAL RESULTS

Figure 12 shows the extracted result of the stable visual landmarks from stereo input images. We give the partition and the white board as the reference plane for the landmark extraction. The left side of the partition and the right side of the white board are located at $(0, 4000)$ and $(-550, 6000)$ mm, respectively. Long vertical lines show the estimated boundary edge positions, and short horizontal lines at the top and the bottom of them show the positional uncertainty.

Figure 13 shows the estimated results by using the EKF. The black bold lines are estimated positional uncertainties of the features by matching to the map. We can see that the uncertainty is constrained on the reference plane.

Table 1 shows the estimate results of extracted features. Features 8 and 7 in the partition and features 14, 15 and 16 in the white board are not valid for use as stable visual landmarks because they are located in the 3σ area from the boundary edge position. Feature 8 in the partition also has low length similarity. We can also see that landmarks have no uncertainty in depth because the features are constrained on the reference plane, i.e. $\sigma_y = 0$.

5. VISUAL NAVIGATION USING THE ON-LINE EXTRACTED LANDMARKS

We conducted experiments for visual navigation using the on-line extracted landmarks in our laboratory environment. As shown in Fig. 14, the robot behavior is simple. Given the 2D obstacle map, the robot first selects the reference plane and then extracts landmarks. After extracting landmarks, the robot moves to the next position using the landmark-based non-stop navigation strategy. Since the visual processing requires a great deal of computation time, the result of the visual observation is obtained after a delay. In order to overcome the delay problem, we

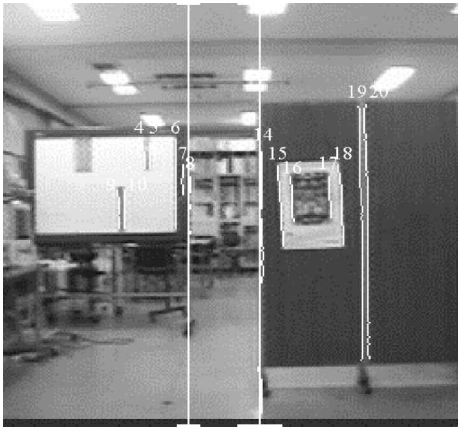


Figure 12. Extracted result of landmarks. Long vertical lines show the estimated boundary edge positions, and the short horizontal lines at the top and the bottom of them show the positional uncertainty of the boundary edge positions.

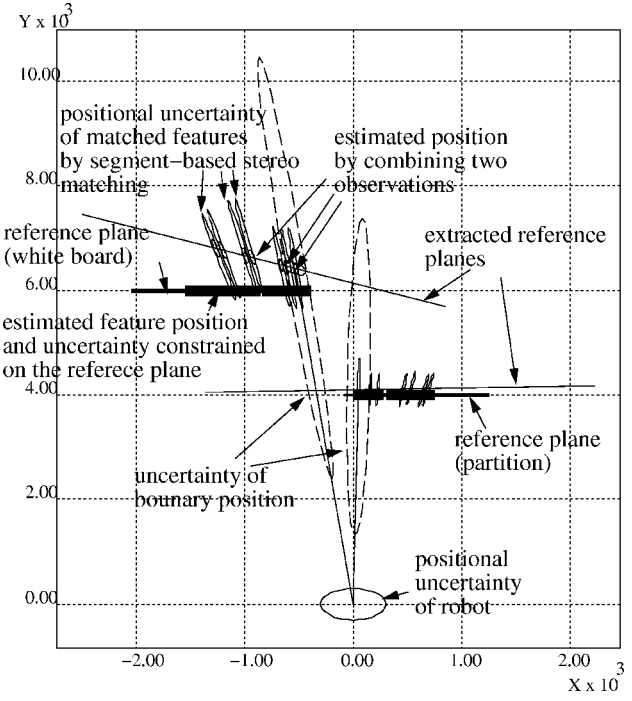


Figure 13. Uncertainty of the extracted landmarks shown in Fig. 12.

Table 1.
Results of extracted landmarks in the partition (p) and white board (w)

plane	segment	x mm	y mm	σ_x	σ_y	LS	validity
p	8	0.1	4000.0	28.3	0.0	0.5	no
p	7	109.4	4000.0	35.0	0.0	1.0	no
p	6	177.8	4000.0	35.0	0.0	0.9	yes
p	5	411.0	4000.0	35.0	0.0	1.0	yes
p	4	483.0	4000.0	35.0	0.0	1.0	yes
p	10	605.1	4000.0	35.0	0.0	1.0	yes
p	9	643.7	4000.0	35.8	0.0	1.0	yes
w	14	−550.2	6000.0	54.2	0.0	0.8	no
w	15	−622.6	6000.0	46.4	0.0	0.9	no
w	16	−698.3	6000.0	47.0	0.0	1.0	no
w	17	−1004.3	6000.0	50.4	0.0	0.9	yes
w	18	−1076.0	6000.0	51.0	0.0	0.9	yes
w	19	−1309.9	6000.0	53.1	0.0	1.0	yes
w	20	−1380.9	6000.0	53.9	0.0	1.0	yes

proposed a novel non-stop navigation method [9] whereby the robot moves to the destination without stopping while processing the visual observation data. When the robot obtains the processing result later, it retroactively integrates the delayed sensory information into the state estimation.

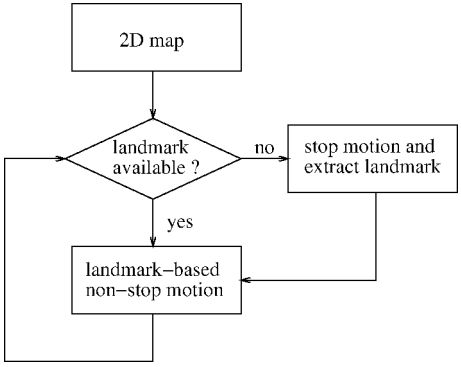


Figure 14. Robot behavior.

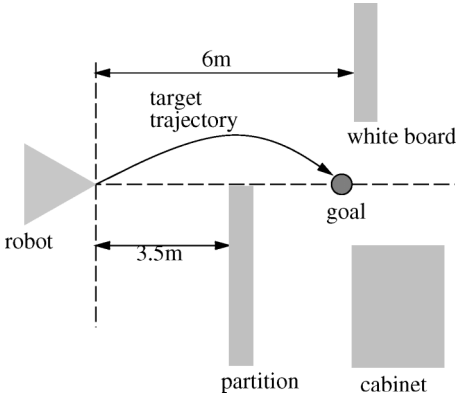


Figure 15. Indoor environment map for experimentation.

Although the robot has landmarks, it must stop and extract new landmarks if it cannot observe them from its current state. In this study the landmark extraction process is performed when the robot meets one of the following three cases: (i) the robot has no extracted landmarks, (ii) the distance to the landmarks is shorter than d_{th} and (iii) the gaze angle on the landmark is large than ϕ_{th}. From experimental results, we set d_{th} and ϕ_{th} to 1.5 m and 25°, respectively.

Once landmarks are extracted, the robot performs the landmark-based non-stop navigation repeatedly. The localization is completed by detecting line segments corresponding to the landmarks from the input image and estimating the current robot state using the EKF. The process for detecting line segments in the landmark-based non-stop navigation is the same as the process for extracting landmarks except for the extraction of the planar surface. One cycle for the non-stop navigation including the vision processing is 3 s and the robot can move 1 m as the maximum motion distance during one cycle.

Figure 15 shows the given map for the experimentation. The robot selects the reference planes as the partition and the white board that are located at 3.5 and 6 m, respectively. In this paper, we proposed an on-line extraction method for the visual

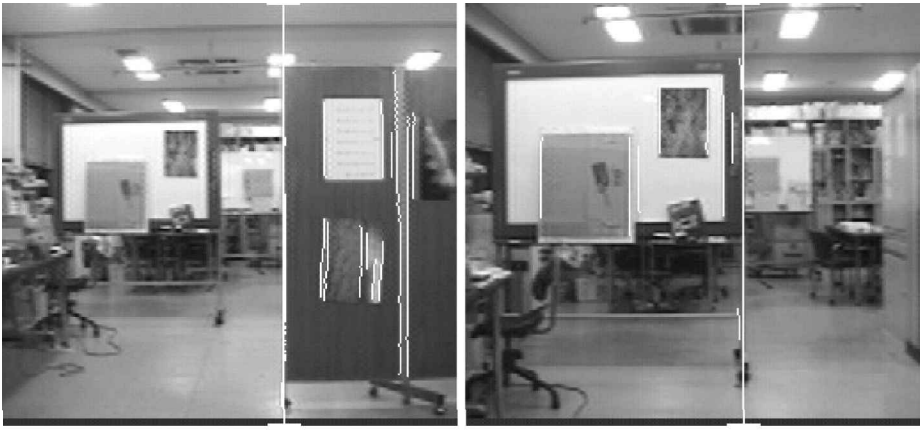


Figure 16. On-line extracted landmarks under normal conditions.

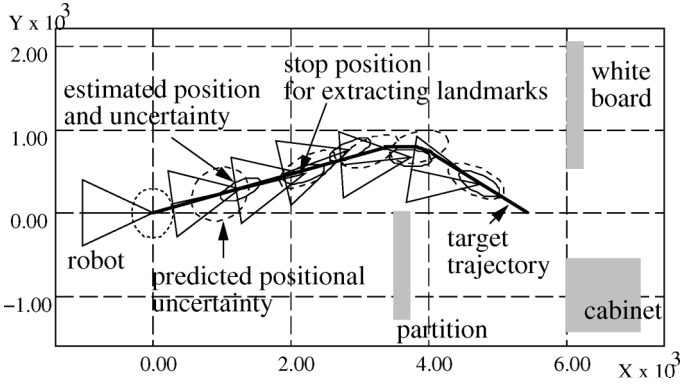


Figure 17. Motion result using on-line extracted landmarks under normal conditions.

landmarks in order to adapt to changes of the environment, such as the illumination condition and the obstacle configuration. Therefore, we conducted two experiments under different conditions; in the first case the robot moves under normal lighting conditions, and in the second case it moves in the same environment but the feature positions and the illumination conditions are changed.

Figure 16a and b shows the results of landmark extraction under normal conditions. White vertical lines are the extracted landmarks. The long vertical lines are the extracted boundary edges. Figure 17 shows the motion results using the on-line extracted visual landmarks. The target trajectory is automatically generated from the given map considering the uncertainty and width of the robot [16]. The ellipses drawn by the thin line show the estimated positional uncertainty of the robot. In this experiment, the robot did stop and extract new landmarks after moving two cycles. Figure 16b shows the result of the second extraction after moving two cycles from the start position. From the motion result shown in Fig. 17, we can see that

Figure 18. On-line extracted landmarks under dark illumination.

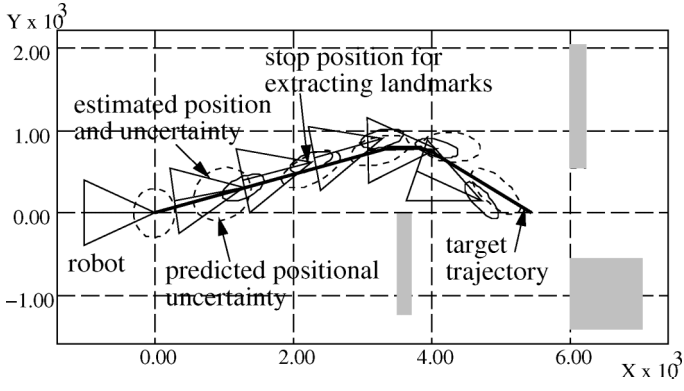


Figure 19. Motion result under dark illumination conditions.

navigation using the on-line extracted visual landmarks is successful. Figures 18 and 19 show the results of the landmark extraction and the robot's motion when the illumination conditions and the configuration of features are changed. Since the robot extracts visual landmarks on-line, the robot could reach the goal position without failure.

6. CONCLUSION

In this paper, we proposed a method for a vision-based mobile robot to extract visual landmarks automatically under uncertainty. The stable visual landmarks are extracted on-line and the robot could adapt visual observations to actual changes of the environment. As the result, it could perform stable visual observations without failure and reach the goal position successfully. Experimental results from various real scenes show the feasibility of the proposed method.

However, we did test our method to the nearly normal reference plane from the robot's view. Thus we are now studying a robust method to extract various planar surfaces including boundary edge information, and we also assumed that the environment is known and the obstacles are stationary. As described in this paper, the robot can roughly detect unknown obstacles not written in the map from the disparity data [14]. Therefore, if the robot can localize using the on-line extracted visual landmarks from the known obstacles, and if it can detect and avoid unknown obstacles from the disparity data, we can achieve more safe and efficient visual navigation. To extend this study to more natural indoor environment with unknown obstacles is an area of future work.

Acknowledgements

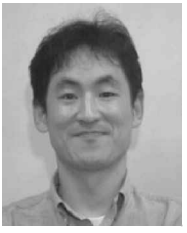
This work was (partly) supported by RRC program (Research Institute for Medical Instrumentation and Rehabilitation Engineering) of MOST and KOSEF.

REFERENCES

1. I. J. Cox, Blanche — an experiment in guidance and navigation of an autonomous robot vehicle, *IEEE Trans. Robotics Automat.* **7** (2), 193–204 (1991).
2. C. Becker, J. Salas, K. Tokusei and J. C. Latombe, Reliable navigation using landmarks, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, Nagoya, Japan, pp. 401–406 (1995).
3. R. Talluri and J. K. Aggarwal, Mobile robot self-location using model-image feature correspondence, *IEEE Trans. Robotics Automat.* **12** (1), 63–77 (1996).
4. G. Hager, D. Kriegman, E. Yeh and C. Rasmussen, Image-based prediction of landmark features for mobile robot navigation, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, Albuquerque, NM, pp. 1040–1046 (1997).
5. R. Sim and G. Dudek, Mobile robot localization from learned landmarks, in: *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Victoria, BC, Canada, pp. 1060–1065 (1998).
6. J. J. Little, J. Lu and D. R. Murray, Selecting stable image features for robot localization using stereo, in: *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 1072–1077 (1998).
7. J. Horn and G. Schmidt, Continuous localization for long-range indoor navigation of mobile robots, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, Nagoya, Japan, pp. 387–394 (1995).
8. D. Fox and W. Burgard, Active Markov localization for mobile robots, *Robotics Autonomous Syst.* **25**, 195–207 (1999).
9. I. Moon, J. Miura and Y. Shirai, On-line viewpoint and motion planning for efficient visual navigation under uncertainty, *Robotics Autonomous Syst.* **28**, 237–248 (1999).
10. J. Miura and Y. Shirai, Vision and motion planning for a mobile robot under uncertainty, *Int. J. Robotics Res.* **16** (6), 806–825 (1997).
11. P. S. Maybeck, The Kalman filter: an introduction to concepts, in: *Autonomous Robot Vehicles*, I. J. Cox and G. T. Wilfong (Eds), pp. 194–204. Springer-Verlag, Berlin (1990).
12. N. Ayache and O. D. Faugeras, Maintaining representations of the environment of a mobile robot, *IEEE Trans. Robotics Automat.* **5** (6), 804–819 (1989).
13. S. H. Lee and J. J. Leou, A dynamic programming approach to line segment matching in stereo vision, *Pattern Recognition* **27** (8), 961–986 (1994).

14. K. Kidono, J. Miura and Y. Shirai, Autonomous visual navigation of a mobile robot using a human-guided experience, in: *Intelligent Autonomous Systems 6*, E. Pagello *et al.* (Eds), pp. 620–627. IOS Press, New York (2000).
15. W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C*, 2nd edn. Cambridge University Press, Cambridge (1992).
16. I. Moon, J. Miura, Y. Yanagi and Y. Shirai, Planning of vision-based navigation for a mobile robot under uncertainty, in: *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Grenoble, France, pp. 1202–1207 (1997).

ABOUT THE AUTHORS



Inhyuk Moon received the BE and ME degrees in Electronics Engineering from Gyeongsang National University, Korea in 1992 and 1994, respectively, and PhD degree in Computer-controlledMechanical Systems from Osaka University, Japan in 1999. From April 1999 to March 2000, he was a Postdoctoral Rresearcher at the Organization for Hamamatsu Technopolis, Shizuoka, Japan. In March 2000, he joined the Technology Innovation Center for Medical Instruments and Department of Medical Engineering, Yonsei University, Wonju, Korea. Currently he is a Research Professor. His research interests include robotics, computer vision and intelligent rehabilitation systems.

Jun Miura received the BE degree in Mechanical Engineering in 1984, and the ME and DE degrees in Information Engineering in 1986 and 1989, respectively, all from the University of Tokyo, Tokyo, Japan. In 1989, he joined the Department of Computer-Controlled Mechanical Systems, Osaka University, Suita, Japan, where he is currently an Associate Professor. From March 1994 to February 1995, he was a Visiting Scientist at the Computer Science Department, Carnegie Mellon University, Pittsburgh, PA. He received the Best Paper Award from the Robotics Society of Japan in 1997. His research interests include robotics, artificial intelligence and computer vision.

Yoshiaki Shirai received the BE degree from Nagoya University in 1964, and the ME and the PhD degrees from the University of Tokyo in 1966 and 1969, respectively. In 1969, he joined the Electrotechnical Laboratory. From 1988 he has been a Professor of the Department of Mechanical Engineering for Computer-Controlled Machinery, Graduate School of Engineering, Osaka University. His research area has been computer vision, robotics and artificial intelligence. He is a member of the IEEE Computer Society, Information Processing Society of Japan, the JSME, the Japanese Society of Robotics, SICE and the Japanese Society of Artificial Intelligence.