Man Chasing Robot by an Environment Recognition Using Stereo Vision*

Yoshinobu SAWANO*, Jun MIURA, Yoshiaki SHIRAI****

* Fujitsu Co.ltd, Kawasaki, Kanagawa 211-8588, Japan
** Dept. of Computer-Controlled Mechanical Systems,
Osaka University, Suita, Osaka 565-0871, Japan
Email : {sawano,jun,shirai}@cv.mech.eng.osaka-u.ac.jp

Abstract

This paper describes a control method of chasing person by a vision guided mobile robot in an unknown indoor environment. The robot obtains the environment information by stereo vision and makes a map. From both the map and the observed person position, the robot decides the optimal path to chase the person while avoiding obstacles. Since the observation by stereo vision may include uncertainty, we model the uncertainty to calculate the reliability of the map to be obtained by integrating multiple observation results.

1. INTRODUCTION

One of the main issues of autonomous mobile robot studies is sensor data processing and planning when the robot has uncertainty in position estimation or in sensing. If the robot has a map of the environment, the robot can plan a path and observes landmarks recorded on the map for localization on the path [1]. By matching the observed landmarks to the map, the robot can move safely. There is a study in which a man guides a robot to the destination in an unknown environment [2]; the robot first makes the map of the environment and, then it moves autonomously. In our approach, the robot detects a person using vision. It also observes an environment by stereo vision to make a map. The robot then decides a path to chase the person while avoiding obstacles. The robot does not necessarily move on the path the target has moved, but chooses the optimal path.

Position estimate only by dead reckoning includes uncertainty and this uncertainty is accumulated as the robot moves. To reduce the uncertainty, the robot observes the environment by stereo vision. This observation has also uncertainty (or an error). For example, the robot sometimes observes nonexistent obstacles or sometimes fails to observe an obstacle due to the error in stereo matching. To recognize the environment more reliably, we develop the following method.

By modeling the error of observation based on the stereo vision characteristics, the robot evaluates the reliability of the observation result. If an obstacle is detected, its probability of existence is determined depending on the distance from the robot. The nearer to the robot the obstacle is, the larger the reliability is. We then model the positional uncertainty due to quantization error in stereo vision by using a two-dimensional normal distribution. We update the existence probability of each obstacle using the Bayes' theorem. If the robot detects an obstacle several times, then the existence probability of the obstacle becomes high. On the other hand, if the robot cannot detect an obstacle several times, the probability becomes low. The map contains the existence probability in the environment. The robot makes the map by

* In Proc. 2000 Int. Conf. on Machine Automation, pp. 389-394, 2000.

repeating the observation. We also apply the Extended Kalman Filter [3] to reduce the uncertainty in motion and obstacle position.

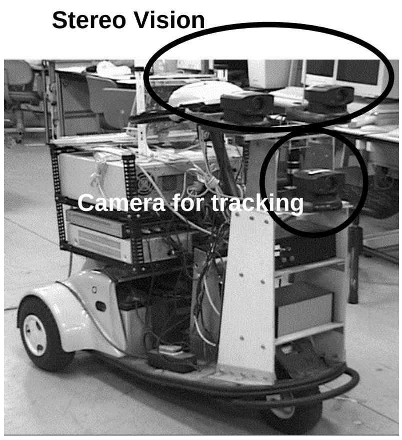


Figure 1. Mobile Robot

The simulation of this stochastic model shows the effectiveness of the uncertainty model. The experiment with the real robot shows the usefulness of the proposed method. Figure 1 shows a robot with three wheels. The robot can control a steering wheel, a camera head, and speed. A stereo pair of cameras is fixed on the camera head. Using the steering input and the odometer reading, the robot estimates its position by dead reckoning.

2. DETECTING OBSTACLES BY STEREO VISION

2.1 Stereo Vision System

We use a stereo vision to obtain range data of the environment. In stereo vision, it is necessary to find the correspondence between feature points in the left and the right images (see Figure 2). After smoothing the input image to suppress noises, we make edge images (see Figure 3) ; edges with high contrast are used as the feature points. A pair of points is considered to match if the sum of absolute difference (SAD) of the intensity value of 5 **x** 5 windows *W* around the points is small enough and minimum among SAD values computed for the possible disparity range. SAD is given by the following

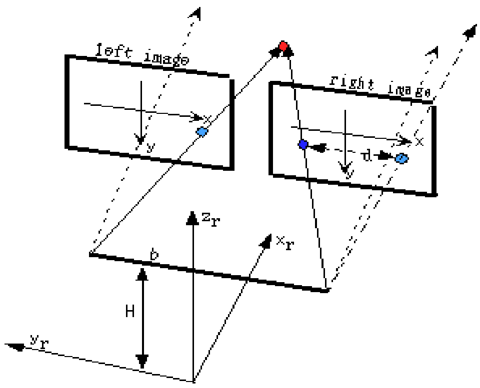
$$\sum_{(i,j)\in W} I_R(i,j) - I_L(i+d,j) \quad (1)$$

where $I_R(i,j)$ and $I_L(i,j)$ represent the intensity values of the point (i,j) in the right and the left images, respectively, and d represents the disparity. Figure 4 shows a result of the stereo method. In this figure, darker points indicate larger disparities.

Figure 2. Input Image

Figure 3. Edge image

Figure 4. Disparity image



2.2 Uncertainty of Observation

Once a set of matched feature points are obtained, their three-dimensional positions are computed by triangulation. If we set each parameter as Figure 5, these are given as

$$\begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = \frac{b}{d}\begin{bmatrix} f \\ -x \\ -y \end{bmatrix} + \begin{bmatrix} 0 \\ -\frac{1}{2}b \\ H \end{bmatrix} \quad (2)$$

Figure 5. Geometry of Stereo Vision

where $\mathbf{x}_r=(x_r,y_r,z_r)^t$ is a three-dimensional position; $\mathbf{x}=(x,y)$ is the point of right image; H is the height of the cameras from the; b is the distance between the cameras (called *baseline*); f and d are the focal length and the disparity, respectively.

We consider the three-dimensional position error due to the quantization error in the image. Using the Taylor series expansion and neglecting higher-order terms, we linearize equation (2). We then model the uncertainty of $\mathbf{x_r}$ as a two-dimensional normal distribution. The relation between the covariance matrix of disparity space ($\mathbf{x}_d=(x,y,d)^t$) and that of real space is given as

$$x_r = Ax_d + c, \qquad (3) \qquad\qquad \Lambda_r = A\Lambda_d A^T, \qquad (4)$$

where

$$A = \frac{b}{\hat{d}^2}\begin{bmatrix} 0 & 0 & -f \\ -\hat{d} & 0 & \hat{x} \\ 0 & -\hat{d} & \hat{y} \end{bmatrix} \qquad c = \frac{b}{\hat{d}}\begin{bmatrix} f \\ -\hat{x} \\ -\hat{y} \end{bmatrix} + \begin{bmatrix} 0 \\ -\frac{b}{2} \\ H \end{bmatrix}$$

3. DETECTING TARGET POSITION

We assume that the target person wears a colored cap and his height is known. We use the auto-tracking function of the camera (EVI-D30:SONY) to track the target. This camera can track the target quickly using color information. From the position of the target in the image and the pan and tilt angles, the target position $\mathbf{x_p}=(x_p,y_p)$ is given as (see Figure 6)

$$x_p = \begin{bmatrix} \dfrac{f + y_\omega \tan\theta_t}{f\tan\theta_t - y_\omega}(H_p - H_c) \\ -\dfrac{f\tan\theta_p + x_\omega}{f - x_\omega\tan\theta_p}\,\dfrac{f + y_\omega\tan\theta_t}{f\tan\theta_t - y_\omega}(H_p - H_c) \end{bmatrix}, \qquad (5)$$

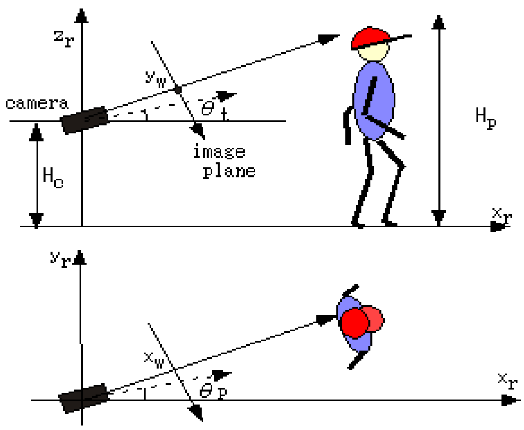


Figure 6. Observation of target

where x_W and y_W are the position in the image; f is the focal length, θ_p and θ_t are the pan and the tilt angle, respectively. Since the height of the cap may change as the target walks, the target position may include uncertainty. The positional uncertainty due to the height change is calculated similarly to the case of stereo uncertainty.

4 UPDATE POSITION OF ROBOT AND OBSTACLES

4.1 Uncertainty of Robot Motion

The position of robot is represented by vector $\mathbf{x_t}=[x_t,y_t, \boldsymbol{\theta}_t]^t$. Since the robot motion includes uncertainty, we estimate the position of the robot with uncertainty. Figure 7 shows the relationship between a control value (steering value $\boldsymbol{\lambda}_t$ and the movement distance d_t) and the movement of the

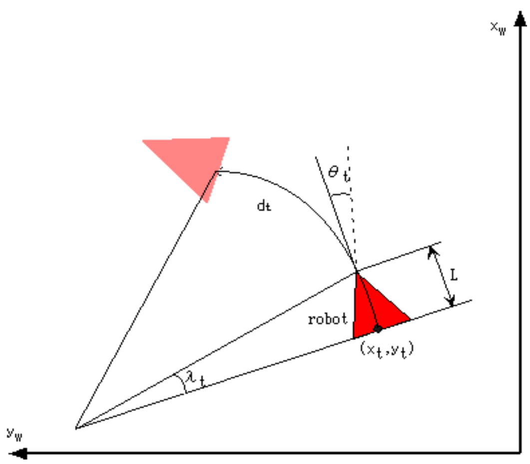


Figure 7. Movement of robot

robot. The position after movement is represented by the following equation [1] (for the case $\boldsymbol{\lambda} = 0$).

$$x_{t+1} = f(x_t, d_t) = \begin{bmatrix} x_t + \dfrac{L}{\sin\lambda_t}(\sin(\theta_t + \lambda_t + \dfrac{d_t}{L}\tan\lambda_t) - \sin(\theta_t + \lambda)) \\ y_t + \dfrac{L}{\sin\lambda_t}(\cos(\theta_t + \lambda_t) - \cos(\theta_t + \lambda + \dfrac{d_t}{L}\tan\lambda_t)) \\ \theta_t + \dfrac{d_t}{L}\tan\lambda_t \end{bmatrix} (\lambda \neq 0) \quad (6)$$

Since this equation is the non-linear function, we make this function linear as follows.

$$x_{t+1} = f(\hat{x}_{t/t}, \hat{d}_t) + \frac{\partial \hat{f}}{\partial x}(x_t - \hat{x}_{t/t}) + \frac{\partial \hat{f}}{\partial d}(d_t - \hat{d}_t) \quad (7)$$

We estimate the uncertainty of the robot motion by

$$\hat{x}_{t+1/t} = f(\hat{x}_{t/t}, \hat{d}_t) \qquad \Lambda^x_{t+1/t} = F_x \Lambda^x_{t/t} F_x^T + F_d \Lambda^d_t F_d^T$$
$$\textit{where} \qquad F_x = \frac{\partial \hat{f}}{\partial x},\ F_d = \frac{\partial \hat{f}}{\partial d} \quad (8)$$

4.2 Robot Position Estimate by Extended Kalman Filter

The robot records the position of vertical observed segments on the map. Using this segment, we update the position of segments on the map and the position of the robot using Extended Kalman Filter [3]. In Figure 8, suppose the robot observes segments M_i ($i=1 \sim n$) recorded in the map. Equation (9) is represents the relationships among the position and direction of the robot $\boldsymbol{x_t}=[x_t, y_t, \theta_t]^T$, the position of segments in the robot coordinate system $\boldsymbol{r_i}=[r_{xi}, r_{yi}, r_{zi}]^t$ and those in the world coordinate system $\boldsymbol{R_i}=[R_{xi}, R_{yi}, R_{zi}]^t$.

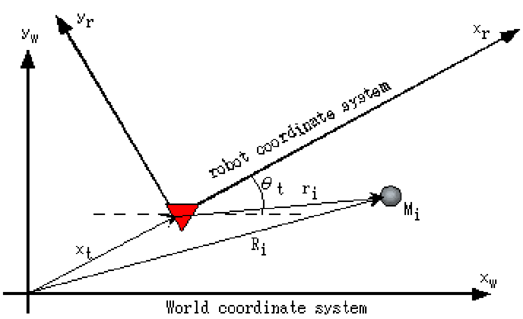


Figure 8. Observing a segment

$$h(X_t, r) = \begin{bmatrix} x_t + r_{x1}\cos\theta_t - r_{y1}\sin\theta_t - R_{x1} \\ y_t + r_{x1}\sin\theta_t - r_{y1}\cos\theta_t - R_{x1} \\ r_{z1} - R_{z1} \\ \vdots \\ x_t + r_{xn}\cos\theta_t - r_{yn}\sin\theta_t - R_{xn} \\ y_t + r_{xn}\sin\theta_t - r_{yn}\cos\theta_t - R_{xn} \\ r_{zn} - R_{zn} \end{bmatrix} = 0$$

$$\textit{where} \quad (9)$$

$$X_t = [x_t, R_1, \ldots, R_n]^t \qquad r = [r_1, \ldots, r_n]^t$$

Since equation (9) is a non-linear equation, we make this equation linear as follows.

$$h(X_t, r_t) = h(\hat{X}_{t/t-1}, \hat{r}_t) + \frac{\partial \hat{h}}{\partial X_t}(X_t - \hat{X}_{t/t-1}) + \frac{\partial \hat{h}}{\partial r_t}(r_t - \hat{r}_t) = 0 \quad (10)$$

We rewrite the equation (10) as follows.

$$-h(\hat{X}_{t/t-1},\hat{r}_t)+H_x\hat{X}_{t/t-1}=H_xX_t+H_r(r_t-\hat{r}_t)$$
$$H_x=\frac{\partial\hat{h}}{\partial X_t},\quad H_r=\frac{\partial\hat{h}}{\partial r_t} \tag{11}$$

In equation (12), we can consider that the left side is the observation and the right side is the state of the robot. We construct the Kalman Filter based on this linear system as follows.

$$\begin{aligned} K_t &= \Lambda^X_{t/t-1}H_X^T\left[H_X\Lambda^X_{t/t-1}H_X^T+H_r\Lambda^r_tH_r^T\right]^{-1}\\ \hat{X}_{t/t} &= \hat{X}_{t/t-1}-K_th\left(\hat{X}_{t/t-1},\hat{r}_t\right)\\ \Lambda^X_{t/t} &= \Lambda^X_{t/t-1}-K_tH_X\Lambda^X_{t/t-1}\end{aligned} \tag{12}$$

5. MODELING OBSERVATION AMBIGUITY

Since the observation result includes matching error, we cannot always trust the detected obstacles just by one observation. If the robot detects false obstacles and records them on the map, they could be problematic in generating a safe path to the target person. To solve this problem, we take the following strategy.

1. We consider far range data are less reliable than near range data.
2. We would like to reject accidentally-detected range data.

To evaluate the reliability of an observation result, we model the ambiguity of stereo vision statistically. The reliability is estimated by equation (13) and equation (14) based on the Bayes' theorem. Equation (13) applies to the obstacles that the robot succeeded to detect, and equation (14) applies to the region between the robot and the detected obstacle.

$$P_n\left(E(x)\mid O(x)\right)=\frac{P_n(O(x)\mid E(x))\,P_{n-1}(E(x))}{P_n(O(x))} \tag{13}$$

$$P_n\left(E(x)\mid \overline{O}(x)\right)=\frac{P_n(\overline{O}(x)\mid E(x))\,P_{n-1}(E(x))}{P_n(\overline{O}(x))} \tag{14}$$

where *P(E(x))* represents the probability that the obstacle exists at position ***x***, and its initial value is 0.5. *P(O(x))* represents the probability that the obstacle is observed at position ***x***.

P(O(x)|E(x)) is given as Figure 9; *P(O(x)|E(x))* changes depending on the distance from robot. Moreover *P(O(x)|E(x))* is distributed around the detected position using the model of the positional uncertainty of stereo vision. $P(O(x)\mid\overline{E}(x))$ is set to constant 0.05 at every position. Figure 10 shows the result of a simulation using this method; the existence probability distribution in the environment after the robot made 7 observations are shown. In the figure, the whiter a point is, the higher the existence probability of obstacle at the point is. The robot decides the position of obstacles by thresholding the existence probability.

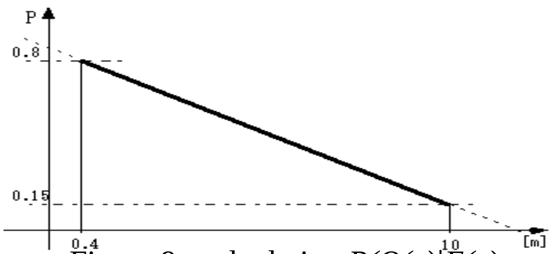


Figure 9. calculating *P(O(x)|E(x)*

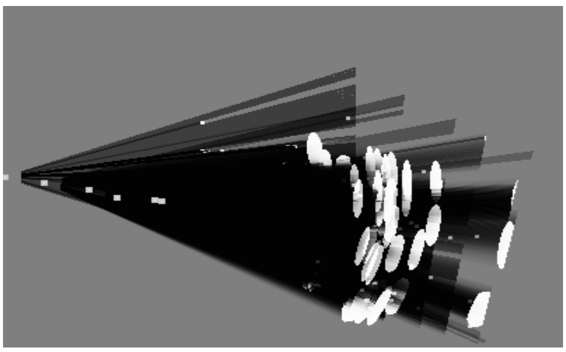


Figure 10. Existence probabilities after 7 observations.

6. PATH GENERATION

Since the robot has a motion uncertainty and the robot is bigger than the target, the path where target has moved is not necessarily safe for the robot movement. The obstacle regions on the map are enlarged by a half of the robot width and the estimated motion uncertainty. We call these enlarged regions dangerous regions as shown in Figure 11. The robot decides the path to avoid this region. To decide the optimal path for chasing depending on the situation, the robot takes following steps (see Figure 11).

The robot chooses the direct path to the last position of the target if obstacles do not exist on the path. If it is impossible to take such a path, the robot tries to generate a direct path to one of the previous target positions if the path is not obstructed by any obstacles. If the robot cannot take such paths, it chooses the edge of an obstacle which is nearest to the target position as the subgoal for chasing.

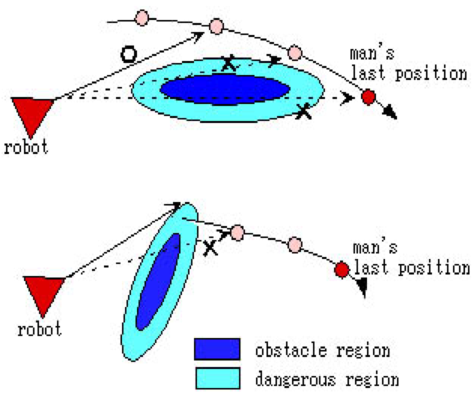


Figure 11. The shortest path

7. CHASING EXPERIMENT

The robot repeats the following steps for chasing: (1) detecting target; (2) obtaining range data; (3) updating the map; (4) generating the path; and (5) moving. In this experiment, we used a simplified version of map making, where the vision uncertainty due to quantization error is not considered; this can still generate a safe path because a enough margin is considered in path generation. Figure 12 shows snapshots of a chasing experiment.

Figure 12. Chasing experiment

8. SUMMARY

We have proposed a method that the robot can chase a person in an unknown environment even if the environment information includes uncertainty. To cope with the ambiguity in obstacle detection by stereo vision, we have developed a stochastic model of the ambiguity to evaluate the reliability of observation. Experimental results show the effectiveness of the method.

REFERENCES

1. I. Moon, J. Miura, and Y. Shirai, "On-line Viewpoint and Motion Planning for Efficient Visual Navigation under Uncertainty," *Robotics and Autonomous Systems*, Vol. 28, pp. 237-248, 1999.
2. K. Kidono, J. Miura, and Y. Shirai, “Autonomous Navigation of a Mobile Robot Using a Human-Guided Experience”, Proc. Asian Conference on Computer Vision, Vol.1, pp. 449-454, 2000.
3. N. Ayache and O.D. Faugeras, "Maintaining Representaions of the Environment of a Mobile Robot", *IEEE Trans. on Robotics and Automat.*, Vol. 5, No. 6, pp. 804-819, 1989.