Proc. 1992 IEEE Int. Conf. on Robotics and Automation,
pp. 1772-1777, Nice, France, 1992.

Vision-Motion Planning with Uncertainty

Jun MIURA Yoshiaki SHIRAI
Dept. of Mech. Eng. for Computer-Controlled Machinery,
Osaka University, Suita, Osaka 565, Japan
jun@ccm.osaka-u.ac.jp

Abstract

This paper describes a new framework for a planning of vision and motion for a mobile robot. For a planning in a real world, the uncertainty and the cost of visual recognition are important issues. A robot has to consider a trade-off between the cost of visual recognition and the effect of information obtained by recognition. A problem is to generate a sequence of vision and motion operations based on a sensor information which is an integration of the current information and the predicted next sensor data. The problem is solved by recursive prediction of sensor information and the recursive search of operations. As an example of sensor modeling, we describe a model of stereo vision in which correspondence of wrong pairs of features as well as quantization error is considered. Using the framework, a robot can successfully generate a plan for a real world problem.

I. Introduction

Mobile robot with vision is one of the most interesting subjects in robotics research. For a planning in a real world, the uncertainty and the cost of visual recognition are important issues. A robot has to consider a trade-off between the cost of visual recognition (including the cost of motion for recognition) and the effect of information obtained by vision. A robot also has to consider both the cost of visual recognition and that of motion to generate an optimal sequence of vision and motion operations.

There have been many works on the path planning of mobile robots [2][9]. Most of them assume that the environment is completely known. Clearly, this assumption is inappropriate in a real world.

Cameron et al.[3] applied Bayesian decision theory to selecting the optimal sensing. Hutchinson et al.[5] used Dempster-Shafer theory to represent uncertainty in object identification. These methods do not consider the cost of recognition. Besides, since they select a sensing operation one by one using some utility function, solutions are only locally optimal. In the planning system by Dean et al.[4], a mobile robot selects one sensing operation that minimizes the expectation of the total cost of the task. The selected sensing operation is still locally optimal.

If multiple sensing operations and motions are required to achieve a given goal, both recursive prediction of sensor information and the recursive search of optimal operations are necessary for making a globally optimal plan. This paper describes the three important topics:

- *method of predicting sensor information*,
- *formulation of planning problem*, and
- *modeling of uncertainty of vision*.

II. Motion Selection with Uncertain Information

Let us consider figure 1. A robot is going to *Goal*. There are two paths, *pass* and *detour*, and the former is shorter. When the robot estimates the distance between objects, the estimated value will be distributed because of uncertainty of vision. In our framework, the robot, whose width is W_{robot}, decides on the next motion based on the following criterion. In case (a) of figure 2, the robot decides that the *pass* is passable and takes it. In case (c), the robot decides that *pass* is impassable and takes *detour*. In case (b), the robot cannot decide whether *pass* is passable, and further sensing is needed. The robot, however, will take *detour* if taking it has less cost than doing extra sensing and taking *pass*. For a planning, therefore, it is necessary to consider the uncertainty and the cost of visual recognition and those of motion.

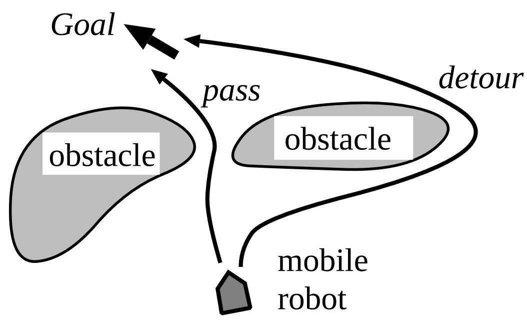

Fig. 1: A robot passing through obstacles.

III. Prediction of Sensor Information

To make a plan including sensing, a robot must be able to predict sensor information. We here explain how to predict sensor information using an example. Suppose that a robot is determining the location of a feature F and a two-

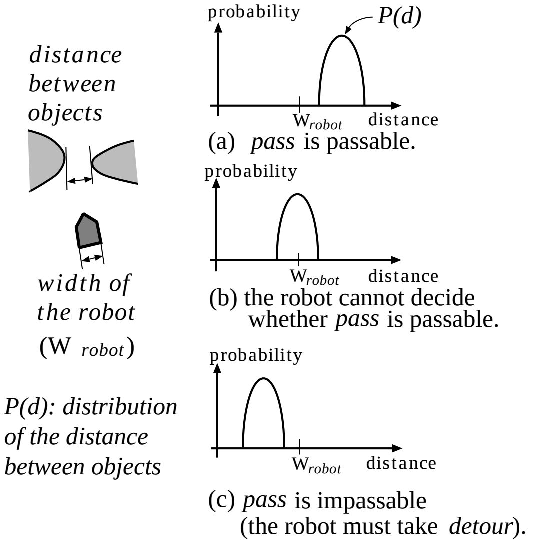

Fig. 2: Relation between the robot width and the distribution of the distance between objects.

dimensional distribution which represents the positional uncertainty of F has been acquired by sensing so far. Let P_1 in figure 3 denote this distribution. If the true position of F is $\boldsymbol{x}_0$, the actual observed position from another viewpoint will be distributed. Let P_2 in the figure denote this distribution. P_2 depends on observation conditions such as the distance and the direction of observation. Now, let assume that the robot actually observed F at $\boldsymbol{y}_0$ and we call this event Y. The new positional distribution of F after Y is computed by integrating the prior distribution P_1 and the event Y. This integration is carried out using Bayes rule. The new distribution becomes

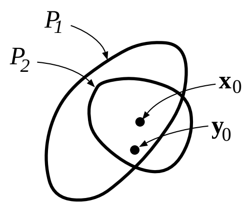

Fig. 3: Distributions of the position of F.

$$P_{after}(\boldsymbol{x}) = P(\boldsymbol{x}|Y) = \frac{P(\boldsymbol{x})P(Y|\boldsymbol{x})}{\int_{\boldsymbol{x}} P(\boldsymbol{x})P(Y|\boldsymbol{x})d\boldsymbol{x}}. \quad (1)$$

Suppose P_2 is represented as $P_2(\boldsymbol{u})$ using $\boldsymbol{u}$, a position vector of the observed position with respect to the true position. Then, $P(Y|\boldsymbol{x})$ becomes $P_2(\boldsymbol{y}_0 - \boldsymbol{x})$ and equation (1) is rewritten as

$$P_{after}(\boldsymbol{x}) = \frac{P_1(\boldsymbol{x})P_2(\boldsymbol{y}_0 - \boldsymbol{x})}{\int_{\boldsymbol{x}} P_1(\boldsymbol{x})P_2(\boldsymbol{y}_0 - \boldsymbol{x})d\boldsymbol{x}}. \quad (2)$$

The probability of acquiring this information is the product of the probability that true point is $\boldsymbol{x}_0$ and the probability that the observed position is $\boldsymbol{y}_0$, that is, $P_1(\boldsymbol{x}_0)P_2(\boldsymbol{y}_0 - \boldsymbol{x}_0)$. A robot can predict what information is acquired with how much probability by computing equation (2) for every possible combination of $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$.

When a robot observes the environment at many positions, the uncertainty of motion also needs to be considered. In such case, P_2 is calculated considering such uncertainty and then, equation (2) is applied.

IV. Formulation of the Vision-Motion Planning Problem

An optimal plan of vision and motion operations of a robot depends on the environment. It is difficult to generate the optimal plan in advance when only uncertain information about the environment is available. Here, we define that the optimal plan is a plan which minimizes the expectation of the total cost. We assume that the state of the environment is represented by a multivariate distribution, each variable of which is a property of the environment. We call such multivariate distribution *distribution information* and describe it by $\boldsymbol{D}$. We here derive a recurrence formula which relates the current position $\boldsymbol{x}_i$ and distribution information $\boldsymbol{D}_i$ with the optimal next observation position $\boldsymbol{x}_{i+1}$ and observation $\boldsymbol{o}_{i+1}$.

Since an observed data is a vector of properties, the uncertainty of an observed data becomes a multivariate distribution. Let $P_{obsd}(\boldsymbol{s}; \boldsymbol{D}_i, \boldsymbol{x}_{i+1}, \boldsymbol{o}_{i+1})$ denote the probability of getting an observed vector $\boldsymbol{s}$. Also, let $fuse(\boldsymbol{D}, \boldsymbol{s}, \boldsymbol{x}, \boldsymbol{o})$ be a function which computes a new distribution information from $\boldsymbol{D}$, $\boldsymbol{s}$, $\boldsymbol{x}$ and $\boldsymbol{o}$ using the method described in the previous section. A robot can predict that a distribution information $fuse(\boldsymbol{D}_i, \boldsymbol{s}, \boldsymbol{x}_{i+1}, \boldsymbol{o}_{i+1})$ will be acquired with the probability $P_{obsd}(\boldsymbol{s}; \boldsymbol{D}_i, \boldsymbol{x}_{i+1}, \boldsymbol{o}_{i+1})$ after an observation $\boldsymbol{o}_{i+1}$ at $\boldsymbol{x}_{i+1}$.

Because Bayes rule is used for integrating information, the integration result includes all information acquired in recognition processes in the past. Therefore, the optimal plan based on some specific position and distribution information is independent of how such information has been acquired. Consequently, the minimum cost at $\boldsymbol{x}_i$ with distribution information $\boldsymbol{D}_i$ becomes the minimum of the sum of the following:

1. the cost of motion to the next observation point $\boldsymbol{x}_{i+1}$.
2. the cost of the next observation $\boldsymbol{o}_{i+1}$.
3. the minimum cost from $\boldsymbol{x}_{i+1}$ to the goal point. This is a weighted sum of minimum expectations of the cost, each of which depends on each possible sensor information $\boldsymbol{s}$ obtained by $\boldsymbol{o}_{i+1}$ and weighted with the probability of $\boldsymbol{s}$.

Therefore, the problem of vision-motion planning with uncertainty is formulated as follows:

$$
\begin{aligned}
&C_o(\boldsymbol{x}_i, \boldsymbol{D}_i) = \\
&\qquad \min_{\substack{\boldsymbol{x}_{i+1} \in \mathcal{X}, \\ \boldsymbol{o}_{i+1} \in \mathcal{O}}} \begin{array}{l} (C_m(\boldsymbol{x}_i, \boldsymbol{x}_{i+1}) + C_v(\boldsymbol{o}_{i+1}) + \\ \qquad min\_cost(\boldsymbol{x}_{i+1}, goal)), \end{array} \\
&min\_cost(\boldsymbol{x}_{i+1}, goal) = \\
&\quad \int P_{obsd}(\boldsymbol{s}; \boldsymbol{D}_i, \boldsymbol{x}_{i+1}, \boldsymbol{o}_{i+1}) \cdot \\
&\qquad\qquad C_o(\boldsymbol{x}_{i+1}, fuse(\boldsymbol{D}_i, \boldsymbol{s}, \boldsymbol{x}_{i+1}, \boldsymbol{o}_{i+1})) d\boldsymbol{s}
\end{aligned} \tag{3}
$$

$C_o(\boldsymbol{x}, \boldsymbol{D})$: The optimal cost with $\boldsymbol{D}$ at $\boldsymbol{x}$.
$C_m(\boldsymbol{x}, \boldsymbol{y})$: The cost of motion from $\boldsymbol{x}$ to $\boldsymbol{y}$.
$C_v(\boldsymbol{o})$: The cost of observation $\boldsymbol{o}$.
$\mathcal{X}$: A possible range of $\boldsymbol{x}_{i+1}$.
$\mathcal{O}$: A possible range of $\boldsymbol{o}_{i+1}$.

Since the planning problem is represented by a recurrence formula, a robot can get the optimal sequence of observation points using dynamic programming (DP). As mentioned in section II, a robot is sometimes able to decide on the final motion without further observation (cases (a) and (c) in figure 2). In such situations, the cost to the destination can be computed without using equation (3) and the recursive computation terminates.

In [8], we analyzed a simple vision-motion planning problem and concluded that hill-climbing is useful to limit a search space at each stage of DP. However, even if we combine DP with hill-climbing, the planning problem is not yet free from combinatorial explosion. To reduce the computational cost, pruning using problem specific constraints is necessary. Pruning branches with little probability is also useful although plans become suboptimal.

V. Modeling Uncertainties of Stereo Vision

Our framework does not make any assumptions on a model of sensing uncertainty except that uncertainties are represented by probabilistic distributions. However, a concrete model of sensors is needed to solve a planning problem of an actual robot. We here describe a model of stereo vision as an example of modeling uncertainty. Although most researches [1][6][7] deal with only quantization error in edge detection, uncertainty caused by correspondence of wrong pairs of features also needs to be investigated because this uncertainty is much greater than quantization errors.

A. Quantization Error in Edge Detection

Suppose a point at (x, z) be observed at X_L and X_R on the left and the right images, respectively. We assume that the real positions of X_L and X_R are normally and independently distributed, and linearize the equation of the image projection [6]. Then, the uncertainty of (x, z) can be represented by a two-dimensional normal distribution. Although the distribution of (X_L, X_R) depends on the contrast of each edge, we currently assume that this distribution always has the same covariance matrix.

Using this model, distributions of other properties can be computed. For example, suppose that a robot is measuring the distance between two features. Letting $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ be positions of two features and their distributions be $P_1(\boldsymbol{x}_1)$ and $P_2(\boldsymbol{x}_2)$, the distribution of the distance d, $P(d)$, is given by

$$
P(d) = \int\int_{\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| = d} P(\boldsymbol{x}_1) P(\boldsymbol{x}_2) d\boldsymbol{x}_1 d\boldsymbol{x}_2. \tag{4}
$$

Using discretized distributions, we calculated the uncertainty of the distance (the difference of the maximum and the minimum distance) at each viewpoint. Figure 4 shows the result.

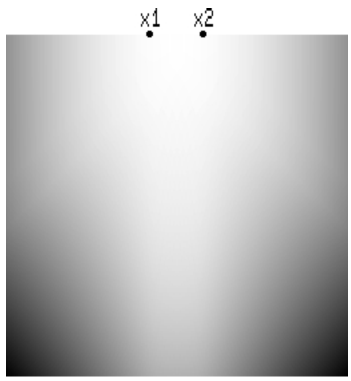

Fig. 4: Change of the uncertainty of the distance between objects. The brighter a point is, the less uncertain an measurement is.

B. Estimating Uncertainty Caused by Correspondence of Wrong Pairs

The most important problem in stereo vision is to determine matching pairs of features. We employ a local disparity histogram (LDH)-based method [10] to determine matching pairs. The outline of this method is as follows. First, for every features in the left image, candidates of the corresponding features in the right image are obtained on the basis of the similarity of the contrast and the direction. Then, the image is divided into small overlapping areas (windows) and the LDH in each window is computed from candidates of matching pairs. If the LDH has one prominent peak, the disparity of a window is determined and the matching pairs of features are established in the window.

When an LDH has multiple prominent peaks, a robot cannot decide which is true unless any other clues are available. In such case, a robot consider that each disparity corresponding to each peak may be true. Suppose a robot is estimating the position of a feature and the LDH of a window including the feature is figure 5. First, a robot eliminates clusters which are too small (A) or which do not include the disparities for the candidates of corresponding features (B and D). Two clusters (C and E) are accepted and peaks of those clusters indicate possible disparities. For each disparity, a normal distribution for the positional uncertainty of the feature can be computed. Then, a robot considers that the positional distribution of the feature is represented by the weighted sum of such normal distributions, where each weight is given by the normalized number of points in each cluster.

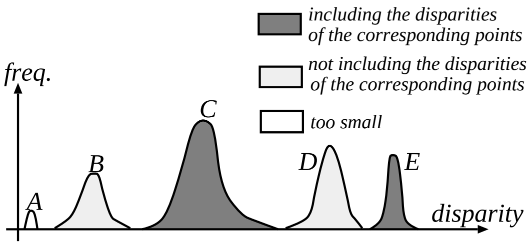

Fig. 5: An example LDH.

We applied this procedure to the real data. Figure 6 shows a pair of stereo images. Suppose a robot is estimating the distance between boxes (the distance between the rightmost point of the left box and the leftmost point of the right box). Small squares in the left image indicate windows for calculating LDHs. Figure 7 shows the computed LDHs. There are two acceptable clusters (black ones in the figure) for both windows. Peak disparities and weights are also indicated. The true disparity is 53. Arrows in the right image indicate four positions of corresponding features (two for true matchings and two for false ones) for computed disparities. Figure 8 shows the positional distributions of the two features. The distribution of the left feature is composed of L_1 and L_2 and that of the right one is composed of R_1 and R_2, where L_2 and R_2 are for false points.

C. Selecting Views to Resolve Ambiguity Caused by Multiple Matching Pairs

When there are multiple possible disparities, it is necessary for a robot to choose the next viewpoint from which only one of such disparities will be known to be true. Let us consider figure 9. There are two candidate positions for a feature estimated by the first observation. Suppose T is the true position and F is the false one. Of course, for the robot, the position of the feature is still the weighted sum of two normal distributions centered at T and F. If a wrong matching about the feature occurs again by the second observation, the false position will lie on the line determined by T and either of two lens centers. Let F' be the second false position. As mentioned in section III, the information after two observations is computed by integrating the first and the second observed data. If F and F' are so apart from each other that they cannot be integrated, only T will remain and the matching ambiguity will be resolved.

F and F' can be integrated if their corresponding distributions can be integrated. We can check whether two normal distributions can be integrated by using Mahalanobis

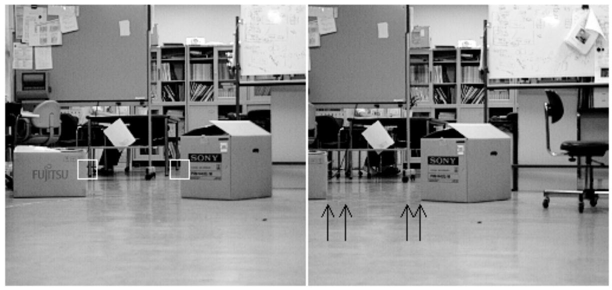

Fig. 6: A pair of stereo images.

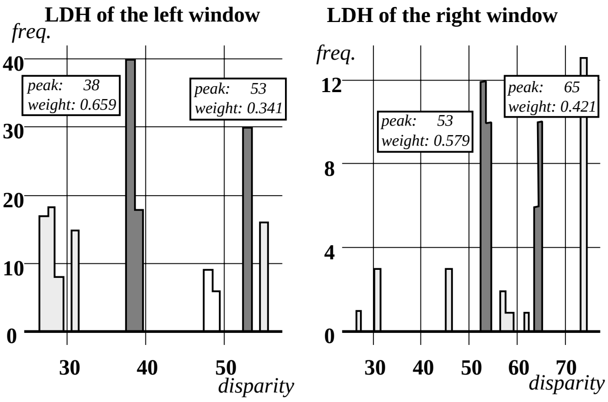

Fig. 7: LDHs for the left and the right feature points.

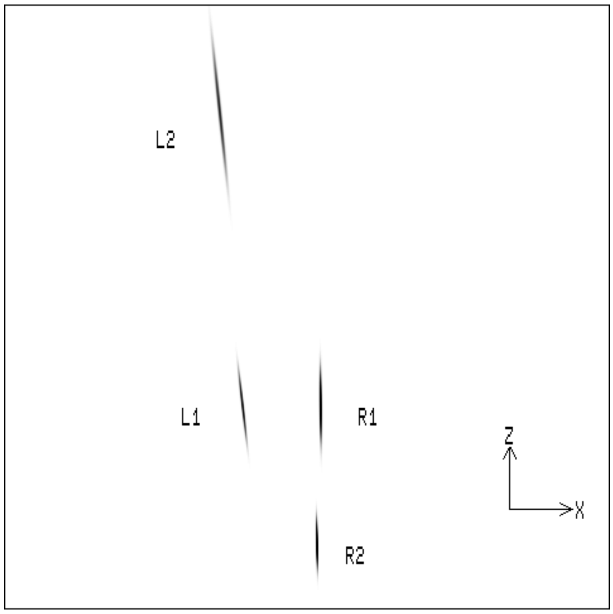

Fig. 8: Distributions of the two feature points. The darker a point is, the higher its probability is.

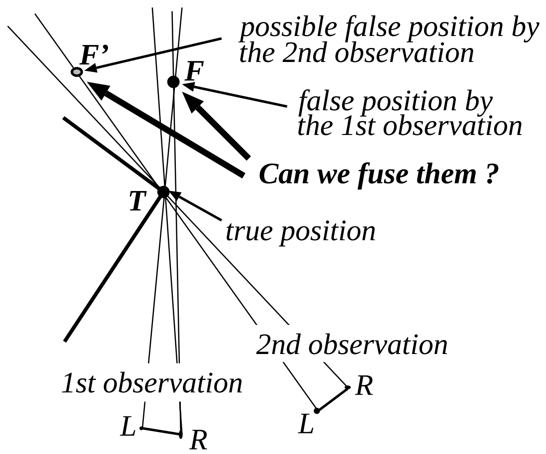

Fig. 9: Viewpoints for resolving ambiguity.

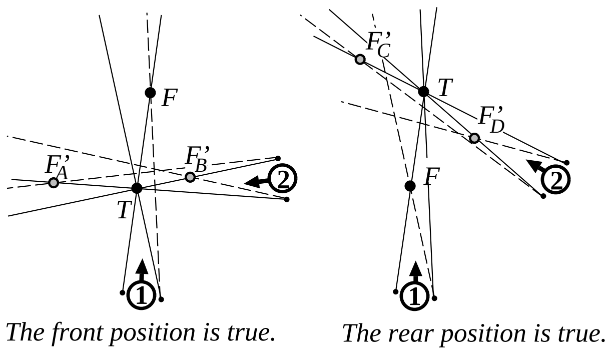

Fig. 10: Four cases for the test of the Mahalanobis distance.

distance [1]. Let $\boldsymbol{\mu}_F$, $\boldsymbol{\mu}_{F'}$, C_F and $C_{F'}$ denote means and covariance matrices of distributions of F and F'. The Mahalanobis distance between two distributions, d_m, is

$$d_m = (\boldsymbol{\mu}_F - \boldsymbol{\mu}_{F'})^T (C_F + C_{F'})^{-1} (\boldsymbol{\mu}_F - \boldsymbol{\mu}_{F'}). \quad (5)$$

d_m has a χ^2 distribution with two degrees of freedom. Looking at a χ^2 table, it is possible to check whether two distributions can be integrated with certain confidence by setting an appropriate threshold on d_m.

Let assume that the observation direction is limited to the direction from the robot toward T. If we check whether each view is appropriate by investigating the worst case, checking four cases ($F'_A \sim F'_D$) in figure 10 is sufficient for this purpose. In case of F'_A, for example, F'_A lies on the line determined by T and the left lens center. For each possible position of F'_A, we first compute the distribution of F'_A considering quantization error. We then compute Mahalanobis distance between this distribution and the distribution of F which has been already calculated in the first observation. We repeat this computation for all possible positions of F' and takes the minimum value. If the minimum value throughout the four cases is larger than some threshold, F and F' cannot be integrated, and consequently this second observation position is appropriate for resolving ambiguity.

VI. Experimental Results

Figure 11 shows the experimental environment. A robot is going to $destination$. There are two boxes in front of the robot. If the distance between boxes is large enough, the robot can take a short path to $destination$. Otherwise, the robot must take a longer path through the passageway. The problem is to find the optimal sequence of observation and motion operations which minimizes the expectation of the total cost. Figure 6 is a pair of stereo images taken at the current position. The robot makes the plan based on this pair of images.

The robot searched the optimal sequence of observation points in the $Search\,Area$ shown in figure 11. To reduce computation, observation points were limited to only grid points set on $Search\,Area$ and discretized distributions were used. In addition, the following assumptions were made: Time for estimating the distance between boxes once is constant. Uncertainty of motion is negligible. Time for motion is proportional to the distance. In the outside of $Search\,Area$, the costs for moves are predetermined. The robot passes the center between boxes, if possible. The paths of the robot consist of straight-line segments.

Figure 12 shows the calculated optimal plan. Before searching an optimal plan, the robot computed the areas where observations may not be able to resolve ambiguity of multiple correspondence or where observations may cause occlusion of the feature of the one box by the that of another. These areas, drawn in figure 12, were excluded from the actual search area. In the figure, planned paths of the robot are indicated by lines and observation points are indicated by branches. According to this plan, the robot moves as follows. The robot observes only once at the indicated point and decides on the final motion based on the information acquired at that point. When the true positions are L_1 and R_1, the robot passes between boxes if the distance between them are large enough and otherwise, the robot will take a detour through passageway. In other three cases, that is, in the cases that the combination of true positions is $L_1 - R_2$, $L_2 - R_1$ or $L_2 - R_2$, the robot can pass between boxes. While there are five possible paths in the plan, it is undecided in advance which path the robot will actually take.

VII. Conclusions and Discussions

This paper describes a new framework of planning of vision and motion for a mobile robot under uncertainty. We proposed a method of predicting sensor information and formulated the planning problem in a recurrence formula. As an example of sensor modeling, we described a model of the uncertainty of stereo vision in which correspondence of wrong pairs of features as well as quantization errors are considered. Using our framework, a robot can successfully make a plan for a real world problem.

Our current implementation requires much computation even for a simple problem. We have to introduce methods of reducing computation such as:

- To divide the problem into subproblems and to apply our method to each subproblems *or*
- To classify calculated plans for many situations into patterns to generate planning rules.

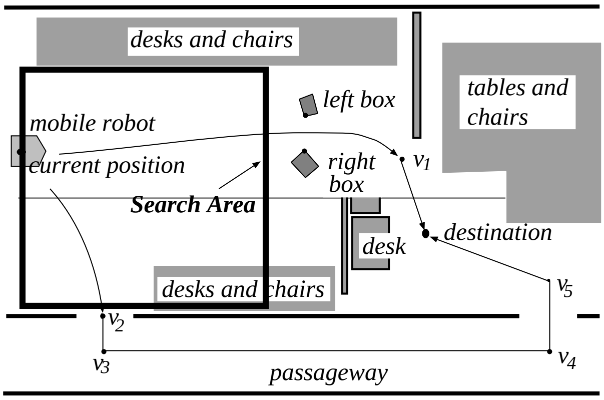

Fig. 11: Experimental environment.

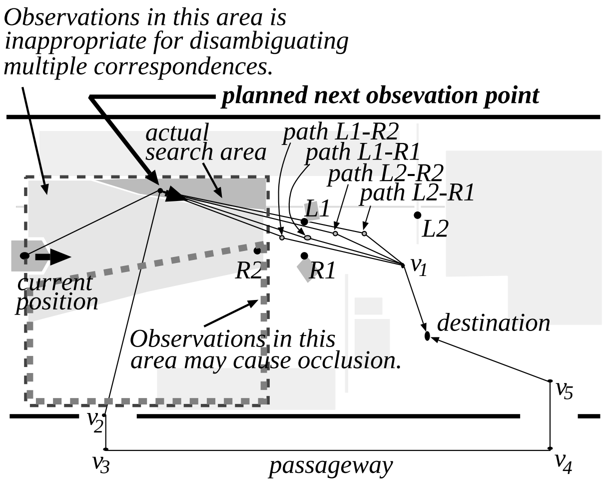

Fig. 12: The final plan.

References

[1] Ayache, N. and Faugeras, O.D., "Maintaining Representations of the Environment of a Mobile Robot," *IEEE Trans. Robotics and Automat.*, vol.5, no.6, pp.804-819 (1989).

[2] Brooks, R.A. and Lozano-Pérez, T., "A Subdivision Algorithm in Configuration Space for Findpath with Rotation," *IEEE Trans. Systems Man Cybanet.*, vol.15, pp.224-233 (1985).

[3] Cameron, A. and Durrant-Whyte, H., "A Bayesian Approach to Optimal Sensor Placement," *Int'l J. Robotics Res.*, vol.9, pp.70-88 (1990).

[4] Dean, T., Basye, K., Chekaluk, R., Hyun, S., Lejter, M. and Randazza, M., "Coping with Uncertainty in a Control System for Navigation and Exploration," *AAAI-90*, pp.1010-1015 (1990).

[5] Hutchinson, S.A and Kak, A.C., "Planning Sensing Strategies in a Robot Work Cell with Multi-Sensor Capabilities," *IEEE Trans. Robotics and Automat.*, RA-5, no.6, pp.765-783 (1989).

[6] Kriegman, D.J., Triendl, E. and Binford, T.O., "Stereo Vision and Navigation in Buildings for Mobile Robots," *IEEE Trans. Robotics and Automat.*, vol.5, no.6, pp.792-803 (1989).

[7] Matthies, L. and Shafer, S.A., "Error Modeling in Stereo Navigation," *IEEE J. Robotics and Automat.*, vol.RA-3, no.3, pp.239-248 (1987).

[8] Miura, J. and Shirai, Y, "Planning of Vision and Motion for a Mobile Robot Using a Probabilistic Model of Uncertainty," *1991 IEEE Int'l Workshop on Intelligent Robots and Systems*, (1991).

[9] Noborio, H., Naniwa, T. and Arimoto, S., "A Feasible Motion-Planning Algorithm Using the Quadtree Representation," *1988 IEEE Int'l Workshop on Intelligent Robots and Systems*, pp.769-774 (1988).

[10] Shirai, Y. and Nishimoto, Y., "A Stereo Method Using Disparity Histograms of Multi-Resolution Channels," *Robotics Research 3*, The MIT Press, pp.27-32 (1986).