Alec Cameron
Philips Laboratories
North American Philips Corporation
Briarcliff Manor, New York 10510

Hugh Durrant-Whyte
Robotics Research Group, Department of Engineering Science
University of Oxford
Oxford OX1 3PJ, United Kingdom

A Bayesian Approach to Optimal Sensor Placement

Abstract

By "intelligently" locating a sensor with respect to its environment, it is possible to minimize the number of sensing operations required to perform many tasks. This is particularly important for sensing media, such as tactile sensors and sonar, that provide only "sparse" data. In this paper, a system is described that uses the principles of statistical decision theory to determine the optimal sensing locations for performing recognition and localization operations. The system uses a Bayesian approach to utilize any prior object information (including object models or previously acquired sensory data) in choosing the sensing locations.

1. Introduction

An important feature of many sensing systems is that only a small portion of the object of interest can be sensed at any instant, and hence it is necessary to relocate the sensor to build up a model of the object. Typical systems include tactile sensing, sonar, and range sensing systems. Several researchers (Browse 1987; Gaston and Lozano-Pérez 1984; Grimson and Lozano-Pérez 1984) have addressed the problem of fusing the sparse data provided by such systems and using it to perform tasks such as object recognition, localization, and path planning. Although it is clear that judicious placement of the sensor will reduce the number of sensing operations required and the quantity of data to be processed, the strategy of where optimally to place the sensor has not been thoroughly investigated. Some researchers (Grimson 1986; Ellis 1987; Schneiter 1986) have shown how guided sensing can be used to discriminate between a finite number of admissible interpretations following previously random sensor placement, but the requirements of the early sensing stages, where there may be infinitely many interpretations of the sensory data, were not considered. In this paper, a system is described that addresses the problem of where to locate the sensor to obtain the most useful (maximally constraining) data for the model of the sensed object. The solution proposed has more general application in the domain of problems involving directed search.

In developing "intelligent" sensing strategies, a statistical decision theoretic approach is used (see Berger 1980; Blackwell and Girshick 1954; DeGroot 1970). Using prior knowledge of the set of admissible objects and data from one or more sensed points, an interpretation discrimination function for the work space is developed, and the point that maximizes this function is chosen as the next sensing location. The object models and sensed data may be probabilistic functions using the approach of Durrant-Whyte (1985).

The system described in this paper is flexible enough to be able to take advantage of any special capabilities of the sensor used. If, for example, the sensor is capable of detecting specific geometric features (such as edges and corners), then the system can utilize the saliency of such features for object recognition or localization and ascribe to them a suitably high priority when choosing optimal sensing locations.

This work was performed while A.C. was a member of the Robotics Research Group at the University of Oxford supported by a Rhodes Scholarship.

The International Journal of Robotics Research,
Vol. 9, No. 5, October 1990,

The system also uses all sensory data to update its expectations regarding the operating environment. The approach used is a probabilistic form of the interpretation tree approach used in many artificial intelligence algorithms, including Grimson and Lozano-Pérez (1984), which assigns an expectation to every branch of the tree. This enables the system to order any admissible interpretations of the sensory data at any stage of the sensing process and to ascribe confidence levels to any conclusions that are reached.

In this paper, only two-dimensional domains are addressed; however, the same techniques can theoretically be generalized to three-dimensional environments (Cameron 1989).

Section 2 provides a brief description of statistical decision theory and introduces the concepts of prior information and the utility function. The Bayes decision principle, which is pursued in this paper, is also presented.

In section 3, it is shown how information about an object's shape can be combined with sensory data to produce a probabilistic membership function on the work space. This function represents the knowledge of the sensed environment and can be used as a prior information function in the sensing strategy. Separate functions can be generated for specific geometric features to enable different utilities to be applied to different sensing outcomes.

Utility functions on the work space are derived in section 4. These functions qualitatively describe the constraining value of obtaining sensory data at each location in the environment. The membership function can be used together with the utility function to make a Bayes decision about the "best" location at which to sense in the work space.

In section 5, the techniques of sequential analysis (see Wald 1947) are used to process new sensory information as it is acquired. The new data determine whether a decision can be made (within certain error bounds) or whether data from a new sensory location are required. This section also explains how new data are used to update current expectations and describes the probabilistic interpretation tree approach.

The final two sections describe an implementation of the principles described in this paper and some suggested refinements and generalizations of the system.

2. Bayesian Decision Theory

Statistical decision theory formalizes the problem of taking optimal actions with uncertain information. The basic elements of statistical decision theory include the state of the environment (which is only partially known), and a choice of decisions or actions, the outcomes of which will depend on the current state. These outcomes can be quantified by a utility function $U(\theta, a)$ which is a function of the state θ and the action a. Usually the aim is to choose the action a so as to maximize the utility function $U(\theta, a)$.

The method by which the action is chosen is called the decision principle. The decision principle pursued in this paper is the Bayes principle.[1] This principle relies heavily on any prior information that is possessed about the state of nature. The prior information is represented as a probability distribution on possible states, $\pi(\theta)$.

This leads to a quantity called the Bayes risk of a decision given by

$$r(\pi, a) = E^{\pi}[U(\theta, a)]$$

The action that maximizes $r(\pi, a)$ (over all a) is called the Bayes decision.

To apply these concepts to the sensor placement problem, it is necessary to identify the elements θ, a, $U(\theta, a)$, and $\pi(\theta)$. For this problem, θ represents the space occupied by the sensed object (which is only partially known), and a is the location of the next sensed point. At any given location the object is either present or not, and hence θ can be represented by a two-valued function on the work space corresponding to detection of the object or failure to detect. [Such a function is also known as the characteristic function of the object (Horn 1986).]

The optimal value of a is sought to maximally constrain the object model. Hence the quantity $U(\theta, a)$ must reflect the constraining power acquired by sensing the state θ at the location a. The current expectation of sensing the object at a particular location, based

1. For a discussion of the merits of the Bayesian approach, see Berger (1980).

on any knowledge of the object's shape and the location of previously sensed points, is described by the probability distribution function $\pi(\theta)$.

In the next section, a probabilistic membership function on the work space is developed that describes the likelihood of sensing the object at a given location. Hence this function fills the role of $\pi(\theta)$ in decision theoretic terminology. The utility functions developed in section 4 are functions of the location of the point of interest (the action a) and the presence or absence of the object at that location (the two-valued state θ).

3. Probabilistic Membership Functions

Once a sensor has made contact with an object, some knowledge of the location of the object has been acquired. In particular, if a recognizable feature, such as an edge, has been detected, a powerful constraint on the object's position has been obtained that can be further constrained by sensory information at other locations. Most importantly, once the first sensed point has been obtained, the knowledge of the shape of the object can be used to guide the subsequent search for further contact points. In this section, it is shown how knowledge of the object's shape, in conjunction with the location of a sensed point, can be used to determine the probability of sensing the object at each point in the sensing environment. The function describing these probabilities will be termed the probabilistic membership function (PMF).

The methods described in this section lend themselves readily to automatic generation and updating of the PMF for any given polygonal shape, as was performed for the implementation described in section 6.

3.1. Generating PMFs

In a two-dimensional domain, the model of an object o may be described by a membership (or characteristic) function $O(x, y)$ (associated with a "home" placement for the object, usually at the origin of the coordinate system). If the shape of o is completely known, $O(x, y)$ is defined by

$$O(x, y) = \begin{cases} 1 & \text{if } (x, y) \in o \\ 0 & \text{otherwise} \end{cases}$$

If the object shape is not completely known, a value can be assigned to $O(x, y)$ that represents the confidence that $(x, y) \in o$, subject to the conditions

$$0 \leq O(x, y) \leq 1 \quad \forall (x, y)$$

and

$$\int \int O(x, y)\, dx\, dy = \text{area of } o \text{ (or its expected value).}$$

The accuracy of the data supplied by the local sensor can also be modeled. The function $S(x, y)$ will be used to represent the uncertainty implicit in any data returned by the sensor. $S(x, y)$ is a distribution function and satisfies the conditions

$$S(x, y) \geq 0 \quad \forall (x, y)$$

and

$$\int \int S(x, y)\, dx\, dy = 1.$$

For a perfect sensor (one that returns data with no implicit uncertainty), $S(x, y) = \delta(x, y)$. All data returned by the sensor should be convolved with $S(x, y)$ to correctly take account of the inherent limitations of the sensor.

In addition to uncertainties introduced by the sensor, any attempts to detect accurately the sensed object are also affected by uncertainty in the relative position of the sensor and object, which is dependent on the shape of the sensed feature. For example, if a corner is sensed, the location of the object may be completely determined. However, if an edge is detected without one of its end points, there will be uncertainty in the object's position in the direction parallel to the edge. If a point within the boundaries of the object is sensed without detecting an edge (or a point on a planar face of a three-dimensional object), there will be uncertainty in three degrees of freedom (two translational and one rotational). This "feature uncertainty" can be represented by a distribution function $F(x, y)$ satisfying

$$F(x, y) \geq 0 \quad \forall(x, y)$$

and

$$\int\int F(x, y)\, dx\, dy = 1.$$

For the cases of corner and edge mentioned above, $F(x, y)$ will, in general, be represented by a delta function and a uniform distribution in one variable, respectively.

Once contact has been established with a particular feature of the object, the object model can be transformed to align with the corresponding feature on the model with the sensed point. This gives rise to a new function $O'(x, y)$ describing the transformed model. To incorporate the uncertainty in the object's location, the model must be convolved with the sensor and feature uncertainty functions, producing the PMF, $\Pi(x, y)$.[2]

$$\Pi(x, y) = O'(x, y) * F(x, y) * S(x, y)$$

$\Pi(x, y)$ has the same qualities as $O(x, y)$ (see appendix A), namely

$$0 \leq \Pi(x, y) \leq 1 \quad \forall(x, y)$$

and

$$\int\int \Pi\, dx\, dy = \text{area of } o.$$

In addition, $\Pi(x, y)$ represents the probability of the location (x, y) lying within the boundary of the object, taking into account uncertainties in the sensed data and object model.

2. Throughout this paper the symbol * will be used to represent the convolution integral

$$a(x, y) * b(x, y) = \int\int a(x-\mu, y-\eta)\, b(\mu, \eta)\, d\mu\, d\eta$$

Some examples follow, showing how the PMF can be generated from the membership function of an object.

Example 1: A Symmetric Object

Consider a square of side length l. The object model for such a shape can be defined as[3]

$$O(x, y) = \begin{cases} 1 & \text{if } 0 \leq x \leq l \text{ and } 0 \leq y \leq l \\ 0 & \text{otherwise} \end{cases}$$

It will be assumed that the sensor has no implicit uncertainty, so

$$S(x, y) = \delta(x, y).$$

Further assume that contact is made with an edge at the point (x_0, y_0) with the edge lying parallel to the x-axis of the coordinate system. As the object is completely symmetric, the sensed edge can be associated with any edge in the model. For simplicity, it will be associated with the model edge whose end points are $(0, 0)$ and $(l, 0)$.

Furthermore, the sensed point can be associated with any point lying on the chosen edge in the model. This uncertainty in relative position of the sensed point is described by the feature uncertainty function. Assume the sensed point (x_0, y_0) has been associated with the midpoint of the edge in the model, the point $(l/2, 0)$. The resulting transformed model function would be

$$O'(x, y) = O(x - x_0 + l/2, y - y_0).$$

The actual position of the sensed point could with equal probability be located anywhere within a range of values $\pm l/2$ of this location.

The feature uncertainty function, $F(x, y)$, could be described by a uniform distribution of its x-coordinate $U_{(-l/2,l/2)}(x)$, as shown in Figure 1.

Alternatively, the sensed point may be associated with one of the end points of the modeled edge, say the

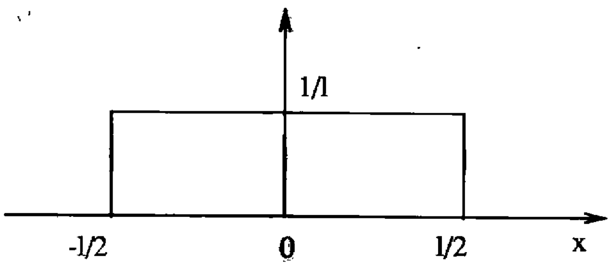
3. The specification of $O(x, y)$ is unique only up to arbitrary translations and rotations of the model.

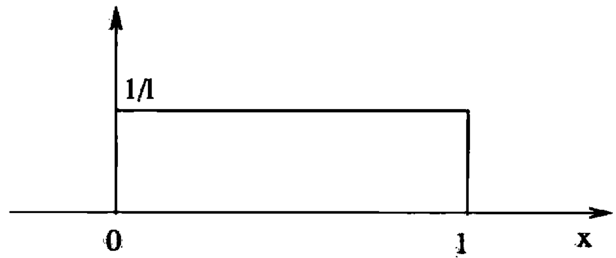
Fig. 1. Feature uncertainty function.

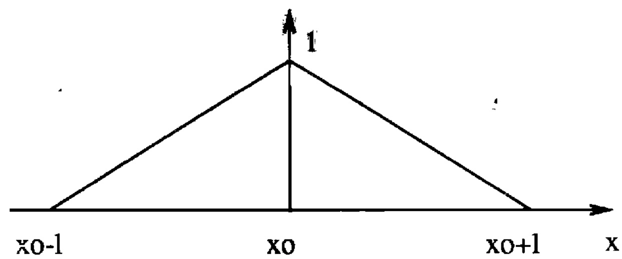
Fig. 2. Alternative feature uncertainty function.

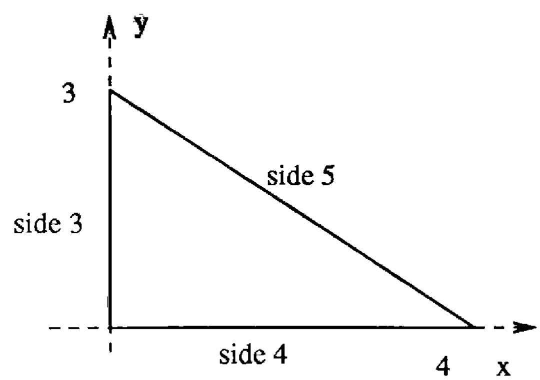
Fig. 3. Probabilistic membership function.

Fig. 4. Object model (or membership function).









point (0, 0). This would give rise to a different transformed model function

$$O'(x, y) = O(x - x_0, y - y_0).$$

In this case, however, the uncertainty in the x-value is within the range 0 to $+l$ of the chosen value, resulting in a different feature uncertainty function, as shown in Figure 2.

No matter which point is used, the resulting function $\Pi(x, y)$ is unchanged.

$$\begin{aligned}
\Pi &= O'(x, y) * F(x, y) * S(x, y) \\
&= O'(x, y) * F(x, y) \\
&= \int \int O'(x - \mu, y - \eta)\, F(\mu, \eta)\, d\mu\, d\eta \\
&= \int O'(x - \mu, y)\, F(\mu, 0)\, d\mu \\
&= 1/l \int_0^l O(x - x_0 - \mu, y - y_0)\, d\mu \\
&= 1/l \int_{-l/2}^{l/2} O(x - x_0 + l/2 - \mu, y - y_0)\, d\mu \\
&= \begin{cases} 0 & \text{if } y < y_0,\ y > y_0 + l,\ x < x_0 - l, \text{ or } x > x_0 + l \\ 1 - \dfrac{x_0 - x}{l} & \text{if } y_0 \le y \le y_0 + l \text{ and } x_0 - l \le x \le x_0 \\ 1 - \dfrac{x - x_0}{l} & \text{if } y_0 \le y \le y_0 + l \text{ and } x_0 \le x \le x_0 + l \end{cases}
\end{aligned}$$

A plot of $\Pi(x, y_1)$ for $y_0 \le y_1 \le y_0 + l$ is given in Figure 3.

As would be expected, the probability of sensing the object again at (x_0, y_1) is one (assuming a perfect sensor). The probability decreases with distance from x_0 until it becomes zero, when the distance from the sensed point is greater than the side length, l. Thus the results obtained concur with general expectations.

Example 2: An Asymmetric Object

Consider a right triangle with sides of length 3, 4, and 5 units, respectively. A possible "home" location is shown in Figure 4 for the object, giving rise to the object model

$$O(x, y) = \begin{cases} 1 & \text{if } 0 \le y \le 3 \text{ and } 0 \le x \le 4(1 - y/3) \\ 0 & \text{otherwise.} \end{cases}$$

Fig. 5. Probabilistic membership function for side *4.*

Fig. 6. Object model for contact on side *3.*

Assume again that $S(x, y) = \delta(x, y)$, and that contact is made with an edge at the point (x_0, y_0) with the edge lying parallel to the x-axis of the coordinate system. The form of $O'(x, y)$ and $F(x, y)$ will depend on which edge is associated with the sensed point, and hence the three edges must be treated as separate cases.

Consider first associating the sensed point with the side of length 4 (termed *side* 4). Further, the sensed point shall be associated with the leftmost point of the edge (although it was seen in the previous example that any point can be chosen by an appropriate formulation of the feature uncertainty function).

Hence $O'(x, y)$ can be calculated.

$$O'(x, y) = \begin{cases} 1 & \text{if } y_0 \le y \le y_0 + 3 \text{ and } x_0 \le x \le x_0 + 4\left(1 - \dfrac{y - y_0}{3}\right) \\ & \text{otherwise.} \end{cases}$$

The feature uncertainty function is given by

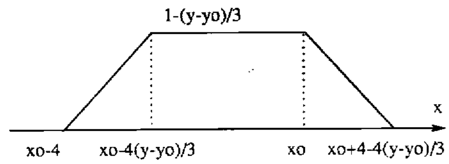
$$F_{side4}(x, y) = \begin{cases} U_{(0,4)}(x) & \text{if } y = 0 \\ 0 & \text{otherwise.} \end{cases}$$

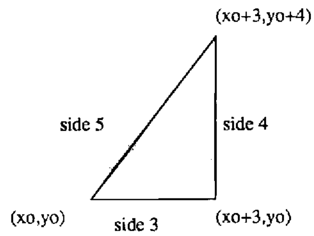
resulting in the PMF

$$\begin{aligned} \Pi_{side4}&(x, y) \\ &= O'(x, y) * F_{side4}(x, y) * S(x, y) \\ &= O'(x, y) * F_{side4}(x, y) \\ &= \int\int O'(x - \mu, y - \eta)\, F(\mu, \eta)\, d\mu\, d\eta \\ &= \int_0^4 1/4 O'(x - \mu, y)\, d\mu \end{aligned}$$

$$= \begin{cases} 0 & \text{if } y \le y_0,\ y \ge y_0 + 3,\ x \le x_0 - 4, \text{ or } x \ge x_0 + 4\left(1 - \dfrac{y - y_0}{3}\right) \\ 1 + \dfrac{x - x_0}{4} & \text{if } y_0 \le y \le y_0 + 3 \text{ and } x_0 - 4 \le x \le x_0 - 4\dfrac{y - y_0}{3} \\ 1 - \dfrac{y - y_0}{3} & \text{if } y_0 \le y \le y_0 + 3 \text{ and } x_0 - 4\dfrac{y - y_0}{3} \le x \le x_0 \\ 1 - \dfrac{y - y_0}{3} - \dfrac{x - x_0}{4} & \text{if } y_0 \le y \le y_0 + 3 \text{ and } x_0 \le x \le x_0 + 4 - 4\dfrac{y - y_0}{3} \end{cases}$$

This function is plotted in Figure 5.





Similar calculations associating the sensed point with the other two sides can be performed. For the side of length 3 (termed *side* 3), the tranformed model is shown in Figure 6, and its describing equations is

$$O'(x, y) = \begin{cases} 1 & \text{if } y_0 \le y \le y_0 + 4 \text{ and } \\ & x_0 + 3\dfrac{y - y_0}{4} \le x \le x_0 + 3 \\ 0 & \text{otherwise.} \end{cases}$$

The feature uncertainty function is

$$F_{side3}(x, y) = \begin{cases} U_{(0,\,3)}(x) & \text{if } y = 0 \\ 0 & \text{otherwise} \end{cases}$$

resulting in the PMF

$$\Pi_{side3}(x, y) = \int_0^3 1/3 O'(x - \mu, y)\, d\mu$$

$$= \begin{cases} 0 & \text{if } y \le y_0,\ y \ge y_0 + 4,\ x \le x_0 - 3\left(1 - \dfrac{y - y_0}{4}\right),\ x \ge x_0 + 3 \\ 1 - \dfrac{y - y_0}{4} + \dfrac{x - x_0}{3} & \text{if } y_0 \le y \le y_0 + 4 \text{ and } x_0 - 3\left(1 - \dfrac{y - y_0}{4}\right) \le x \le x_0 \\ 1 - \dfrac{y - y_0}{4} & \text{if } y_0 \le y \le y_0 + 4 \text{ and } x_0 \le x \le x_0 + 3\dfrac{y - y_0}{4} \\ 1 - \dfrac{x - x_0}{3} & \text{if } y_0 \le y \le y_0 + 4 \text{ and } x_0 + 3\dfrac{y - y_0}{4} \le x \le x_0 + 3 \end{cases}$$

This function is plotted in Figure 7.

For the side of length 5 (see Figure 8),

$$O'(x, y) = \begin{cases} 1 & \text{if } y_0 \le y \le y_0 + 2.4 \text{ and} \\ & x_0 + 3.2\dfrac{y - y_0}{2.4} \le x \le x_0 + 5 - 1.8\dfrac{y - y_0}{2.4} \\ 0 & \text{otherwise.} \end{cases}$$

The feature uncertainty function is

$$F_{side5}(x, y) = \begin{cases} U_{(0,5)}(x) & \text{if } y = 0 \\ 0 & \text{otherwise,} \end{cases}$$

resulting in the PMF

$$\Pi_{side5}(x, y) = \int_0^5 1/5\, O'(x - \mu, y)\, d\mu$$

$$= \begin{cases} 0 & \text{if } y \le y_0,\ y \ge y_0 + 2.4, \\ & x \le x_0 - 5 + 3.2\dfrac{y - y_0}{2.4}, \text{ or} \\ & x \ge x_0 + 5 - 1.8\dfrac{y - y_0}{2.4} \\ 1 + \dfrac{x - x_0}{5} - 3.2\dfrac{y - y_0}{12} & \text{if } y_0 \le y \le y_0 + 2.4 \text{ and} \\ & x_0 - 5 + 3.2\dfrac{y - y_0}{2.4} \le x \\ & \le x_0 - 1.8\dfrac{y - y_0}{2.4} \\ 1 - \dfrac{y - y_0}{2.4} & \text{if } y_0 \le y \le y_0 + 2.4 \text{ and} \\ & x_0 - 1.8\dfrac{y - y_0}{2.4} \le x \\ & \le x_0 + \dfrac{y - y_0}{2.4} \\ 1 - \dfrac{x - x_0}{5} - 1.8\dfrac{y - y_0}{12} & \text{if } y_0 \le y \le y_0 + 2.4 \text{ and} \\ & x_0 + \dfrac{y - y_0}{2.4} \le x \\ & \le x_0 + 5 - 1.8\dfrac{y - y_0}{2.4} \end{cases}$$

This function is plotted in Figure 9.

To obtain the total probability of the object being present at a particular location in the work space without any knowledge of which edge has been contacted, the functions obtained for the three sides must be combined. The total probability is evaluated as follows:

$$\begin{aligned} \Pi_{total}(x, y) = {} & P(side\ 3)\Pi_{side3}(x, y) \\ & + P(side\ 4)\Pi_{side4}(x, y) \\ & + P(side\ 5)\Pi_{side5}(x, y). \end{aligned}$$

To evaluate the probability of the sensed side having a particular identity, in the absence of any other information, one approach is to assume that this probability is proportional to the length of the side. Thus the probability for each side is the ratio of the side length to the perimeter of the object. For the triangle considered in this example,

$$P(side\ 3) = 1/4,\ P(side\ 4) = 1/3,\ P(side\ 5) = 5/12.$$

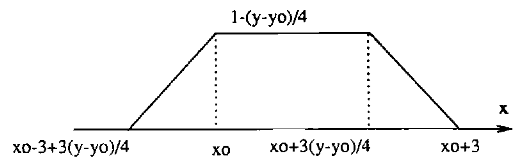


Fig. 7. Probabilistic membership function for side *3.*

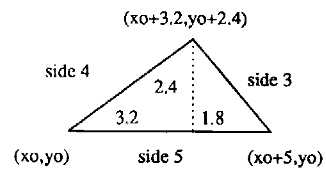


Fig. 8. Object model for contact on side *5.*

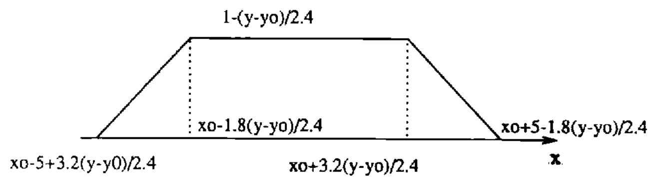


Fig. 9. Probabilistic membership function for side *5.*

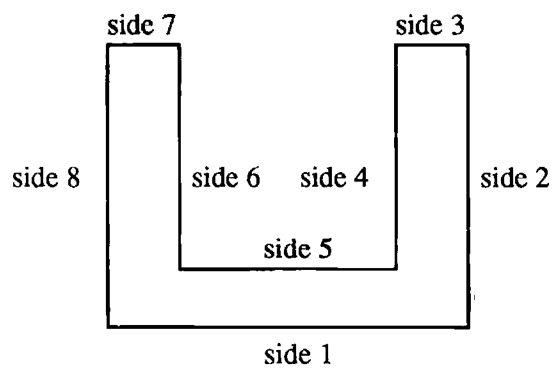


Fig. 10. Object characteristic function.

Fig. 11. PMF for side *1.*

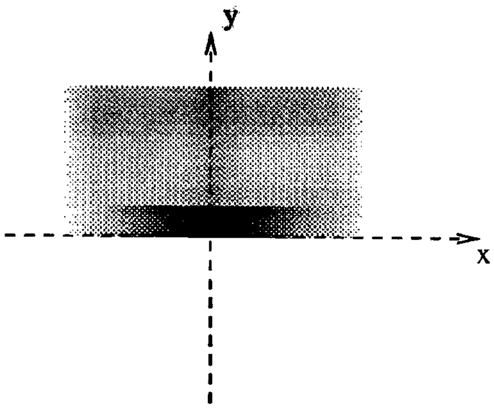


Fig. 12. PMF for side *2.*

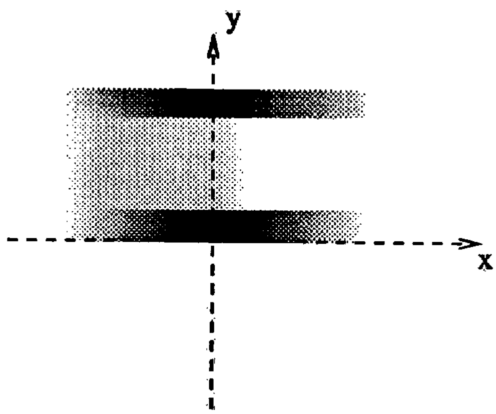


Fig. 13. PMF for side *3.*

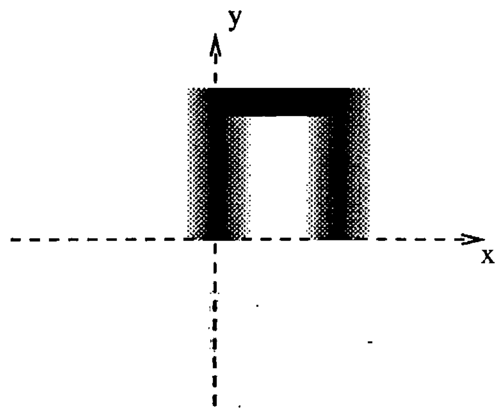
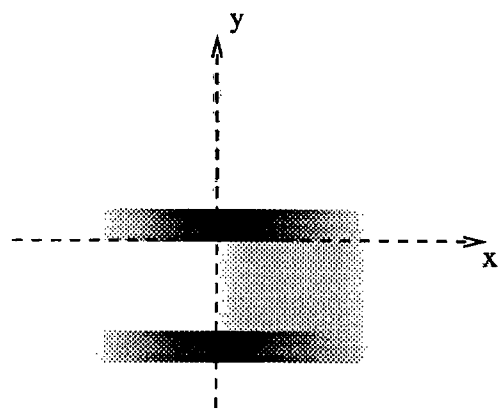


Fig. 14. PMF for side *4.*



Example 3: Automatic Generation for a Concave Object

This example considers the PMF for a more complex object. Rather than deriving analytically an expression for this function, automatically generated plots are displayed showing the PMFs for each contact side.

Consider the concave object shown in Figure 10. It has eight sides, and as there is no rotational symmetry for the object, each must be considered separately in generating the PMF.

The PMFs generated automatically for each side are shown in Figures 11 to 18. The plots use a gray scale to indicate likelihood of detection, with black indicating certain detection (PMF value of one), and white indicating no prospect of detection (PMF value of zero). The overall PMF (obtained by obtaining a sum of the side PMFs, each weighed according to the length of the corresponding side), is shown in Figure 19.

The figures demonstrate how the PMFs explicitly represent the shape of the object, blurred by uncertainty along the designated contact edge.

The automatic generation of the PMF is readily performed from a list of vertex coordinates representing the object model. As the objects become more complex, the amount of processing and data storage required to compute the PMF increases linearly with the number of sides (i.e., the algorithm used to generate the PMF automatically is only $O(n)$ complex with respect to the number of sides n).

3.2. Updating PMFs

As has been shown in the previous subsection, the only data required to construct a PMF are a model of

Fig. 15. PMF for side *5.*

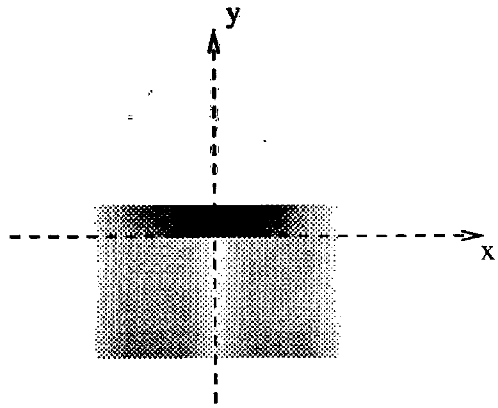
Fig. 16. PMF for side *6.*

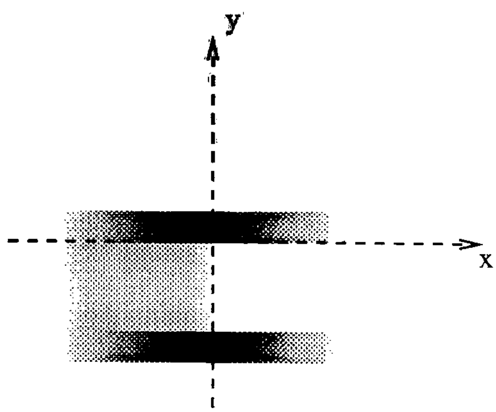
Fig. 17. PMF for side *7.*

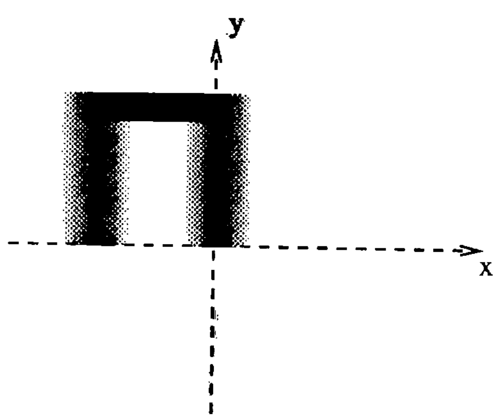
Fig. 18. PMF for side *8.*

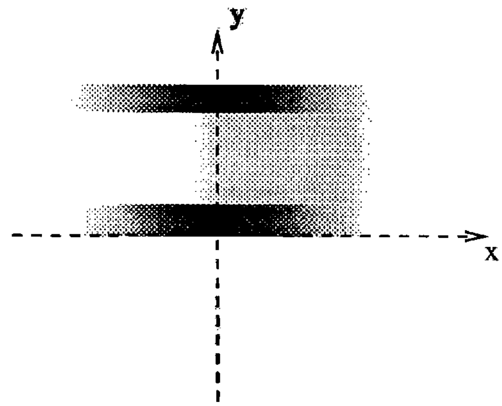
Fig. 19. Resultant PMF.

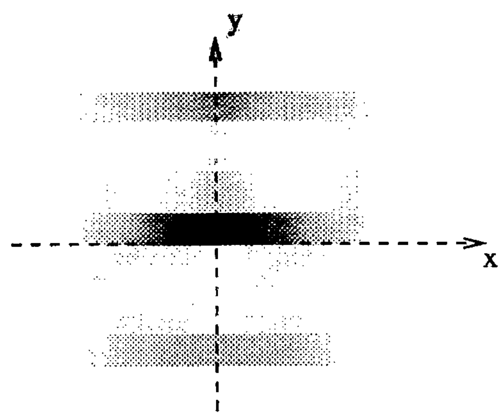
Fig. 20. Probabilistic membership functions.

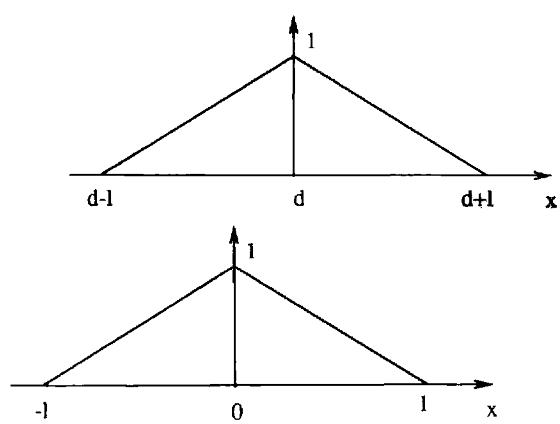












the sensed object and the location of a sensed object feature. However, as data from more sensed points become available, they must be used to update the PMF so that it accurately represents the knowledge of the work space.

One appealing solution to this problem would be to generate a PMF for each sensed point and combine them. In practice, however, this combination proves to be nontrivial and requires knowledge of the object's features.

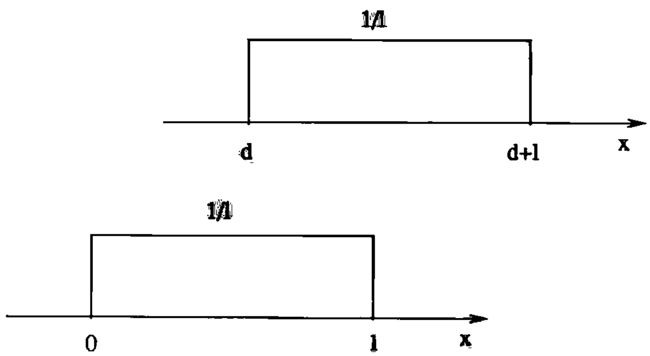


Fig. 21. Feature uncertainty functions.

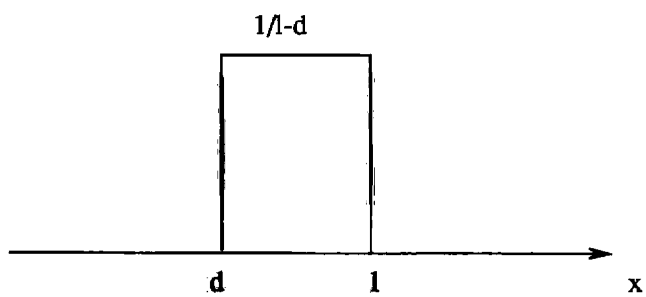


Fig. 22. Updated feature uncertainty function.

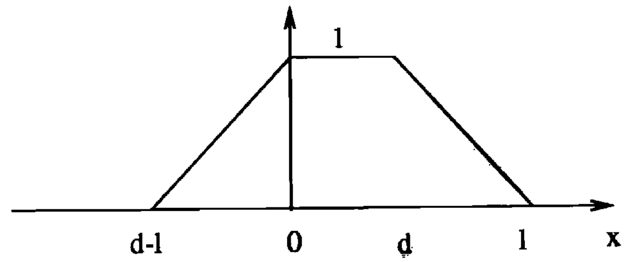


Fig. 23. Updated probabilistic membership function.

Consider the case of a square of side length l whose edge has been sensed at two locations, separated by a distance d. The resulting PMFs can be represented graphically as shown in Figure 20.

To determine the appropriate PMF that represents the information provided by both sensed points, the feature uncertainty functions for the two locations relative to the sensed point at (x_0, y_0) must be considered (Fig. 21).

Observing these functions, it can be deduced that the combined uncertainty in the feature's position is represented by the uniform distribution $U_{(d,l)}(x)$, shown in Figure 22.

Convolving this function with the object model function gives the PMF shown in Figure 23.

This function cannot be readily obtained from the two single-point PMFs shown earlier. It is also necessary to reconstruct the feature uncertainty functions to take account of locations at which no feature is detected (and hence have no associated PMF, although they constrain the object's position) and also observations of different features, such as interiors (which do not provide strong constraints on the object location by themselves, but do when combined with a sensed edge).

A robust method is required for integrating information provided by further sensed points to update the feature uncertainty function.

Consider again the scenario where the initial sensed point was an edge contact. Thus the feature uncertainty function is as shown in Figure 2. In being convolved with this function, the object model $O'(x - \mu, y)$ is integrated with respect to μ as μ varies in the (acceptable) range $(0, l)$. Consider the location (x_1, y_1). Now, unless $\Pi(x_1, y_1)$ is zero or one, then the function $O'(x_1 - \mu, y_1)$ for $\mu \in (0, l)$ will take the value zero for some subset of values of μ and the value one for the remainder. Once sensor data have been acquired at (x_1, y_1) and the actual value of $\Pi(x_1, y_1)$ determined, those values of μ that produced the wrong value can be rejected. A uniform distribution on the (remaining) acceptable values of μ can be assumed, producing the new feature uncertainty function.

3.3. Effect of Finite Sensor Size

All the analysis carried out thus far has assumed that it is possible to obtain sensory information at a given point in the work space. Because of the finite sensing area of all sensors, however, this is never true in practice, nor is it always desirable. Often by sensing over a finite area in a single action, more information can be acquired, enabling characteristic point features, such as corners, to be detected, which would not be possible with a point sensor. As corners are often salient features of an object set (as will be discussed in section 4.2), they can play an important role in the sensor placement strategy. It will be clear that in these in-

Fig. 24. Sensor data and possible object locations.

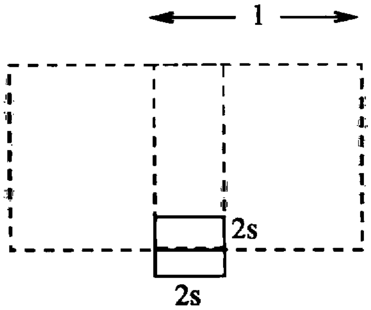
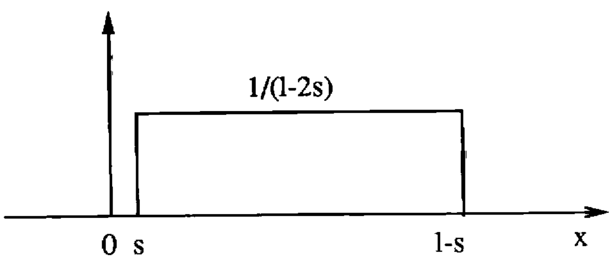


Fig. 25. Feature uncertainty function.



stances, uncertainty is reduced (and hence sensing power increased) as the sensor scope increases.

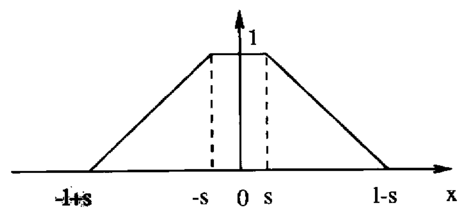
By sensing over a finite area, the sensing is effectively being performed at many adjacent points simultaneously. This gives rise to a different PMF than is obtained by sensing at a single point. It was seen in the previous subsection that sensor data from different sensor locations combine to alter the feature uncertainty function, F. It is to be expected that the feature uncertainty function due to sensing over a finite area is the same as that due to sensing at point locations spread over the same sensing area. The finite size of the sensor does not affect the object model, O, or the sensor uncertainty function, S.[4] Hence the probabilistic membership function, Π, can be computed as before

$$\Pi = O' * F * S,$$

where F is the only quantity that has altered.

4. It may be expected that altering a property of the sensor, such as its size, would have an effect on the function S. S, however, is the sensor uncertainty function and models the accuracy of (or the confidence in) the sensor data. F, on the other hand, is the feature uncertainty function and represents the uncertainty in the location of the model due to the properties of the sensed feature and the manner in which it is sensed.

Fig. 26. Probabilistic membership function.



Example

In this example, a sensor with a square sensing area of small, but finite, extent is considered.

Assume that the edge of an object of length l has been sensed by a sensor of dimensions $2s$ by $2s$. The sensor output is as shown in Figure 24, together with possible locations of the object. This leads to the feature uncertainty function along the dimension parallel to the sensed edge as shown in Figure 25. Assuming no sensor uncertainty [$S(x, y) = \delta(x, y)$], the resulting PMF along this dimension is as plotted in Figure 26.

Note that the PMF obtained is equivalent to that obtained by sensing with a point sensor at points displaced by $\pm s$ from the midpoint of the sensor.

3.4. Feature PMFs

The PMFs discussed thus far have represented the probability of detecting the sensed object at a given location in the work space. If the sensor is capable of detecting features (such as edges or corners), then it may be useful to represent the probability of detecting a given feature at a given location in the work space. Thus it is possible to generate PMFs for each feature of interest and use this information in evaluating potential sensing locations.

4. Utility Functions

As was discussed in section 2, one of the vital elements of decision theory is a utility function that will enable the merits of different actions to be compared for a particular state of the sensing environment. For the tasks of object recognition or localization, the utility functions that are required must provide some measure of the ability of the sensing location to discriminate among various object identities or locations.

It will be assumed that the state of the environment can be described by a membership function on the work space representing the presence or absence of the object at each location. The prior knowledge of the membership function is the probabilistic membership function described in the previous section.

A suitable utility function will depend on the task to be performed. In the following subsections, some utility functions are generated for several recognition and localization problems.

Example 1: The Two-Object Set

Consider the task of identifying the sensed object from a set of two objects, $\{o_1, o_2\}$. The utility function is a function of two variables: the sensing location, (x, y) (the action), and whether the object is present at the sensing location (the state or outcome). The state can be described by a two-valued function: *detection* = 0 (object not present at the sensing location) and *detection* = 1 (object present).

If a location (x_0, y_0) exists at which detection of the object will enable its identity to be deduced as o_1, then the utility function should have a high value for (x_0, y_0) and *detection* = 1. Similarly, if a location (x_1, y_1) exists at which detection of the object implies its identity is not o_1, then the utility function should also be high for (x_1, y_1) and *detection* = 1. These two situations represent extreme values (i.e., 0 and 1) of the function

$$P(\text{object} = o_1 | detection = 1).$$

Failure to detect an object can be just as important and useful as detecting the object. If the object is not detected at a location where $P(object = o_1 | detection = 0)$ has an extreme value, then this information can be used to deduce the object's identity.

Consider the following function as a utility function.

$$\begin{aligned} U(state, (x, y)) &= |P(o_1|state)(x, y) - P(o_1)| \\ &\quad + |P(o_2|state)(x, y) - P(o_2)| \\ &= 2|P(o_1|state)(x, y) - P(o_1)| \end{aligned}$$

This function attains its maximum value when the a posteriori expectation of the object's identity being o_1 $[P(o_1|state)]$ differs greatest from the a priori expectation $[P(o_1)]$, and attains its minimum value of zero when the expectations are the same (no new information).

By expanding equation (1) using Bayes' Rule and weighting the function over the prior expectations on the states, the discrimination function can be expanded and simplified to obtain the result (see appendix B):

$$D(x, y) \propto |\Pi_{o_1}(x, y) - \Pi_{o_2}(x, y)|. \qquad (2)$$

Hence for this task, the optimal sensing location is the location at which the difference between the probabilistic membership functions is greatest. This location is easily determined and corresponds with an intuitive sense of the characteristics of the optimal sensing location.

If we sense at a location where there is a high probability of detecting the object if it is of a given identity, but a low probability if it is of the other identity, then we can expect the outcome to influence our expectations significantly. Clearly, if the object is detected, our expectation of it being the first object will be strengthened at the expense of the second object. Conversely, nondetection will favor our expectation of the second object.

On the other hand, we can expect to acquire little information at locations where we have a similar expectation of detection of both objects. There is no point in sensing at a location where detection is assured for both objects, as the expectations will not be altered by the sensory data. The same will be true at locations where nondetection is assured.

Example 2: The n*-Object Set*

Consider the task of identifying a sensed object from a set of n candidates $\{o_1, o_2, \ldots, o_n\}$. The utility function that is sought will be a function of the prospective sensing location and the (unknown) state at that location. An appropriate function is one that will return a high value for those values of location and state where the detection of the given state at the given location is sufficient to deduce the identity of the sensed object. More formally, high values of [*state*, (x, y)] would occur at locations where $P(o_k|state)$ was high for one of the objects, o_k.

One possible utility function is the function

$$U[state, (x, y)] = \sum_i |P(o_i|state) - P(o_i)|, \qquad (3)$$

where $P(o_i)$ represents the current expectation of the sensed object's identity being o_i. This function assigns a high value to sensing situations that drastically alter the expectations of the object's identity and no value to those situations that do not alter the expectations (as such locations provide no new information).

The Bayesian sensing location is that location that maximizes the discrimination function, obtained by determining the expected value of the utility function over the distribution of possible sensing states:

$$D(x, y) = \sum_j U[state_j, (x, y)]P(state_j). \qquad (4)$$

Substituting equation (3) in equation (4) gives

$$D = \sum_j \left[\sum_i |P(o_i|state_j) - P(o_i)|\right] P(state_j)$$
$$= \sum_i P(o_i) \left[\sum_j |P(state_j|o_i) - P(state_j)|\right].$$

For the case where the value *state* can take only two values, *detection* = 1 and *detection* = 0, the conditional probability $P(detection = 1|o_i)$ is simply the PMF Π_i described in the previous section. ($P(detection = 0|o_i)$ is also readily evaluated as $1 - \Pi_i$.)

The discrimination function can thus be evaluated at all possible sensing locations, and the location that maximizes its value can be chosen as the best next sensing location.

Example 3: Localizing an Object

Consider the problem of localizing a single object, whose shape is known, in the work space. This requires constraining the feature uncertainty function to a delta function.

Hence detection of (or failure to detect) the object at a location that significantly reduces the spread of the feature uncertainty function will result in a high value for the utility function. Sensing at locations that have no effect on the position uncertainty will result in low values.

Generally, detecting the object at a location at which there is an a priori low probability of detection will result in a high loss value. This suggests the utility function

$$U(detection = 1, (x, y)) = 1 - P(detection = 1)(x, y)$$
$$= 1 - \Pi(x, y).$$

Similarly, failure to detect the object at a location at which there is a high probability of detection will result in a high utility value. This suggests the following utility function.

$$U(detection = 0, (x, y)) = P(detection = 1)(x, y)$$
$$= \Pi(x, y).$$

This leads to the discrimination function, $D(x, y)$

$$D = U(detection = 0)P(detection = 0)$$
$$+ U(detection = 1)P(detection = 1)$$
$$= \Pi(1 - \Pi) + (1 - \Pi)\Pi$$
$$= 2\Pi(1 - \Pi).$$

The form of this quadratic function is well known and attains its maximum value when $\Pi = 1/2$. Hence for the example of the square of side length l, whose PMF was found in section 3.1, the optimal location for sensing to localize the object is to sense at a location removed a distance $l/2$ along the direction of the edge from the first sensed point. This is the expected result for constraining the position of the object with the minimum number of sensing locations.

4.1. Effect of Finite Sensor Size

It is important to note that a PMF gives the probability of detecting the object at a given point in the work space. With a sensor of finite scope, it is possible to sense over a finite area. Hence it is necessary to be able to calculate the probability of detecting a given feature with a sensor of finite scope when the center of the sensor is located at a given location.

4.2. Salient Features

It has been recognized by several researchers that object features can play an important role in object recognition. The ability of a specific feature to constrain an object interpretation has been termed its saliency and has been utilized in recognizing visually occluded objects by Turney, Mudge, and Volz (1985). The work of Bolles and Cain (1982) also uses certain object features as a focal point for recognizing and locating partially visible objects. There is similarity between the problems of occluded visual data and other sensing methods in which the whole object cannot be sensed at any instant, and it is necessary to base any conclusions on data from certain subsets of the object.

Although the problem of recognizing objects that may be partially occluded by other (known) objects is beyond the scope of this paper, the decision theoretic methods described in this paper can be applied to the problem by simple management of the utility function. Because of the necessity of relying on salient features for reliable object identification, these features must be heavily weighted by the utility function to the exclusion of those features that may occur in the overlapping objects.

5. Sequential Analysis and Updating Expectations

Generally, the decision to be made after each new set of data is acquired is to arrive at a conclusion or to seek further data. The approach taken to solve this problem is that of sequential analysis as described in Wald (1947).

All data that are acquired must then be incorporated with any prior information to update the current knowledge about the environment. For this purpose a probabilistic interpretation tree structure is developed that encapsulates the expectations regarding the associations of sensed points with object model features.

5.1. Sequential Analysis

Using the terminology of sequential analysis, the task of identifying an object can be considered as the testing of a hypothesis regarding the object's identity. If there is a single hypothesis, H_0, to be tested, there is a choice of three options to be made every time new data are obtained. These are to accept H_0, reject H_0, or defer judgment until further data are acquired. Generally, it will not be possible to state with certainty whether H_0 is true from a finite amount of data, but it will be possible to give some probabilistic measure of its likelihood. It is thus possible to specify, in arriving at a decision, the probability that the decision is incorrect.

For the single hypothesis case, there are two possible errors: rejection of H_0 when it is true (called an error of the first kind, whose probability is denoted by α); and acceptance of H_0 when it is not true (called an error of the second kind, whose probability is denoted by β). It is common in sequential analysis to specify the acceptable errors beforehand and to continue acquiring more data until a decision can be made for which α and β fall within the stated bounds.

The application of these principles to the object recognition problem is illuminated in the example at the end of this section.

5.2. Probabilistic Interpretation Trees

The concept of an interpretation tree assigning each sensed point to each instance of the sensed feature in an object model was used by Gaston and Lozano-Pérez (1984) and Grimson and Lozano-Pérez (1984) in recognizing a sensed object from a given set. Their approach used several tests to cull inadmissible inter-

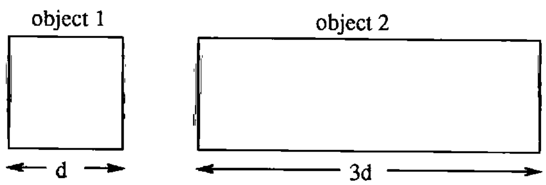


Fig. 27. Object shapes.

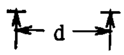


Fig. 28. Sensory data.

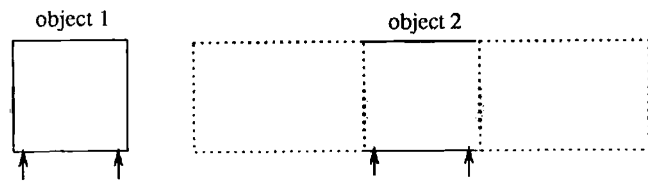


Fig. 29. Possible object interpretations.

pretations from the tree but placed no ordering on the remaining interpretations. Because the techniques proposed in this paper enable the conditional probabilities of sensing particular features to be calculated, these probabilities can be used after sensing has taken place to ascribe an expectation to each admissible interpretation by applying Bayes' Theorem. These a posteriori probabilities can be associated with each remaining branch of the interpretation tree and updated as more data are acquired.

One advantage of this method is that it utilizes the constraints imposed by failing to sense an object at a given location, which cannot be easily integrated into the standard interpretation tree approach. Most importantly, at each stage of the process, the probabilistic method provides information regarding the current probabilities of each interpretation from the admissible set. Using the method of Grimson and Lozano-Pérez (1984), no information is offered as to the relative likelihoods of different interpretations—either an interpretation is admissible or it is pruned from the tree. However, by introducing expectations, the possible interpretations can be ranked according to their likelihood at any stage of the sensing process and a hypothesis accepted or rejected with some known possibility of error.

To compare the two approaches, consider the task of discriminating between the two objects shown in Figure 27, assuming the edge of the object has been sensed at the two locations shown in Figure 28. The possible interpretations for the two objects are shown superimposed on the sensory data in Figure 29.

On the evidence of these two sensed points, it is most likely that the sensed object is *object* 2. This would be reflected in the a posteriori probabilities of the two objects. However, as both objects do give rise to possible interpretations, neither would be culled from the interpretation tree using the approach of Grimson and Lozano-Pérez (1984), and it would be impossible to reach any conclusion with respect to the object's probable identity.

In the following example, an application of these techniques is demonstrated in the problem of recognizing an object from a two-object set.

Example: The Two-Object Set

Consider the problem described in example 1 of the previous section. The hypothesis, H_0, is defined to be "that the sensed object is o_1." It has already been seen that the optimal sensing location is the location that maximizes

$$D(x, y) = |\Pi_{o_1}(x, y) - \Pi_{o_2}(x, y)|.$$

After sensing, it will have been determined whether the object is present at this location. The a posteriori probability can then be calculated[5]:

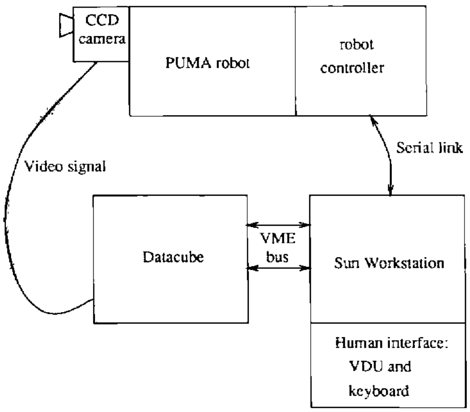
$$P'(o_1) = P(o_1|state) = \frac{P(state|o_1)P(o_1)}{P(state|o_1)P(o_1) + P(state|o_2)P(o_2)}$$

Assume that the error bounds α and β are considered admissible. Then if $P'(o_1) > 1 - \beta$, the hypothesis can be accepted, and if $P'(o_1) < \alpha$, it can be rejected. If $\alpha < P'(o_1) < 1 - \beta$, no decision can be reached, and more data are required.

First, the PMF must be updated using the techniques described in section 3.2 and the distribution function recalculated. More sensory data can then be acquired at the new optimal location and the a poste-

5. The symbol P' is used to denote the a posteriori probabilities in order to distinguish them from the a priori probabilities P.

Fig. 30. Configuration of the experimental equipment.



riori probabilities updated. This procedure can be repeated until the a posteriori probabilities fall within the acceptable bounds, and a decision can be made.

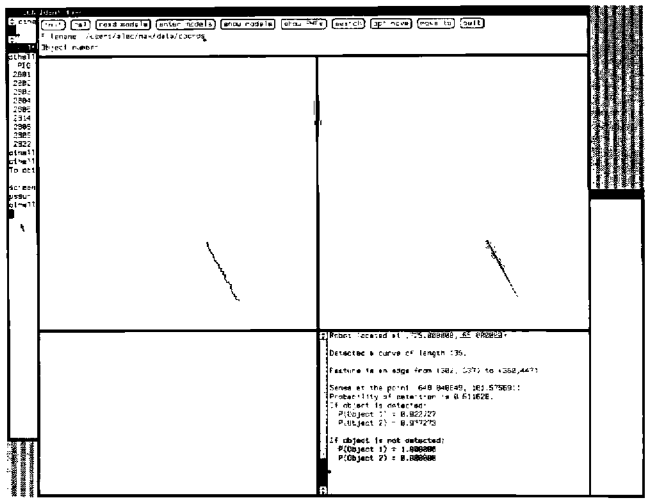
6. An Implementation and Results

To test the principles outlined in this paper, a CCD array camera was mounted on a robot arm and used to obtain sensory data from polygonal shapes placed in the work space of the robot. To stimulate the sparse data and repositioning problems associated with local sensing, the camera's field of view was restricted so that it could only sense a small portion of the object at any instant. The task addressed was to identify the shape from a set of shapes known to the system, and the camera was only moved in a plane parallel to the plane of the shape, giving rise to a strictly two-dimensional problem. The configuration of the robot, camera, and work space is shown in Figure 30.

The images obtained by the camera could be classified as to whether the object had been contacted, and if so, whether an edge or corner feature had been sensed. The image processing was performed on a Sun 3 workstation with the aid of a Datacube vision processing system.

The first task of the system was to move the sensor toward the shape from an extremity of the work space until an edge was detected. Once this contact had been made, powerful constraints limited the possible placements of the shape, and the probabilistic membership

Fig. 31. Screendump from the system in operation.



functions and utility functions described in this paper could be readily evaluated. The processing time required to perform these computations was small in relation to the time taken for the image processing and relocation of the sensor.

A screendump of the information supplied by the implemented system while operating is provided in Figure 31.

For the purposes of this implementation, any uncertainty in the camera data was not considered, which enabled the shapes to be identified with absolute certainty from a finite number of sensed points.

This system was successful in identifying the shapes present in the work space in what was felt to be an efficient manner. The number of sensing locations required to discriminate among a set of objects is clearly dependent on the number of objects and their similarity. The increase in the average number of sensing locations required to identify the object with increasing number of admissible objects in the set was less than a linear relationship. Although it was rarely possible to discriminate between two objects on the basis of two sensory points, it was usually possible to discriminate between n objects on the basis of n points, once n was of the order of five or greater. Clearly more points are required to discriminate between objects which are of very similar size and shape.

For the problem of identifying between the objects shown in Figure 32, an example of a typical strategy generated by the system is shown in Figure 33.

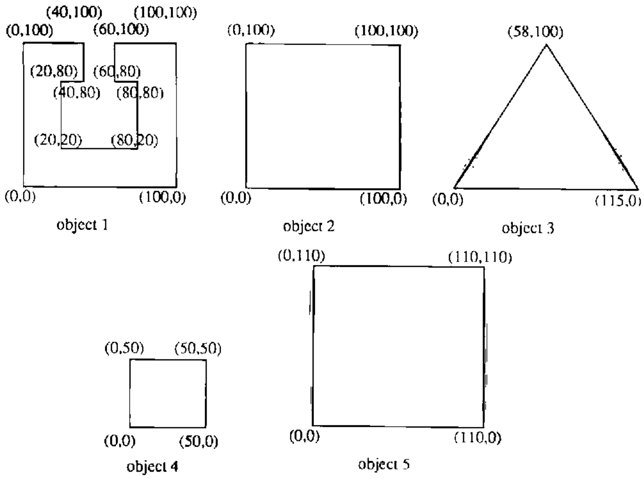


Fig. 32. Object set.

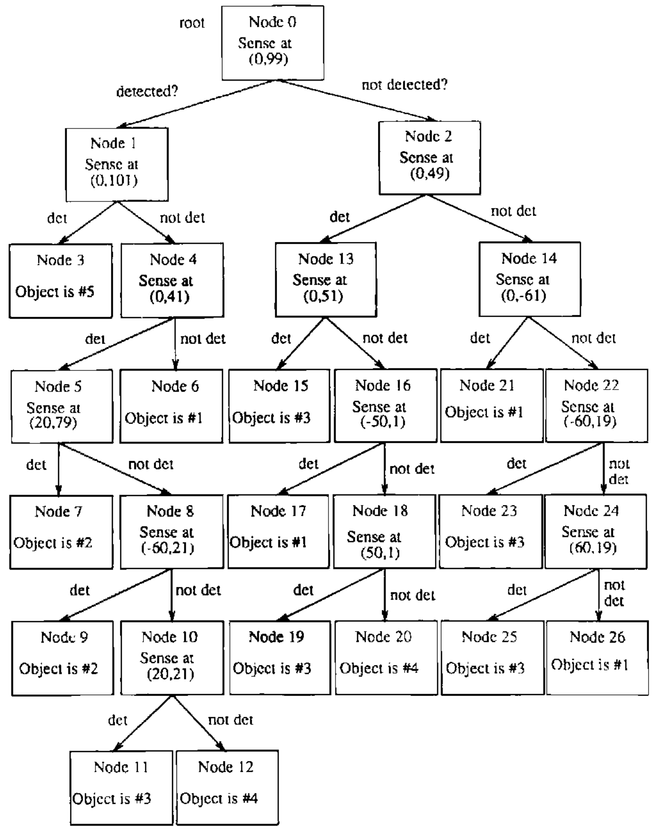


Fig. 33. Typical strategy tree.

7. Conclusions and Future Work

Several features of the system for acquiring and using sensory data described in this paper represent significant advancements over current techniques.

First, the sensing is guided and hence aims to be optimal with respect to the number of sensed points. Rather than acquiring sensor data at a random location, the described method evaluates all possible sensing locations for their ability to constrain the model uncertainties and senses at the chosen location. After sensory data have been acquired, they are then used to update all expectations about the sensing environment, which can then be used to order possible interpretations and can be referenced as prior information in choosing the next sensing location.

These techniques should have particular application in systems in which the sensory data are inherently "sparse" and multiple sensory operations are required. Such sensory systems include tactile sensing and sonar, and potential tasks include object identification and localization and robot navigation.

The most significant work that needs to be done with this approach is to apply the techniques to a three-dimensional work space. This is currently being pursued. Work is also required to determine a more precise approach to generating appropriate utility functions for a given task along with methods of comparing possible utility functions and defining and measuring their "optimality."

Several extensions and refinements to this work, including additional implementations, are described in Cameron (1989).

Appendix A: Characteristics of the PMF

Consider the case where $S(x, y) = \delta(x, y)$. Then

$$\begin{aligned}\Pi(x, y) &= O'(x, y) * F(x, y)\\ &= \int\int O'(x-\mu, y-\eta)\, F(\mu, \eta)\, d\mu\, d\eta\\ &\le \int\int 1 F(\mu, \eta)\, d\mu\, d\eta\\ &= \int\int F(\mu, \eta)\, d\mu\, d\eta\\ &= 1.\end{aligned}$$

Therefore,

$$\Pi(x, y) \leq 1 \quad \forall(x, y).$$

Clearly,

$$\Pi(x, y) \geq 0 \quad \forall(x, y).$$

Thus,

$$0 \leq \Pi(x, y) \leq 1 \quad \forall(x, y).$$

Also,

$$\begin{aligned}
\int\int \Pi(x, y)\, dx\, dy &= \int\int O'(x, y) * F(x, y)\, dx\, dy \\
&= \int\int\left[\int\int O'(x-\mu, y-\eta)\, F(\mu, \eta)\, d\mu\, d\eta\right] dx\, dy \\
&= \int\int\left[\int\int O'(x-\mu, y-\eta)\, dx\, dy\right] F(\mu, \eta)\, d\mu\, d\eta \\
&= \int\int [\text{area of } o]\, F(\mu, \eta)\, d\mu\, d\eta \\
&= [\text{area of } o] \int\int F(\mu, \eta)\, d\mu\, d\eta \\
&= \text{area of } o.
\end{aligned}$$

Appendix B: Discrimination Function for a Two-Object Set

Consider the proposed utility function for the task of identifying an object from the set $\{o_1, o_2\}$:

$$\begin{aligned}
&U[state, (x, y)] \\
&\quad = |P(o_1|state) - P(o_1)| + |P(o_2|state) - P(o_2)| \\
&\quad = |P(o_1|state) - P(o_1)| \\
&\qquad + |(1 - P(o_1|state)) - (1 - P(o_1))| \\
&\quad = 2|P(o_1|state) - P(o_1)|.
\end{aligned}$$

The discrimination function (assuming only two sensing states) is:

$$\begin{aligned}
D(x, y) = {} & U[detection = 1, (x, y)]P(detection = 1) \\
& + U[detection = 0, (x, y)]P(detection = 0).
\end{aligned}$$

Consider the term,

$$\begin{aligned}
&U(state, (x, y))P(state) \\
&\quad = 2|P(o_1|state) - P(o_1)|P(state) \\
&\quad = 2\left|\frac{P(state|o_1)P(o_1)}{P(state)} - P(o_1)\right| P(state) \quad \text{(by Bayes Rule)} \\
&\quad = 2P(o_1)|P(state|o_1) - P(state)| \\
&\qquad \text{(as } P(state) \text{ and } P(o_1) \geq 0 \text{ always)} \\
&\quad = 2P(o_1)|P(state|o_1) - (P(state|o_1)P(o_1) + P(state|o_2)P(o_2))| \\
&\quad = 2P(o_1)|P(state|o_1)(1 - P(o_1)) - P(state|o_2)P(o_2)| \\
&\quad = 2P(o_1)P(o_2)|P(state|o_1) - P(state|o_2)| \quad \text{(as } P(o_2) \geq 0 \text{ always)} \\
&\quad = 2P(o_1)P(o_2)|\Pi_{o_1} - \Pi_{o_2}|.
\end{aligned}$$

Using this result to evaluate the discrimination function:

$$\begin{aligned}
D = {} & U[detection = 1, (x, y)]P(detection = 1) \\
& + U[detection = 0, (x, y)]P(detection = 0) \\
= {} & 2P(o_1)P(o_2)|\Pi_{o_1} - \Pi_{o_2}| + 2P(o_1)P(o_2)|\Pi_{o_1} - \Pi_{o_2}| \\
= {} & 4P(o_1)P(o_2)|\Pi_{o_1} - \Pi_{o_2}|.
\end{aligned}$$

As $P(o_1)$ and $P(o_2)$ are constant over (x, y), it follows that

$$D(x, y) \propto |\Pi_{o_1}(x, y) - \Pi_{o_2}(x, y)|,$$

and the optimal sensing location will be the value of (x, y) where the difference between the PMFs for the two objects is greatest.

References

Berger, J. O. 1980. *Statistical Decision Theory.* New York: Springer-Verlag.

Blackwell, D., and Girshick, M. A. 1954. *Theory of Games and Statistical Decisions.* Mineola, N.Y.: Dover Publications.

Bolles, R. C., and Cain, R. A. 1982. Recognizing and locating partially visible objects: The local-feature-focus method. *Int. J. Robot. Res.* 1(3):57–82.

Browse, R. A. 1987. Feature-based tactile object recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 9(6):779–786.

Cameron, A. J. 1989. A Bayesian approach to optimal sensor placement. University of Oxford, Ph.D. thesis.

DeGroot, M. H. 1970. *Optimal Statistical Decisions.* New York: McGraw-Hill.

Durrant-Whyte, H. F. 1985. *Integration, Coordination and Control of Multi-Sensor Robot Systems.* Boston: Kluwer Academic Publishers.

Ellis, R. E. 1987 (Raleigh, N.C., April). Acquiring tactile data for the recognition of planar objects. *IEEE Int. Conf. Robot. Automat.*, pp. 1799–1805.

Gaston, P. C., and Lozano-Pérez, T. 1984. Tactile recognition and localization using object models: The case of a polyhedra on a plane. *IEEE Trans. Pattern Anal. Mach. Intell.* 6(3):257–266.

Grimson, W. E. L., and Lozano-Pérez, T. 1984. Model-based recognition and localization from sparse range or tactile data. *Int. J. Robot. Res.* 3(3):3–35.

Grimson, W. E. L. 1986. Sensing strategies for disambiguating among multiple objects in known poses. *IEEE J. Robot. Automat.* 2(4):196–213.

Horn, B. K. P. 1986. *Robot Vision.* Cambridge, Mass.: MIT Press.

Schneiter, J. L. 1986 (San Francisco, April). An objective tactile sensing strategy for object recognition and localization. *IEEE Int. Conf. Robot. Automat.*, pp. 1262–1267.

Turney, J. L., Mudge, T. N., and Volz, R. A. 1985. Recognizing partially occluded parts. *IEEE Trans. Pattern Anal. Mach. Intell.* 7(4):410–421.

Wald, A., 1947. *Sequential Analysis.* Mineola, N.Y.: Dover Publications.