An Optimal Sensing Strategy of a Proximity Sensor System for Recognition and Localization of Polyhedral Objects

*Sukhan Lee** and *Hern S. Hahn***

Jet Propulsion Laboratory*
California Institute of Technology
Pasadena, CA 91109

Dept. of EE-Systems*,**
University of Southern California
Los Angeles, CA 90089-0273

Abstract

This paper presents an algorithm for the recognition and localization of 3D polyhedral objects based on an optical proximity sensor system capable of measuring the depth and orientation of a local area of an object surface. Emphasis is given to the determination of an optimal sensor trajectory or an optimal probing, for efficient discrimination among all the possible interpretations. The determination of an optimal sensor trajectory for the next probing consists of 1) the selection of optimal beam orientations based on the Surface Normal Vector(SNV) distribution of the Multiple Interpretation Image(MII) and 2) the selection of an optimal probing plane by projecting the MII onto the projection plane perpendicular to a selected beam orientation and deriving the optimal path on the projection plane. The selection of optimal beam orientation and probing plane is based on the measure of discrimination power of a cluster of surfaces of an MII. The measure of discrimination power is obtained by computing the utility of a cluster of surfaces, representing the expected number of interpretations that can be pruned out. Simulation results are shown.

1 Introduction

There has recently been a growing interest in developing a methodology for the recognition and localization of a 3-D object based on "active sensing" in which a sensor system collects data by actively scanning or probing over the object. The main goal of the above approach is, as pointed out by Grimson[1], the rapid determination of an object model and pose from sparse, noisy, and occluded sensory data by predicting an optimal sensor pose or trajectory that will force a unique interpretation of the object with as few data points as possible.

The accomplishment of the goal is based on the following procedure: 1) With the known features already obtained from the previous sensor readings, the system constructs a multiple interpretation image(MII) or a set of hypotheses representing all the possible interpretations on the object model and pose. 2) The system then determines an optimal sensor pose or trajectory for subsequent data collection, which provides the maximum disambiguability among all the possible interpretations. The research issues involved in implementing the above procedure include 1) an efficient matching of a test object with object models with only partial information on the test object available for the system, 2) an efficient search of the Cartesian space for selecting the optimal sensor pose or trajectory, 3) an object representation which best supports the required matching and searching processes, and 4) a measure of disambiguability of a given sensor pose or trajectory, under the existing sensing errors. The implementation details, however, should take into account the capability of a sensor system, specially with regard to the range of areas it can detect, and the type of features it can extract. For instance, an optimal trajectory or path should be determined for a point sensor, whereas an optimal sensor pose may suffice for an area sensor.

In scheduling tactile sensor moves, Schneiter[2] presented a method to select an optimal path by testing the disambiguability of individual paths ranged from the outside of the union boundary to the inside of the intersection boundary of an MII. Grimson[1] proposed determining of an optimal sensor position on the sensing line (orthogonal to the given sensing direction) by testing the disambiguability of individual segments of the sensing line. Hutchinson et.al[3] partitioned the space into individual segments inside which a sensor detects the same features, and moved the view point to the segment providing the most discrimination. Ellis et.al[4] proposed to represent a family of feasible paths passing through an unsensed edge by an area of trapesoid. An optimal path is selected as a point within the maximum overlap of the trapezoids from all the candidate edges. The above approaches, however, either do not consider an optimal continuous sensor trajectory which takes into account intermediate sensor movements, or do not provide a general and rigorous definition of disambiguability, or have serious limitations in 3-D applications due to excessive computational complexity.

This paper presents an optimal sensing strategy of an optical proximity sensor system[7] for recognizing and localizing a 3-D polyhedral object. In particular, this paper provides the determination of an optimal continuous sensor trajectory based on 1) the selection of optimal beam orientations from a surface normal vector distribution of an MII, and 2) the selection of an optimal probing plane by projecting an MII onto the projection plane perpendicular to a selected beam orientation and deriving an optimal path on the projection plane. This paper also presents the measure of discrimination power of a cluster of surfaces of an MII as a general definition of disambiguability, based on the concept of utility, for the selection of optimal beam orientations and a probing plane.

2 Object Localization

A hierarchical representation of a 3-D polyhedral object is used[8], where the upper level describes the inter-surface relation of a polyhedral object, while the lower level describes individual surfaces by defining the inter-edge relation of each polygonal surface. The representation also includes the Surface Normal Vector(SNV) distribution graph, obtained by translating all the unit SNVs of the surfaces of the object to the origin of the object frame. The SNV emphasizes the structural characteristics of the object and to achieve computational efficiency in determining optimal probing plane.

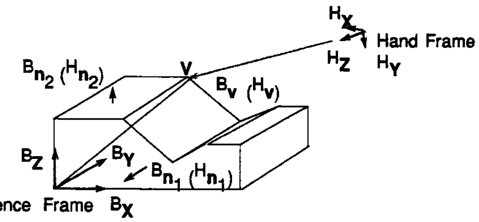


Figure 1: Defining the present hand position referring to reference vertices

Since there may exist many interpretations of the test object, further probings are needed to prune out infeasible interpretations until a unique interpretation remains. This prompts the necessity of selecting an optimal probing which provides maximum discrimination among all the possible interpretations. The selection of an optimal probing is based on a Multiple Interpretation Image, obtained by superimposing the images of the object models corresponding to individual interpretations in Cartesian space with reference to the world frame. An image of the object model can be obtained by using rotational transformation between the object frame and the sensor frame, which is obtained from localizing the test object based on a particular interpretation.

The construction of an MII requires the localization of the test object based on individual interpretations. The sufficient conditions for unambiguous localization of a test object are as follows: 1) the SNVs and distances of three independent surfaces are known, 2) the SNVs and the distances of two non-parallel surfaces and one vertex position are known, or 3) the SNV and distance of one surface and two vertex positions are known.

Assuming the sufficient conditions are satisfied, the transformation ^HR_B from the object frame to the sensor frame and the position vector ^HP can be derived (refer to Fig. 1), where $^{B(H)}V$ is the position vector of vertex V w.r.t. the object (hand) frame and $^{B(H)}n_i$ is the SNV of surface i w.r.t. the object (hand) frame:

$$^HV = {}^HP + {}^HR_B\,{}^BV \tag{1}$$

$$^Hn_1 = {}^HR_B\,{}^Bn_1 \tag{2}$$

$$^Hn_2 = {}^HR_B\,{}^Bn_2 \tag{3}$$

^HR_B is obtained by

$$^HR_B = {}^HN\,{}^BN^{-1} \tag{4}$$

where

$$^HN = [\,^Hn_1,\ ^Hn_2,\ ^Hn_1 \times {}^Hn_2] \tag{5}$$

$$^BN = [\,^Bn_1,\ ^Bn_2,\ ^Bn_1 \times {}^Bn_2] \tag{6}$$

Once ^HR_B is obtained, ^HP can be computed from Eq. (3).

3 Optimal Probing

An additional probing of the test object is required in the case where previous probings have failed to provide sufficient information for the unique recognition and localization of the test object. A probing plane π, in which the probing trajectory lies, is defined by a beam orientation vector ^Wl and a probing direction vector ^Wv: $^Wn_\pi = {}^Wl \times {}^Wv$, where $^Wn_\pi$ represents the surface normal vector of π and the superscript W indicates that the vectors are represented w.r.t. the world frame. Then, we can consider the determination of the optimal probing plane π^* as the problem of finding the optimal combination of a light beam orientation $^Wl^*$ and a probing direction $^Wv^*$.

3.1 Optimal Light Beam Orientation

Given a light beam orientation, the sensor system can measure only those surfaces whose SNVs reside inside the cone, as illustrated in Fig. 2. The cone is referred to as "a SNV range cone".

We first select a list of candidates for the optimal beam orientation, based on the "SNV Distribution Graph (SNVDG) of a Multiple Interpretation Image (MII)". The generation of a list of candidates for the optimal cone orientation is based on the following "SNV clustering algorithm".

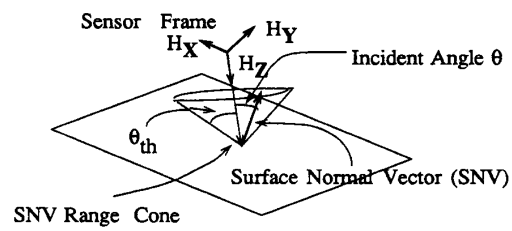


Figure 2: SNV Measurement Range Cone for a given Beam Orientation

SNV Clustering Algorithm

Phase 1: Generation of the list of Tentative Candidates for Optimal Cone Orientation

1. Select a SNV S_k in the SNVDG of the MII, and pose the cone apex of the SNV range cone at the origin of the SNVDG of an MII by making the cone axis coincide with the selected SNV S_k.

2. Generate a SNV cluster $C(S_k)$ by collecting all the SNVs inside the cone, and evaluate the cluster in terms of its discrimination power $M_p[C(S_k)]$. (Refer to Eq. (7))

3. Repeat 1) and 2) for all the individual SNVs, S_k, in the SNV Distribution Graph of an MII.

4. Order the SNV clusters based on their discrimination power; Select those SNV clusters whose discrimination power M_p is greater than the predefined threshold M_{pTH}, and put the SNVs corresponding to the selected SNV clusters in the list of tentative candidates for optimal cone orientation.

Phase 2: Refinement of Tentative Candidates for the Final Candidate List

1. Select a SNV S_k from the tentative candidate list if S_k is not included in the list of tentative candidates at the end of Phase 1, and set $S_k^{old} = S_k$. If all the SNVs in the candidate list are already selected, goto step 6.

2. For the SNV cluster $C(S_k)$ corresponding to a SNV S_k, generate a weighted cluster center $S_k^{weighted}$ by calculating the weighted average of all the SNVs in the cluster, where the weight of a SNV is proportional to the discrimination power of the corresponding SNV cluster: $S_k^{weighted} = \sum_{S_i \in C(S_k^{old})} \sigma_i S_i$,

$\sigma_i = M_p[C(S_i)] / \sum_{S_i \in C(S_k)} M_p[C(S_i)]$, where $M_p[C(S_i)]$ represents the discrimination power of $C(S_i)$.

3. Move the cone axis from S_k^{old} to S_k^{new} incrementally toward the weighted cluster center $S_k^{weighted}$:
$S_k^{new} = S_k^{old} + \triangle S_k$, $\triangle S_k = \alpha(S_k^{weighted} - S_k^{old})$, where α is the scaling constant.

4. Generate a new cluster $C(S_k^{new})$ corresponding to S_k^{new} and evaluate its discrimination power $M_p[C(S_k^{new})]$.

5. If $M_p[C(S_k^{new})] \geq M_p[C(S_k^{old})]$, then set $S_k^{old} = S_k^{new}$ and goto step 2. If $M_p[C(S_k^{new})] < M_p[C(S_k^{old})]$,
then set $S_k^{refined} = S_k^{old}$ and put $S_k^{refined}$ in the list of refined candidates. Go to step 1.

6. Finally, select the predetermined number of candidates from the combined list of refined candidates and the tentative candidates (from the phase 1), in a decreasing order of discrimination power.

Note that the phase 2 of the SNV clustering algorithm can be omitted, in case minimal computational complexity is a premium, without major degradation of system performance.

3.2 The Discrimination Power of a SNV Cluster

The measure of discrimination power, evaluating the performance of individual SNV clusters, is a function of 1) the number of SNVs in the cluster, 2) the utility of individual SNVs in the cluster, and 3) the distribution of SNVs included in the clusters over the individual interpretations. To consider these factors, let us first define the measure of discrimination power of a SNV cluster, $C(S_i)$, in terms of the utility of $C(S_i)$, which represents the expected number of interpretations that can be pruned out by the individual SNVs in the cluster.

<u>Notational Conventions</u>:

- $C(S_i) = \{S_1, S_2, ..., S_p\}$: A cluster of SNVs with S_i as the reference SNV.

- $S_1 \sim M_1$, $S_2 \sim M_2$,...., $S_p \sim M_p$
where $M_a \subset N \triangleq \{1, 2, .., n \mid$ the list of all the interpretations included in the current $MII\}$, $S_a \sim M_a$ implies that S_a belongs to the interpretations specified by the set M_a, i.e. S_a is shared by a number of interpretations in M_a.

- $m_a \triangleq Card(M_a)$: the number of interpretations listed in the set M_a
$m_a \cap m_b \triangleq Card(M_a \cap M_b)$, $m_a \cup m_b \triangleq Card(M_a \cup M_b)$, $n \triangleq Card(N)$

- Prob(S_b/S_a)[Prob($S_b/\overline{S_a}$)] $\triangleq$ Probability that S_b is detected on the condition that S_a has been detected [has not been detected].

- Prune($S_a[\overline{S_a}]$)$\triangleq$ the number of interpretations that can be pruned out when S_a is [not] detected.

- Prune(S_b/S_a)[Prune($S_b/\overline{S_a}$)]$\triangleq$ the number of interpretations that can be pruned out additionally by the detection of S_b after the initial pruning of interpretations by the detection of S_a [the non-detection of S_a].

The utility of a SNV can be defined as follows:

Definition: *The utility of a SNV S_a, $U(S_a)$*

$$U(S_a) \triangleq Prob(S_a)Prune(S_a) + Prob(\overline{S_a})Prune(\overline{S_a})$$
$$\triangleq U_{\frac{1}{2}}(S_a) + U_{\frac{1}{2}}(\overline{S_a})$$

Without *a priori* knowledge on the probability of which interpretation is correct, let us assign a uniform probability distribution among individual interpretations: $Prob(S_a) = m_a/n$ and $Prob(\overline{S_a}) = 1 - m_a/n$. Since $Prune(S_a) = n - m_a$ and $Prune(\overline{S_a}) = m_a$, we have $U_{\frac{1}{2}}(S_a) = \frac{m_a}{n}(n - m_a)$, and $U_{\frac{1}{2}}(\overline{S_a}) = \frac{n-m_a}{n} m_a$. Therefore, $U_{\frac{1}{2}}(S_a) = U_{\frac{1}{2}}(\overline{S_a}) = (1 - \frac{m_a}{n})m_a$.

The utility of a pair of SNVs $\{S_a$ and $S_b\}$, $U(S_a, S_b)$, can be defined as follows:

Definition: *The utility of a pair of SNVs $\{S_a$ and $S_b\}$, $U(S_a, S_b)$*

$$\begin{aligned} U(S_a, S_b) \triangleq\ & Prob(S_a)Prune(S_a) + Prob(\overline{S_a})Prune(\overline{S_a}) \\ & + Prob(S_a)[Prob(S_b/S_a)Prune(S_b/S_a) \\ & + Prob(\overline{S_b}/S_a)Prune(\overline{S_b}/S_a)] \\ & + Prob(\overline{S_a})[Prob(S_b/\overline{S_a})Prune(S_b/\overline{S_a}) \\ & + Prob(\overline{S_b}/\overline{S_a})Prune(\overline{S_b}/\overline{S_a})] \\ \triangleq\ & U_{\frac{1}{2}}(S_a) + U_{\frac{1}{2}}(\overline{S_a}) + Prob(S_a)[U_{\frac{1}{2}}(S_b/S_a) \\ & + U_{\frac{1}{2}}(\overline{S_b}/S_a)] + Prob(\overline{S_a})[U_{\frac{1}{2}}(S_b/\overline{S_a}) + U_{\frac{1}{2}}(\overline{S_b}/\overline{S_a})] \end{aligned}$$

where $U_{\frac{1}{2}}(S_b/S_a) = U_{\frac{1}{2}}(\overline{S_b}/S_a)$ and $U_{\frac{1}{2}}(S_b/\overline{S_a}) = U_{\frac{1}{2}}(\overline{S_b}/\overline{S_a})$.

With a uniform probability distribution among individual interpretations, we have $Prob(S_a)\ U_{\frac{1}{2}}(S_b/S_a) = (m_a \cap m_b)\ [m_a - (m_a \cap m_b)]/n$ and $Prob(\overline{S_a})U_{\frac{1}{2}}(S_b/\overline{S_a}) = [m_b - (m_a \cap m_b)][n - (m_a \cap m_b)]/n$.

The utility of an arbitrary number of SNVs can be defined in this manner by directly extending the above definition. However, there exists a simpler way of computing the utility of an arbitrary number of SNVs, as shown in the following theorems:

Theorem 1: *$S_j \sim M_j, j = 1, .., p, M_j s$ are mutually disjoint and $\cup_{j=1}^{p} M_j = N$, then $U(\{S_j, j = 1, .., p\}) = \sum_{j=1}^{p} U_{\frac{1}{2}}(S_j)$.*

- *Proof:* When $p = 2$, from Lemma 3 (refer to Appendix), $U(\{S_j, j = 1, 2\}) = \sum_{j=1}^{2} U_{\frac{1}{2}}(S_j)$. Let us assume that Theorem 1 holds when $p = k$, i.e.

$$U(\{S_j, j = 1, .., k\}) = \sum_{j=1}^{k} U_{\frac{1}{2}}(S_j)$$

Then, when $p = k + 1$, from Lemma 5

$U(\{S_j, j = 1, .., (k + 1)\})$
$= U_{\frac{1}{2}}(S_{k+1}) + U_{\frac{1}{2}}(\overline{S_{k+1}}) + Prob(\overline{S_{k+1}})U((S_i, i = 1, .., k)/\overline{S_{k+1}})$
$= U_{\frac{1}{2}}(S_{k+1}) + U_{\frac{1}{2}}(\overline{S_{k+1}}) + Prob(\overline{S_{k+1}}) \sum_{j=1}^{k} U_{\frac{1}{2}}(s_j/\overline{S_{k+1}})$
$= Prob(S_j)[Prune(S_j) -$

$$Prune(\overline{S_{k+1}})]Prob(\overline{S_{k+1}})\sum_{j=1}^{k}U_{\frac{1}{2}}(s_j/\overline{S_{k+1}})$$
$$= \sum_{j=1}^{k} Prob(S_j)[Prune(S_j) - Prune(\overline{S_{k+1}})]$$
$$= \sum_{j=1}^{k} U_{\frac{1}{2}}(S_j) - Prob(\overline{S_{k+1}})Prune(\overline{S_{k+1}})$$

Therefore, we have

$$U(\{S_j, j = 1, .., p\}) = \sum_{j=1}^{p} U_{\frac{1}{2}}(S_j)$$

Q.E.D.

To consider the more general case where M_js are not disjoint, let us introduce the following definitions:

Definition: *Disjoint Minimal covering, DMC(N) of a set N*
A collection of disjoint subsets, $\{M_l^{'}, l = 1, .., q\}, M_{l_1}^{'} \cap M_{l_2}^{'} = \phi, \forall M_{l_1}^{'}, M_{l_2}^{'} \in \{M_l^{'}, l = 1, .., q\}$ and $l_1 \neq l_2$, which covers the set N, such that $N = \cup_{l=1}^{q} M_l^{'}$ is called the disjoint minimal covering, DMC(N), of N:

$$DMC(N) \triangleq \{M_l^{'}, l = 1, .., q\}$$

Definition: $DMC(N/C(S_j))$
A disjoint minimal covering of N, which is derived from a cluster of SNV, $C(S_i)$, $C(S_i) = \{S_j, j = 1, .., p\}$ with $S_j \sim M_j$, is defined as $DMC(N/C(S_j))$.

With the definition of $DMC(N/C(S_i))$, the utility of an arbitrary cluster of SNVs can be computed based on the following theorem:

Theorem 2: *The utility $U(C(S_i))$ of a cluster of SNVs $C(S_i) = \{S_l, j = 1, .., p\}$ and $S_l \sim M_l, l = 1, .., p$, is obtained from $DMC(N/C(S_j))$, $DMC(N/C(S_j)) = \{S_l^{'} \mid S_l^{'} \sim M_l^{'}, l = 1, .., q\}$:*

$$U(C(S_i)) = \sum_{l=1}^{q} U_{\frac{1}{2}}(S_l^{'})$$

- *Proof:* From Lemma 5 (refer to Appendix),if $S_a \sim M_a, S_b \sim M_b, M_a \subset M_b$, then $U(S_a, S_b) = U(S_a, S_b^{'})$, $S_b^{'} = S_b - S_a \cap S_b$. This can be interpreted based on Lemma 6 (refer to Appendix), as follows:

  $$U(S_a, S_b) = U(S_a, S_a \cap S_b, S_b^{'}) = U(S_a, S_b^{'})$$
  $$(since S_a \cap S_b = S_a, S_a \cup S_b^{'} = S_b)$$

  Since $DMC(N/C(S_j)) = \{S_l^{'}, S_l^{'} \sim M_l^{'}, l = 1, .., q\}$ is obtained by considering all the possible subset boundaries, each $M_l, l = 1, .., p$ can be represented by the union of some $M_l^{'}$s. Therefore, from Lemma 6 (refer to Appendix),

  $$U(\{S_l^{'}, l = 1, .., q\}) = U(\{\hat{S}_l^{'}, l = 1, .., q\}).$$

  Now applying Theorem 1, we have

  $$U(C(S_i)) = \sum_{l=1}^{q} U_{\frac{1}{2}}(S_l^{'})$$

  Q.E.D.

For convenience, $U(C(S_i))$ can be normalized based on the fact that the pruning of $n-1$ interpretations implies the complete discrimination among n interpretations:

Definition: *Normalized Utility, $U_n(C(S_i))$, of a SNV cluster $C(S_i)$.*

$$U_n(C(S_i)) \triangleq U(C(S_i))/(n-1)$$

Note that $U_n(C(S_i)) \leq 1$ and $U_n(C(S_i)) = 1$ implies complete disambiguation, i.e. a unique selection of interpretation. Since the detection of SNVs depends not only on the light beam orientation but also on the probing path (direction), some of the SNVs in $C(S_i)$ may not be accurately detectable by the sensor taking a certain probing path, although the sensor uses the beam orientation S_i. This implies that 1) it is necessary to select multiple candidates for optimal beam orientation, so that the optimal combinations of beam orientation and probing path can be determined, and 2) the discrimination power of $C(S_i)$ for selecting optimal beam orientations should indicate the distribution of the utilities of individual SNVs over interpretations to consider the probability of detecting SNVs of different interpretations by an unknown probing path. Thus, the following definition is introduced:

Definition: *Accumulated Utility Histogram:* An accumulated utility of $C(S_i)$ for an interpretation l, $A_l(S_i)$ is defined by summing up the normalized utilities of individual SNVs of $C(S_i)$ associated with the interpretation l: $A_l(C(S_i)) \triangleq \sum U_n(S_k), \forall S_k \in C(S_k)$, and $S_k \sim M_k$ with the interpretation $l \in M_k$. Then, the accumulated utility histogram of $C(S_i)$ over interpretations can be defined by $A(C(S_i)) = \{A_l(C(S_i)), l = 1, .., n\}$. $A(C(S_i))$ is characterized by its mean $m(A(C(S_i)))$ and variance $\sigma^2(A(C(S_i)))$.

Then, the discrimination power, $M_p(C(S_i))$, of a SNV cluster, $C(S_i)$, is defined as follows:

Definition: *Discrimination power $M_p(C(S_i))$ of a SNV cluster $C(S_i)$*

$$M_p(C(S_i)) = \alpha U_n(C(S_i)) + \beta[m(A(C(S_i)))/(\gamma + \sigma(A(C(S_i))))] \quad (7)$$

where α, β, and γ are weighting coefficients.

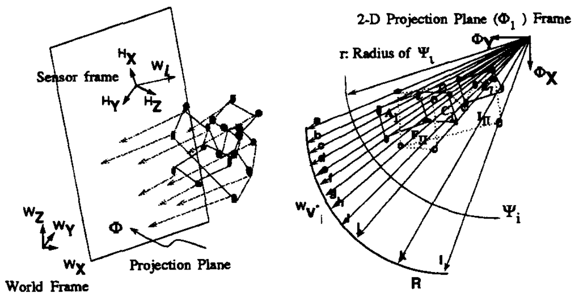
Note that, in case there exists no uncertainty regarding the detectability of all SNVs in $C(S_i)$, $M_p(C(S_i))$ can be defined as

$$M_p(C(S_i)) = U_n(C(S_i)) \quad (8)$$

3.3 Determination of Optimal Probing Plane

For each candidate for the optimal beam orientation ${}^W l_i$ obtained previously, an optimal probing direction ${}^W v_i$ can be selected such that a candidate for the optimal probing plane π, whose surface normal vector ${}^W n_\pi$ is determined by ${}^W n_\pi = {}^W l_i \times {}^W v_i$, can be generated. The determination of ${}^W v_i$, given ${}^W l_i$, is based on the following procedure:

1. From the MII for the test object, select all the vertices of those surfaces that belong to the candidate SNV cluster $C(S_i)$ and form a vertex distribution graph, as shown in Fig. 3(a). Each vertex is labeled with a particular interpretation(s) and a particular surface(s).
2. Define the projection plane Φ_i whose surface normal is the



(a) Vertex Distribution Graph (b) Determination of the optimal probing directon

Figure 3: Generation of a surface image and the determination of the probing direction on the projection plane, Φ_i: ${}^{w}v_i^*$ is the selected probing direction, $S_i, i = 1,,6$ represent the projected-surface-range-arc, and R describes the total-arc-range.

selected beam orientation ${}^{W}l_i$ and which includes the origin of the sensor frame from the last probing, i.e. $\Phi_i = \{{}^{W}X_i \mid {}^{W}l_i \cdot ({}^{W}X_i - {}^{W}S_i) = 0\}$. ${}^{W}S_i$ is the position vector of the origin of the sensor frame w.r.t. the world frame.

3. Project all the vertices located in the vertex distribution graph onto the projection plane Φ_i, as illustrated in Fig. 3(a).

4. Select the probing direction, ${}^{W}v_i$, on the projection plane Φ_i by finding a straight line on Φ_i which can provide the maximum discrimination power.

The details of the step 4 of the above procedure are listed in the following (refer to Fig. 3(b)):

1. All the vertices projected on the projection plane Φ_i are once again projected onto the projection circle Ψ_i by the projection rays emitted from the center of Ψ_i. The center of Ψ_i is the origin of the sensor frame and the radius r is arbitrarily determined. Then, the outer two vertices representing the boundaries of each surface on Ψ_i is connected forming an arc called a projected-surface-range-arc or simply a range-arc. The union of all the range-arcs on Ψ_i defines the total-arc-range.

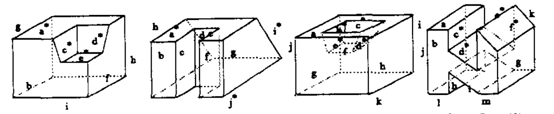
2. The total-arc-range is partitioned into a number of arc segments with the distinctive boundary points of individual range-arcs. An arc-segment represents the intersection of a number of range-arcs.

3. For each arc-segment a_k, collect all the surfaces corresponding to the range-arcs which include an arc-segment a_k, and form an arc-segment cluster $C(a_k)$. Evaluate $C(a_k)$ using $M_p[C(a_k)]$ defined in Eq. (7).

4. Repeat step 3 for all the arc-segments, $a_k, k = 1..m$, and select the one with the highest measure of discrimination power as the optimal arc-segment a^*. Then, ${}^{W}v_i$ is determined by the direction passing through the center of the optimal arc-segment a^*.

The above procedure associates the SNV cluster $C(S_i)$ with the pair $({}^{W}l_i, {}^{W}v_i)$ and associates $M_p[C(a_k^*)]$ to the measure of discrimination power $M_p[C(S_i)]$ for each candidate for the optimal beam orientation. Finally, the optimal probing plane is determined by selecting the optimal $({}^{W}l^*, {}^{W}v^*)$, where $({}^{W}l^*, {}^{W}v^*) = \{({}^{W}l_i, {}^{W}v_i) \mid \max_i M_p[C(S_i)]\}$.

The trajectory of the origin of the sensor frame, called the probing trajectory, is now defined on the optimal probing plane determined above. The actual trajectory of the probing is automatically determined by the automatic surface tracking mechanism of a proximity sensor system[7]. The tracking mechanism maintains constant depth of a sensor from the probing point while avoiding collisions. Along the probing trajectory, the system detects the presence of a certain surface at the assumed position specified by the MII and prunes out infeasible interpretations.



(a) Object A (n=9) (b) Object B (n=10) (c) Object C (n=11) (d) Object D (n=13)

Figure 4: Polyhedral objects used as object models: Object A is also given as the test object in the simulation.

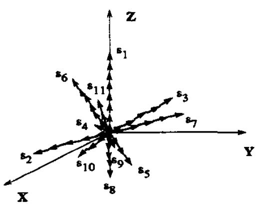


Figure 5: SNV Distribution Graph of the Multiple Interpretation Image for the test object: obtained based on four possible interpretations

4 Simulation

The algorithm presented in this paper has been applied to the recognition and localization of a 3-D test object given four object models, as shown in Fig. 4. In the simulation, we choose two different initial probings on a test object: one generating 4 possible interpretations and the other generating 48 possible interpretations. Each of them provided unambiguous localization. We tested how many probings are needed to prune out all of the incorrect object models and completely localize a test object. The first initial probing generated the MII, given in Fig. 5, of the test object, based on the four possible interpretations obtained from the matching. Two cluster centers, $C(S_5^{new})$ and $C(S_6^{new})$, are selected as the candidates for the optimal light beam orientation, and their corresponding probing directions are derived from the surface images on the projection planes, given in Fig. 6.

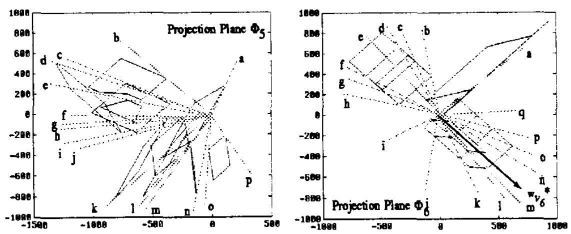
The measures of discrimination power of individual arc-segments in both projection plane were tested, and then the center of arc-segment $\{m - n\}$, ${}^{w}v_6$, having the largest measure of discrimination power, was selected as the next probing direction. Therefore the optimal next probing plane was determined by the two vectors: the center of SNV cluster $C(S_6^{new})$ and ${}^{w}v_6$. The arc-segment associated with ${}^{w}v_6$ includes a set of surfaces

$(S_g^I, S_f^I, S_m^{III}, S_g^{IV})$, and only S_g^I is detected during the probing. Thus interpretation I remained and the other interpretations were pruned out.

The second initial probing generated 48 possible interpretations. Since there existed a large number of candidates for optimal beam orientation, we selected only 10 candidates randomly out of all the possible candidates to reduce the processing time. The number of probings for complete localization varied depending on the selected set of candidates for the beam orientation.

The algorithm has been evaluated based on the number of probings required for a complete recognition and localization. If the goal is to localize a known (model) object and if p interpretations for the test object are possible from the previous k probings, then the maximum number of probings, N, required for complete localization is:

$$N = k + \max_{i=1}^{q} [\text{ card } \{M_{l_i}\}]$$



(a) For SNV Cluster $C(S_5^{new})$ (b) For SNV Cluster $C(S_6^{new})$

Figure 6: Determination of optimal probing direction on the projection plane Φ: ${}^w v_6^*$ in (b) is the selected probing direction.

where M_{l_i} is a disjoint subset defined in the Appendix, and $card\{M_{l_i}\}$ is the number of interpretations that the disjoint minimal set M_{l_i} covers; $\sum_{i=1}^{q} card\{M_{l_i}\} = p$. If we assume that k previous probings detected n_k surfaces among n surfaces, then the number of probings for complete recognition is the number of measurement cones that disjointly covers $(n - n_k)$ SNVs of the model object.

5 Conclusion

The recognition and localization of a 3-D object based on "active sensing" provides several advantages over fixed-view sensing: 1) it can reduce a large amount of data processing time by using a simple sensor detecting simple features, 2) it can actively search for critical features having high discrimination power based on an optimal probing strategy, 3) it can sequentially prune out infeasible interpretations among all the possible interpretations, and 4) it can easily handle an occluded or, even, invisible object. The price to be paid for the advantages is the necessity of determining an optimal probing, as well as sequential sensor movement.

This paper presented an algorithm which determines an optimal probing, i.e. an optimal continuous sensor trajectory of a proximity sensor system, for recognizing and localizing a 3-D object. This paper also introduced the measure of discrimination power based on utility concept to be used as a general form of disambiguability. An extension of this work is underway to include the recognition and localization of a 3-D object having quadric surfaces such as cylinders, spheres, and cones.

References

[1] W. Eric L. Grimson *Sensing Strategies for Disambiguating Among Multiple Objects in Known Poses* IEEE J. of Robotics and Automation, Vol. RA-2, No.4, Dec. 1986

[2] John L. Schneiter *An Objective Tactile Sensing Strategy for Object Recognition and Localization.* pp.1262-1267, Proc. of the 1986 IEEE ICRA

[3] S.A.Hutchinson, R.L.Cromwell,and A.C.Kak *Planning Sensing Strategies in a Robot Work Cell with Multi-Sensor Capabilities* 1988 IEEE ICRA. pp. 1068-1075

[4] R.E.Ellis, Edward M. Riseman, and Allen R. Hanson *Tactile Recognition by Probing: Identifying a Polygon on a Plane* Proc. of AAAI-86, Aug. 1986, pp. 632-637.

[5] Peper C. Gaston and Tomas Lozano-Perez *Tactile Recognition and Localization Using Object Models: The Case of Polyhedra on a Plane.* IEEE Tr. on PAMI, Vol. PAMI-6, No.3, May 1984

[6] M. Dhome and T. Kasvand *Polyhedra Recognition by Hypothesis Accumulation* IEEE Tr. on PAMI, Vol. PAMI-9, No. 3, May 1987

[7] Sukhan Lee and Hern S. Hahn *Design of a Miniaturized Optical Proximity Sensor System: HexEYE* JPL Report. 1987

[8] Sukhan Lee and Hern S. Hahn *Object Recognition and localization Using Optical Proximity Sensor System: Polyhedral Case* IEEE Proc. of Workshop on Interpretation of 3D Scenes. 1989

Appendix: Lemmas

Lemma 1: $S_a \sim M_a, S_b \sim M_b, M_a \cup M_b = N, M_a \cap M_b = \phi$, *then* $U(S_a) = U(S_b)$

Lemma 2: $S_a \sim M_a, S_b \sim M_b$, *and* $M_a = M_b$, *then* $U(S_a, S_b) = U(S_a)$ or $U(S_b)$

Lemma 3: $S_a \sim M_a, S_b \sim M_b, M_a \cap M_b = \phi, M_a \cup M_b = N$, *then* $U(S_a, S_b) = U_{\frac{1}{2}}(S_a) + U_{\frac{1}{2}}(S_b)$

Lemma 4: $S_a \sim M_a, S_b \sim M_b, M_a \cap M_b \neq \phi, M_a \cup M_b = N$, *then* $U(S_a, S_b) = U_{\frac{1}{2}}(S_a') + U_{\frac{1}{2}}(S_{ab}) + U_{\frac{1}{2}}(S_b')$, *where*

$$\begin{aligned} S_a' \sim M_a' &\triangleq M_a - (M_a \cap M_b) \\ S_{ab} \sim M_{ab} &\triangleq M_a \cap M_b \\ S_b' \sim M_b' &\triangleq M_b - (M_a \cap M_b) \end{aligned}$$

Lemma 5: $S_a \sim M_a, S_b \sim M_b, M_a \cap M_b \neq \phi, M_a \cup M_b \subseteq N$, *then* $U(S_a, S_b) = U_{\frac{1}{2}}(S') + U_{\frac{1}{2}}(\overline{S'}) + Prob(\overline{S'}) U((S_a, S_b)/\overline{S_a'}) = U_{\frac{1}{2}}(S_a') + U_{\frac{1}{2}}(S_{ab}) + U_{\frac{1}{2}}(S_b') + U_{\frac{1}{2}}(S')$, *where*

$$\begin{aligned} S_a' \sim M_a' &\triangleq M_a - (M_a \cap M_b) \\ S_{ab} \sim M_{ab} &\triangleq M_a \cap M_b \\ S_b' \sim M_b' &\triangleq M_b - (M_a \cap M_b) \\ S' \sim M_c &\triangleq N - M_a \cup M_b \end{aligned}$$

Lemma 6: $S_a \sim M_a, S_b \sim M_b, S_c \sim M_c, M_a \cup M_b = M_c$, *then* $U(S_a, S_b, S_c) = U_{\frac{1}{2}}(S_a, S_b)$