Performance-Oriented View Planning for Model Acquisition

Scott, William; Roth, Gerhard; Rivest, J.F.

National Research Council Canada
Institute for Information Technology

Conseil national de recherches Canada
Institut de technologie de l'information

NRC · CNRC

*Performance-Oriented View Planning for Model Acquisition **

Scott, W., Roth, G., Rivest, J.F.
May 2000

* published in The International Symposium on Robotics (ISR 2000). pp. 212-219, Montréal, Québec, Canada. May 2000. NRC 45872.

Canada

Performance-Oriented View Planning for Automatic Model Acquisition

William R. Scott[†‡], Gerhard Roth[‡], Jean-François Rivest[†]

[†] Department of Electrical Engineering,
University of Ottawa, Ottawa, Canada, K1N 6N5
rivest@site.uottawa.ca

[‡] Visual Information Technology Group,
National Research Council of Canada, Ottawa, Canada, K1A 0R6
(william.scott,gerhard.roth)@iit.nrc.ca

Abstract

Applications for 3D models of objects and scenes are rapidly growing in number. Active sensors are the most commonly used means of acquiring geometric models. The current acquisition process of view planning, sensing, registration and integration requires a high level of intervention by imaging specialists with extensive training and experience. Automation would improve productivity, freeing humans for higher level tasks. While progress has been made, general purpose, automated model acquisition remains an open problem. View planning, the process of determining a suitable set of sensor viewpoints, is subject to numerous competing constraints. This paper presents a theoretical framework and concept for automated view planning. The goal is to automatically obtain geometric models of a single object with a triangulation-based active range sensor.

{*Keywords: view planning, range sensors, object reconstruction, geometric modelling*}

1 Introduction

The imaging environment comprises three main elements: object, geometric sensor and sensor-object positioning system. View planning, also known as the next-best-view (NBV) problem, involves several challenges, including the following. The optical baseline of a triangulation sensor is often significant with respect to the stand-off distance, with the consequence that shadow effects are an important consideration. Calibration will remove most systematic errors within the calibrated sensor frustum, however there are residual random errors and biases. These non-isotropic effects typically vary quadratically with range, linearly with lateral displacement from the camera boresight and inversely in proportion to the cosine of the grazing angle. Reflectance or geometric discontinuities introduce measurement biases. Wild measurements (outliers) can result from several phenomena.

Viewing perspective can be changed by varying either the sensor or object position. Positioning systems are subject to mechanical positioning uncertainty (greater or lesser than sensor measurement precision), limited range of motion and, frequently, limited degrees of freedom.

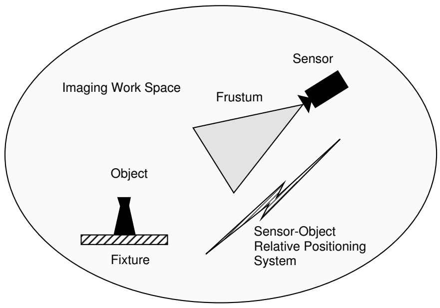

Figure 1: The Geometric Imaging Environment

Surface topology and geometry of real physical objects can be quite complex, with multiple holes, concavities, protrusions and edges. Consequently, the object will self-occlude in complex ways with variation in sensor viewpoint. Furthermore, shape complexity dictates multiple object views for all aspect coverage.

For all these reasons, any solution to the NBV problem must satisfy several requirements, including the following:

- meet specified modeling goals while being efficient, robust and self-terminating,
- demand limited a priori object knowledge and impose few constraints on object shape,
- incorporate a realistic sensor measurement performance model, and
- accommodate a positioning system with a many degrees of freedom and incorporate associated constraints and a positioning performance model.

No current method meets these requirements.

2 View Planning Survey

Conventional non-model-based view planning methods (Figure 2) can perhaps best be categorized by the domain of reasoning about viewpoints - that is, based on surface, imaging volume or global attributes. Some methods combine multiple techniques.

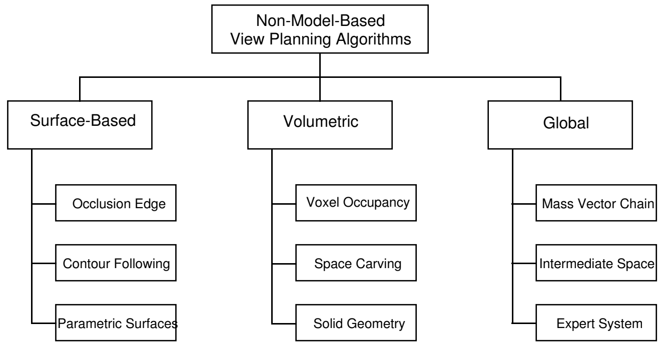

Figure 2: View Planning Algorithms

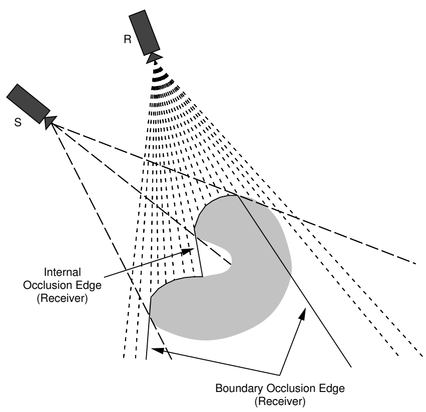

Figure 3: Occluding Edges in Geometric Images

Surface-Based Methods The most commonly exploited NBV mechanism is a geometric jump or *occlusion edge* [5]. As illustrated in Figure 3, the premise is that jump edges internal to an image indicate surface areas not yet sampled, while boundary jump edges represent the boundary of the unobserved volume. Once having located a portion of the object, *contour following* [8] involves "painting" the object surface with the sensor by keeping it in close proximity at all times. The technique has been applied to a unique range sensor sub-class with a limited sensing volume where collision avoidance is inherently a primary concern. Uncertainty in fitting *parameterized superquadrics* to segmented range data [11] has been used to guide viewpoint selection.

Volumetric methods Two *voxelization methods* are in wide usage - voxel occupancy grids and octrees. Recent work [1] selects the NBV oriented to the centroid of the cluster containing the largest number of unknown voxel faces. Voxelization has also been employed [4] with a NBV objective function incorporating a weighted sum of visibility and quality terms. Earlier work [2] employs octrees to more efficiently encode voxel occupancy. The method selects a NBV as the one able to acquire the most information about the unseen imaging volume based on either a global or local visibility analysis. *Space carving* [6] has been applied to the same class of reconstruction problem as contour following. The sensor is swept through the imaging work space in a pre-planned methodical manner, diverting around obstacles, with the objective of reliably labeling work space voxel occupancy. Standard *solid geometry* algorithms available with most CAD packages have been used to model object knowledge [9].

Global View Planning Methods A few methods derive view planning cues from global rather than local characteristics of the geometric data. In [12], NBV selection is based on a global analysis of the *mass vector chain* for the object surface. The *intermediate space technique* [7] has been used to separate visibility analysis of the object surface from that of the sensor. As yet, *expert system* approaches have been restricted mainly to view planning for conventional intensity imaging tasks. In a prototype industrial digitization system [3], the view planning strategy devised by a human operator is implemented semi-automonously by lower level automated scanning primitives using a contouring following scheme after a rough model is first constructed by space carving.

Open Problems As presently developed, none of the current view planning methods are suitable for high performance automated object geometric reconstruction for the following principle reasons: highly constrained viewpoint space, inadequate sensor and positioning system performance characterization, overly simplistic sensor and positioning system models, and excessive computational complexity (even for a highly constrained viewpoint space). The main open view planning problems are efficiency, accuracy and robustness.

3 Theoretical Framework

View planning is the process of determining a suitable set of viewpoints and associated imaging parameters for a specified object reconstruction task using a geometric sensor. The process concerns geometric relationships between two spaces of differing dimensions - the two dimensional [1] surface space $S \subset \Re^2$ of the target object and a multi-dimensional imaging work space, or *viewpoint space* V. For the class of problem addressed by this work, S forms a closed space with complex geometry and possibly complex global topology. A generalized viewpoint $(\mathbf{v}, \lambda_\mathbf{s})$ includes sensor pose $\mathbf{v}$ (position and orientation) plus a set (possibly null) of controllable sensor parameters $\lambda_\mathbf{s}$ such as power and scan length. In general, therefore, a viewpoint has a minimum of six geometric degrees of freedom,

[1] While the surface of a 3D object is strictly a subset of $\Re^3$, a segmented surface always admits a 2D parameterization.

which we show as $V \subset \Re^{6+}$. In practice, viewpoint space V is subject to constraints - typically, an inner boundary determined by collision avoidance considerations, an outer boundary determined by the sensor maximum stand-off range and positioning system limitations with respect to range of motion and possibly degrees of freedom.

Visibility analysis is important, but is not the sole issue. Several factors exacerbate this already computationally difficult task. First, triangulation-based geometric sensors are inherently bi-static, requiring a target region to be simultaneously visible to both the optical transmitter and receiver. Additionally, reconstruction and hence view planning involves measurability not visibility, necessitating models of the measurement physics, an inherently noisy process. Finally, for metrology applications, it is appropriate to specify *objective* modeling goals in addition to all-aspect coverage.

The object surface space S and viewpoint space V can be discretized into suitably small elementary regions, $\mathbf{s}_i$ and $\mathbf{v}_j$ respectively. How discretization takes place is important, but we leave that aside for the moment.

By definition, performance-oriented object reconstruction begins with a set $Q = \{q_k \,; k \in K\}$ of quality factors such as sampling density and measurement precision. The overall reconstruction quality specification $\mathcal{Q}$ will be a user-defined function of quality factors appropriate for the application. Quality lies in the eye of the beholder and can be specified in a variety of ways - analogue or binary, linear or non-linear, uniform or non-uniform. In the uniform case, $\mathcal{Q} = \mathcal{Q}(Q)$. Where quality needs vary for special interest regions such as high curvature zones, $\mathcal{Q} = \mathcal{Q}(Q, S)$.

Additionally, we can think of the quality specification as a function or filter operating on individual surface measurements. Let $\hat{q}_{ijk}$ be the estimated quality of an single surface sample $\mathbf{s}_i$ from viewpoint $\mathbf{v}_j$ with respect to quality factor k. For example, $\hat{q}_{ij1}$ could be estimated local sampling density and $\hat{q}_{ij2}$ estimated local measurement precision. Then, the quality of a single measurement from a specific viewpoint can be expressed as follows:

$$\hat{q}_{ij} = \mathcal{Q}(\{\, \hat{q}_{ijk} \,; k \in K\}) \tag{1}$$

Example specifications include quality as an analogue figure of merit with weighting factors

$$\hat{q}_{ij} = w_1 \hat{q}_{ij1} \, + \, w_2 \, \hat{q}_{ij2} \tag{2}$$

or as a composite binary pass/fail variable:

$$\hat{q}_{ij} = (\hat{q}_{ij1} \geq q_1)\,(\hat{q}_{ij2} \geq q_2) \tag{3}$$

Measurability information can be assembled into a measurability matrix [10] $\mathbf{M}$ as illustrated in Figure 4

$$\mathbf{M} = \mathbf{M}(S \times V, \mathcal{Q}) \tag{4}$$

where $S \times V = \{\, (\mathbf{s}_i, \mathbf{v}_j) \mid \mathbf{s}_i \in S, \mathbf{v}_j \in V \,\}$ and $\mathcal{Q}$ is the quality specification for the modeling task. Then,

$$\mathbf{M} = [m_{ij}] \tag{5}$$

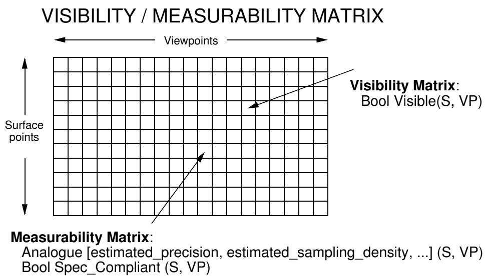

Figure 4: Measurability Matrix

where the elements of $\mathbf{M}$ are measurability estimates

$$m_{ij} = \hat{q}_{ij} \tag{6}$$

It is instructive to partition $\mathbf{M}$ into a set of column vectors $\mathbf{M}_{S,j}$ and row vectors $\mathbf{M}_{i,V}$. The set S_j of surface elements measurable by a single viewpoint $\mathbf{v}_j$ is defined by the corresponding column vector $\mathbf{M}_{S,j}$ of the measurability matrix. This can be expressed as a *measurability mapping* $\mathbf{v}_j \xrightarrow{\mathcal{M}} S_j$ as follows:

$$\begin{aligned} S_j &= \{\mathbf{s}_i\}_j & (7)\\ &= \{\, \mathbf{s}_i \mid m_{ij} \geq q \,,\, \mathbf{v}_j \in V \,\} & (8)\\ &= \mathcal{M}(S \,, \mathbf{v}_j \,, \mathcal{Q}) & (9) \end{aligned}$$

Similarly, the region V_i of viewpoint space from which a given surface element $\mathbf{s}_i$ is measurable is defined by the corresponding row vector $\mathbf{M}_{i,V}$ of the measurability matrix. This can be expressed as an *inverse measurability mapping* $\mathbf{s}_i \xrightarrow{\mathcal{M}^{-1}} V_i$ such that

$$\begin{aligned} V_i &= \{\mathbf{v}_j\}_i & (10)\\ &= \{\, \mathbf{v}_j \mid m_{ij} \geq q \,,\, \mathbf{s}_i \in S \,\} & (11)\\ &= \mathcal{M}^{-1}(\mathbf{s}_i \,, V \,, \mathcal{Q}) & (12) \end{aligned}$$

As illustrated in Figure 5, both $\mathcal{M}$ and $\mathcal{M}^{-1}$ are one-to-many (possibly null) mappings. $\mathcal{M}$ defines regions of the object surface measurable from a given viewpoint, whereas $\mathcal{M}^{-1}$ carves out regions of viewpoint space from which a given surface element is measurable. $\mathcal{M}$ partitions V into regions of equivalent measurability within which measurability typically varies smoothly but non-linearly and which are bounded by abrupt transitions defined by occlusion and measurement phenomena.

The initial work on measurability matrices by Tarbox & Gottschlich [10] was applied to inspection, in which case a detailed object model is available a priori. They compute a measurability matrix encoding a complete visibility analysis over the set of all viewpoints and all surface elements. This conceptually compact representation offers numerous possibilities for intelligent view planning. However, the

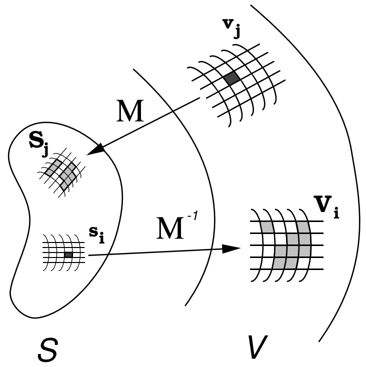

Figure 5: Measurability Mapping

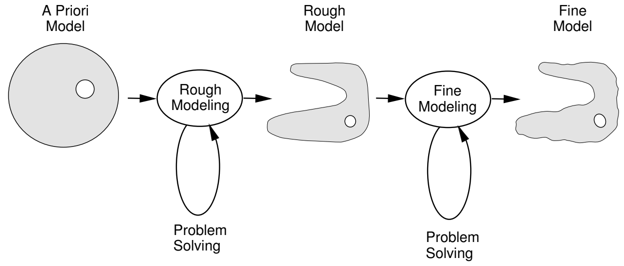

Figure 6: Multi-stage Model-based View Planning

conventional form of a measurability matrix suffers from prohibitive time and memory complexity, even at fairly coarse resolution of object and viewpoint space.

None-the-less, this representation allows us to immediately formulate a compact expression for the view planning problem and from it draw an important conclusion. The view planning task, then, is to find a set of viewpoints measuring (or covering) the object surface:

$$Find \; \{ \mathbf{v}_j \} \; s.t. \; \{ S_j \} \; \supseteq \; S \tag{13}$$

From equation 13, it is evident that the view planning problem is isomorphic to a set covering problem and is therefore NP-complete, as has been observed by [10]. Consequently, an optimal solution is achievable only by an exhaustive search, which is impractical. This motivates the creation of practical, sub-optimal NBV algorithms that satisfy a set of given reconstruction objectives.

The measurability mapping $\mathbf{v}_j \xrightarrow{\mathcal{M}} S_j$ embeds all information essential to solve the view planning task. The issue is how to efficiently represent this construct in computational terms, noting the complexity inherent in a $\Re^{6+} \rightarrow \Re^2$ mapping[2]. Contrasting inspection, in reconstruction the object surface S is not only unknown but is the task objective. Viewpoint space V is defined but it is unclear how best to partition it. These are some of the issues to be addressed to apply the measurability matrix representation to object reconstruction.

4 Multi-Stage Model-Based View Planning

4.1 Overview

Current view planning methods for object reconstruction attempt to incrementally build a model following an iterative cycle. They take an image, acquire some new information, decide on the NBV and repeat the process. We propose a different approach, a *multi-stage model-based view planning strategy* of progressive model refinement. The process begins and ends with a closed object model - the first embodying the relevant and available a priori object knowledge; the last satisfying specified object reconstruction goals. One or more intermediate models represent points on the continuum of spatial resolution and geometric accuracy. Each stage of the process builds upon the global geometric knowledge of the previous stage until a model satisfying the specified requirements is achieved. Our approach is also incremental, but at the model rather than the image level. Within each level, our process is inherently parallel.

In addition to being multi-stage, the strategy applies a variety of techniques at various planning stages and utilizes the principle of least commitment with respect to the most difficult view planning problems. By design, we employ a divide and conquer strategy; endeavoring to scan most of the object surface using fast and simple techniques. The most powerful computationally heavy weight techniques are reserved for small sub-sets of the surface and of viewpoint space, when and if problems are encountered with simpler methods. Thus, the proposed view planning process is layered in two senses - first in the spatial resolution of the problem domain and secondly in the computational cost of the techniques used.

Although additional stages could conceivably be utilized, the present concept envisions a two step object reconstruction process, separating the exploration and precision measurement processes. Knowledge embedded in the *a priori model* is first used to guide fast data acquisition to create an intermediate *rough model* capturing the essence of the object's shape. The rough model is subsequently used to plan high precision, high resolution surface scanning to create the desired output model at the desired level of geometric resolution, the *fine model*. Only the fine model is retained, the rough model being discarded once it has served its purpose. Consequently, the proposed multi-stage model-based view planning strategy (Figure 6) involves two stages, each with an associated problem resolution sub-stage:

[2] Representative quantization levels for high quality reconstruction are $|S| \approx 10^5$ and $|V| \approx 10^{10}$.

- a rough modeling phase for scene exploration, and
- a fine modeling phase to precisely capture geometric detail.

A Priori Model At the minimum, the problem begins with some knowledge of the object's bounding dimensions, however approximate. This knowledge constitutes an initial object model, the a priori model. Experimentation will examine the utility of a modest increase in prior knowledge, such as the number and approximate location of holes and major cavities.

Rough Model The function of the rough modeling phase is to quickly and faithfully capture the object's gross topology and geometry. Spatial resolution and accuracy requirements at this stage are kept to the minimum necessary for it to serve as the basis of fine detail modeling in the next phase. The level of resolution required at the rough modeling stage is expected to depend on several factors, including the size of the final model and the object's topologic and geometric complexity. Data acquisition will be accelerated by acquiring sub-sampled sparse range images. Problem resolution requirements at the rough modeling stage will focus on ensuring the model is closed and has the correct gross topology and geometry.

Fine Model Given a good but approximate rough model, the function of the fine modeling phase will be to sample the surface to a fine level of resolution and high level of precision in compliance with the input specification. This stage will again attempt to rapidly scan all or most of the surface using computationally lightweight view planning techniques. Patches of the rough model for which these tactics fail will be passed to problem resolution modules employing more powerful but slower visibility analysis view planning techniques.

4.2 3M Algorithm

Overview We will now briefly describe the modified measurability matrix (3M) algorithm as applied at the fine modeling stage. The essence of the view planning problem is to efficiently gather sufficient information to encode the measurability mapping at a level of detail satisfactory for the next level of model building. Once that has been done, selecting the NBV is straight forward and fast.

Reflecting on the $\mathbf{v}_j \xrightarrow{\mathcal{M}} S_j$ measurability mapping and measurability matrix constructs presented earlier, we can observe that careful partitioning of S and V is both necessary and achievable. Fairly coarse quantization of S will be sufficient in the rough model for performance-oriented view planning at the next stage, provided the rough model correctly captures the object's gross topology and geometry. Secondly, complex object features such as holes, cavities and protrusions constrain views much more severely than planar or gently convex regions. Thirdly, the physics of the sensing process provides important clues as to optimal viewpoints relative to surface features. Finally, since large portions of the object are occluded from any given viewpoint, it is clearly necessary to compute only a comparatively sparse measurability matrix.

Consequently, we follow a generate and test procedure concentrating most of the effort on generation of a modest number of well-chosen viewpoints. In outline form, the 3M algorithm proceeds as follows at the fine modeling stage:

```
Segment the rough model by surface feature type
While ( Unscanned rough model regions)
  Generate optimal candidate viewpoint set
  Compute measurability; form measurability matrix
  Select NBV(s)
```

Rough Model Segmentation The rough model is segmented into the following surface feature types: cavity, hole or planar/convex patch[3].

Viewpoint Generation For each surface patch, a set of candidate viewpoints is generated that is most likely to measure the surface feature in accordance with the specified criteria. Viewpoint generation is achieved by quantizing the optimal viewpoint zone for the specific sensor and surface feature geometry.

Measurability Estimation Using an appropriate sensor performance model, a measurability matrix is computed for each patch. This is done by calculating the estimated measurement precision and sampling density at each point on the targeted rough model surface patch from each candidate viewpoint in the corresponding viewpoint set. At this stage, measurability determination is based on local visibility analysis only, - that is, considering only the geometry of the target surface patch.

NBV Selection Following a simple set covering algorithm, one or more NBVs are selected as the viewpoint(s) which collectively best sample all of the surface patch within specification.

4.3 Comparison w. Current Methods

The 3M algorithm employs a separate measurability matrix tailored for each segmented rough model surface patch. This approach results in a number of measurability matrices, each comparatively small, for a given object reconstruction task. In contrast, the conventional measurability matrix involves a single, monolithic matrix covering all surface points and all candidate viewpoints. Figure 7 illustrates how the 3M algorithm results in a sharp net reduction in the size of the composite matrix.

[3]Protrusions are not dealt with as a separate type at this stage; rather, residual occlusions are dealt with at the problem resolution sub-stage.

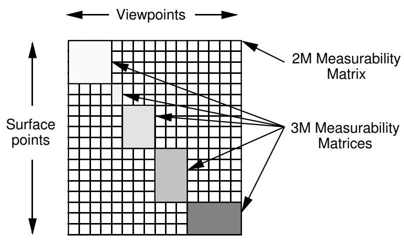

Figure 7: Measurability Matrix Comparison

While a more complete performance analysis awaits completion of development and experimental activity, a number of observations can be made at this time:

- Measurement performance should be superior to conventional methods because:
  - the process incorporates explicit sensor and positioning system performance models,
  - the fine modeling process has the advantage of approximate knowledge of object topology and geometry in the form of the rough model,
  - view planning is targeted and optimized for one specific surface feature at a time, and
  - candidate viewpoints are selected for high prospective viewing quality from a small portion of the imaging workspace.
- The algorithm should be fast relative to conventional methods because:
  - global occlusion analysis is employed, if at all, only for sub-sets of the problem domain, and
  - the domain of both generate and test operations on the 3M matrix is much smaller than conventional measurability matrices.

Efficiency is addressed by several means - problem decomposition inherent to the 3M algorithm, surface segmentation by major geometry feature, limited visibility analysis and generation of a limited set of near-optimum candidate viewpoints. Accuracy and robustness are achieved through modeling of the imaging environment, objective performance criteria and optimal viewpoint set generation.

5 Conclusion

This paper has described a view planning concept whose goal is the automatic creation of high quality geometric models of a single object. A review of current methods identified the main open problems as efficiency, accuracy and robustness. The importance of analyzing measurability versus visibility stressed the need for realistic modeling of imaging physics. A brief theoretical discussion framed the problem in terms of geometric relationships between the two dimensional surface space $S \subset \Re^2$ of the target object and a multi-dimensional imaging work space, or viewpoint space $V \subset \Re^{6+}$. Noting that all information required for view planning is embedded in the measurability mapping $\mathbf{v}_j \xrightarrow{\mathcal{M}} S_j$ and its inverse, led to the observation that the view planning problem is isomorphic to a set covering problem and therefore is NP-complete.

A new concept based on multi-stage, model-based view planning using a modified measurability matrix (3M) algorithm was outlined as a means of addressing the main outstanding open issues. At the time of submission, algorithms implementing the 3M view planning concept have been developed in detail and prototype software is near completion.

References

[1] J. E. Banta and M. A. Abidi. The positioning of a range sensor for optimal reconstruction of three-dimensional models. In *Int. Conf. on Recent Advances in 3-D Digital Imaging and Modeling, Ottawa*, May 12-15 1997.

[2] C. Connolly. The determination of next best views. In *IEEE Int. Conf. on Robotics and Automation*, pages 432–435, 1985.

[3] D. Lamb, D. Baird, and M. Greenspan. An automation system for industrial 3-d laser digitizing. In *Second Int. Conf. on 3-D Digital Imaging and Modeling, Ottawa*, pages 148–157, October 4-8 1999.

[4] N. A. Massios and R. B. Fisher. A best next view selection algorithm incorporating a quality criterion. In *British Machine Vision Conference 1998*, pages 780–789, September, 1998.

[5] J. Maver and R. Bajcsy. Occlusions as a guide for planning the next view. *IEEE Trans. PAMI*, 17(5):417–433, May 1993.

[6] D. Papadopoulos-Orfanos and F. Schmitt. Automatic 3-d digitization using a laser rangefinder with a small field of view. In *Int. Conf. on Recent Advances in 3-D Digital Imaging and Modeling, Ottawa*, pages 60–67, May 12-15 1997.

[7] R. Pito. A sensor based solution to the next-best-view problem. In *IEEE Int. Conf. on Robotics and Automation*, pages 941–945, August 1996.

[8] C. J. Pudney. *Surface Modelling and Surface Following for Robots Equipped with Range Sensors.* PhD thesis, Univ. of Western Australia, Perth, 1994.

[9] M. Reed, P. Allen, and I. Stamos. 3-d modeling from range imagery: An incremental method with a planning component. In *Int. Conf. on Recent Advances in 3-D Digital Imaging and Modeling, Ottawa*, pages 76–83, May 12-15 1997.

[10] G. Tarbox and S. Gottschlich. Planning for complete sensor coverage in inspection. *Computer Vision and Image Understanding*, 61(1):84–111, January 1995.

[11] P. Whaite and F. P. Ferrie. Autonomous exploration: Driven by uncertainty. *IEEE Trans. PAMI*, 19(3):193–205, March 1997.

[12] X. Yuan. A mechanism of automatic 3d object modeling. *IEEE Trans. PAMI*, 17(3):307–311, March 1995.